



FACULTY OF **BIOSCIENCE ENGINEERING**

Department of Forest and Water Management

Assessment of uncertainty in land surface characterization and implications for spatially distributed hydrological modelling

ir. Lien Loosvelt

Thesis submitted in fulfillment of the requirements for the degree of Doctor (Ph.D.) of Applied Biological Sciences

Academic year 2013-2014

Supervisor: Prof. dr. ir. Niko Verhoest

It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of light, it was the season of darkness, it was the spring of hope, it was the winter of despair.

A Tale of Two Cities (1859)
Charles Dickens

Supervisor:	Prof. dr. ir. Niko Verhoest Department of Forest and Water Management Ghent University, Belgium
Examination committee:	Prof. dr. ir. Filip Tack
	Prof. dr. Bernard De Baets Prof. dr. ir. Ann Van Griensven Prof. dr. ir. Gerrit Schoups Prof. dr. ir. Wim Cornelis dr. ir. Frieke Van Coillie
Dean:	Prof. dr. ir. Guido Van Huylenbroeck
Rector:	Prof. dr. Anne De Paepe

ir. Lien Loosvelt

Assessment of uncertainty in land surface characterization and implications for spatially distributed hydrological modelling

Thesis submitted in fulfillment of the requirements for the degree of

Doctor (Ph.D) of Applied Biological Sciences

Academic year 2013-2014

Dutch translation of the title:
Begroten van onzekerheid in de beschrijving van het landoppervlak en implicaties voor het ruimtelijk verdeeld hydrologisch modelleren

Please refer to this work as follows:
Lien Loosvelt (**2013**). *Assessment of uncertainty in land surface characterization and implications for spatially distributed hydrological modelling*, PhD Thesis, Department of Forest and Water Management, Ghent University, Ghent, Belgium.

©Photograph on cover by Lien Loosvelt
Chile, 2006

ISBN 978-90-5989-662-8

The author and the supervisor give the authorization to consult and to copy parts of this work for personal use only. Every other use is subject to the copyright laws. Permission to reproduce any material contained in this work should be obtained from the author.

Acknowledgements - Dankwoord

Op dit moment realiseer ik me dat ik tijdens de afgelopen zes jaar van mijn doctoraat te weinig de mensen rondom mij bedankt heb voor de steun die ze mij gegeven hebben. Dit dankwoord wil ik aangrijpen om te zeggen hoezeer ik hun steun en aanwezigheid geapprecieerd heb.

Hoe kan ik anders dan mijn promotor Niko heel hartelijk te bedanken om mij deze job, inclusief zijn vertrouwen, te geven. Van hem heb ik de voorbije jaren enorm veel bijgeleerd en heel veel input gekregen. Na elk gesprek draaiden mijn hersenen weer op volle toeren. Ook zijn vrouw Hilde wil ik bedanken voor de leuke samenwerking en de koffiekletsjes. Zes jaar later ben ik naast ouder ook wijzer geworden en heb ik geleerd dat doctoreren echt wel een 'philosophical degree' is, een levensles. Ook van mijn collega's heb ik veel bijgeleerd, namelijk dat alcohol niet nefast is om een goed denkvermogen te ontwikkelen en dat vriendschap op het werk heel waardevol is. Dus bij deze wil ik al mijn collega's en ex-collega's bedanken. Het is een enorm voorrecht om met zo'n toffe mensen te mogen samen werken. Jasper heeft mij vier jaar gezelschap gehouden, Marcella was er altijd als niemand anders er was, Davy heeft mij de kunst van het genieten bijgebracht. Speciaal wil ik ook Valentijn bedanken om (zelfs vanuit Australië) mijn programmeerproblemen op te lossen. Bedankt ook aan Bernard voor het vele leeswerk en de feedback.

Verder rest mij de mensen te bedanken die nog niet van mij verlost zijn, namelijk familie en vrienden. Allen hebben ze mij leren relativeren, want wat je ook bereikt in het leven, geluk moet je kunnen delen. Mijn vrienden hebben de weekends doctoraatsvrij gemaakt en veel mooie momenten gecreëerd met huwelijken, geboortes, stuif-ins en 'candle-light-dinners'. In het bijzonder wil ik Jan bedanken omdat hij mij overtuigd heeft om te doctoreren. Mijn familie is voor mij het allerbelangrijkste geweest: mijn grootouders hebben de gunstige maanfasen opgevolgd, mijn zus heeft mij behoed voor een saai leven en mij overtuigd (eerder gesmeekt) om Doetsie te adopteren, en mijn broers zijn vooral zichzelf gebleven want ik kan me geen betere wensen. Speciale dank gaat uiteraard uit naar mijn ouders: dank aan 'hotel mama' waar ik de voorbije zes jaar kosteloos heb mogen verblijven, dank aan 'klusser papa' die de voorbije vier jaar elk weekend heeft helpen verbouwen. Een dikke merci uit de grond van mijn hart!!!

Lien Loosvelt
Gent, 28 november 2013

Contents

Acknowledgements - Dankwoord vii

Dutch summary - Samenvatting 1

Summary 7

1 Introduction 13
- 1.1 History of hydrological modelling . . . 13
- 1.2 Uncertainty issues in hydrological modelling . . . 15
- 1.3 Research objectives, questions and overview . . . 17

I General concepts of uncertainty and sensitivity analysis 21

2 Introduction 23

3 Uncertainty analysis 25
- 3.1 Uncertainty identification . . . 25
  - 3.1.1 Types of uncertainty . . . 26
  - 3.1.2 Sources of uncertainty . . . 27
- 3.2 Uncertainty representation . . . 29
  - 3.2.1 Probability theory . . . 30
  - 3.2.2 Fuzzy set theory . . . 32
  - 3.2.3 Possibility theory . . . 33
- 3.3 Uncertainty propagation . . . 36
  - 3.3.1 Sampling methods . . . 37
  - 3.3.2 Approximation methods . . . 41
  - 3.3.3 Extension principle . . . 42

4 Sensitivity analysis 45
- 4.1 Derivative-based methods . . . 46
  - 4.1.1 Selecting the perturbation factor . . . 48
  - 4.1.2 One-at-a-time paradox . . . 49
- 4.2 Regression-based methods . . . 50
- 4.3 Variance-based methods . . . 52

II Distributed modelling of catchment hydrology 55

5 Introduction 57
5.1 Classification of hydrological models 57
5.2 Operation of Land Surface Models 60

6 Hydrological model description 63
6.1 Model structure 63
6.1.1 Local water balance equations 64
6.1.2 Local energy balance equations 68
6.1.3 The TOPMODEL 69
6.1.4 Soil description 70
6.2 Model operation 70
6.3 Inputs and parameters 72
6.4 Runoff routing 75

7 Catchment and data description 77
7.1 Geographical data 78
7.1.1 Digital elevation model 78
7.1.2 Soil maps 78
7.1.3 Land cover maps 80
7.2 Remote sensing data 81
7.3 Meteorological data 82
7.4 Soil moisture and discharge measurements 83

8 Calibration of the hydrological model 85
8.1 Calibration methods 85
8.1.1 Shuffled complex evolution 86
8.1.2 Dynamically dimensioned search 87
8.1.3 Weight-adaptive recursive parameter estimation 88
8.2 Hydrological model setup 89
8.3 Results and discussion 90
8.3.1 Calibration of the TOPMODEL parameters 91
8.3.2 Calibration of the routing parameters 94
8.3.3 Validation of the calibrated model 95
8.4 Conclusion 97

III Impact of uncertain land cover input on hydrological model predictions 99

9 Introduction 101

10 Effect of land cover generalization on the prediction of water fluxes 103

10.1 Introduction . . . 103
10.2 Techniques for thematic map comparison . . . 105
10.2.1 Cross-walking . . . 105
10.2.2 Average mutual information . . . 106
10.2.3 Fuzzy agreement . . . 107
10.3 Hydrological model set-up . . . 108
10.4 Results and discussion . . . 109
10.4.1 Comparison of land cover information . . . 109
10.4.2 Comparison of predicted water fluxes . . . 114
10.5 Conclusion . . . 117

11 Evaluation of uncertainty in land cover from image classification 119
11.1 Introduction . . . 119
11.2 Study site and data . . . 123
11.2.1 SAR imagery . . . 123
11.2.2 Ground measurements . . . 124
11.3 Methodology . . . 126
11.3.1 Random Forests . . . 126
11.3.2 Model evaluation . . . 130
11.3.3 Definition of polarimetric features . . . 134
11.3.4 Feature reduction . . . 138
11.4 Effect of multi-configuration on the classification quality . . . 140
11.4.1 Map quality analysis . . . 140
11.4.2 Effect of image frequency . . . 148
11.4.3 Effect of image acquisition date . . . 148
11.4.4 Spatial analysis of uncertainty . . . 149
11.5 Effect of feature reduction on the classification quality . . . 151
11.5.1 Feature importance analysis . . . 151
11.5.2 Feature reduction . . . 153
11.5.3 Map quality analysis . . . 157
11.6 Conclusion . . . 166

12 Impact of classification uncertainty on the prediction of water fluxes 169
12.1 Introduction . . . 169
12.2 Methodology . . . 171
12.2.1 Hydrological model set-up . . . 171
12.2.2 Uncertainty analysis . . . 172
12.2.3 Sensitivity analysis . . . 175
12.3 Results and discussion . . . 177
12.3.1 Map quality analysis . . . 177
12.3.2 Effect of uncertain land cover on water flux predictions . . . 180
12.3.3 Fast indicative uncertainty assessment . . . 184

12.3.4 Sensitivity of model predictions to land cover parameters . 189
12.4 Conclusion 192

IV Impact of uncertain soil input on hydrological model predictions 195

13 Introduction 197

14 Impact of soil parameter uncertainty on the prediction of soil moisture 199
14.1 Introduction 199
14.2 Methodology 201
14.2.1 Hydrological model set-up 201
14.2.2 Prediction of the soil hydraulic parameters 202
14.2.3 Ensemble predicting 205
14.3 Results and discussion 208
14.3.1 Within-soil-class variability of the SHPs 208
14.3.2 Uncertainty in the soil moisture predictions 214
14.3.3 Implications for data assimilation 221
14.3.4 Evaluation of the central point as a proxy for the soil class 222
14.4 Conclusion 227

15 Sensitivity of soil moisture to changes in soil texture 229
15.1 Introduction 229
15.2 Hydrological model set-up 231
15.3 Compositional data 232
15.3.1 Basic concept and operations 232
15.3.2 Soil texture in the simplex 233
15.4 Local sensitivity analysis on compositional data 234
15.4.1 Perturbing in the 2D Euclidean space 237
15.4.2 Perturbing in the simplex 237
15.4.3 Calculating the sensitivity index 241
15.4.4 Optimizing the perturbation factor 244
15.5 Results and discussion 244
15.5.1 Identification of the optimal perturbation factor 244
15.5.2 Identification of sensitive regions in the texture triangle . . 246
15.5.3 Identification of the hydrological model behaviour 252
15.6 Conclusion 255

General conclusions and outlook 257
Contribution of the thesis 257
Answer to the research questions 258
Recommendations for hydrological modelling 262

Future perspectives 263

List of Figures

2.1 General scheme that illustrates the establishment of model prediction uncertainty arising from uncertainty in the description of the real world and in the model input data (modified from Wagener and Gupta (2005)). 23
2.2 General scheme of uncertainty analysis and sensitivity analysis with a focus on input uncertainty (modified from Crosetto et al. (2000)). 24

3.1 Example of a probability density function (PDF) (left) and a cumulative distribution function (CDF) (right), with indication of a $a \cdot 100\%$ confidence interval $[u_1, u_2]$. 31
3.2 Example of a trapezoidal (left) and triangular (right) membership function $A(x)$ that maps every element of the universe to the interval [0,1]. 33
3.3 Example of a possibility distribution with indication of the core and the support of the distribution. 34
3.4 Illustration of the possibility Π and the necessity N for the event $X \leq u$. 35
3.5 Schematic of a Monte Carlo analysis of a model with two input variables X_1 and X_2. 38
3.6 Illustration of the FOSM method for propagating uncertainty in a one-dimensional problem. 42
3.7 Illustration of the extension principle to propagate uncertainty in a two-dimensional problem. 43

4.1 In two dimensions, the OAT partially explores a circle by moving along the abscis or the ordinate, while the entire sample space is given by the full square. Moving from the origin along the abscis, returning to the origin and moving along the ordinate does not allow to move outside the circle. 50
4.2 Illustration of the curse of dimensionality. The percentage of the explored space by OAT, relative to the entire sample space goes rapidly to zero as the number of dimensions increases. 51

5.1 Criteria-based classification scheme for hydrological models (modified from Jajarmizadeh et al. (2012)). The five criteria used are (i) laws and assumption, (ii) equations, (iii) parameters, (iv) time and (v) procedure of computation. 58
5.2 Illustration of spatial discretization of a catchment with regard to the application of a distributed model. 60

5.3 Schematic of the operation of a land surface model. 61

6.1 Schematic representation of the water budgets (indicated with filled arrows) and the energy budgets (indicated with open arrows) in the local SVATS (*i.e.* for one computational unit, either a pixel or an interval) within the hydrological model TOPLATS. 65

7.1 Situation of the Bellebeek catchment in Belgium with indication of the river network and the location of the measurement stations. . . 77
7.2 USDA soil classification triangle (top) and Belgian soil classification triangle (bottom). 79
7.3 FAO soil map STM-FAO (left) and Belgian soil map STM-BEL (right) for the Bellebeek catchment. The class labels in the legends are explained in Table 7.1. 80
7.4 General land cover map LCM-GEN (left) and detailed land cover map LCM-CROP with indication of the crop type (right) for the Bellebeek catchment. The class labels from the legend are explained in Table 7.2. 82

8.1 Timeseries of the observed discharge and the simulated discharge for the year 2006 obtained with the calibrated TOPMODEL parameters (TOPLATS run in statistical mode) as obtained with the different calibration methods: (a) SCE, (b) DDS, (c) WARPE. 92
8.2 Scatterplot between the observed and simulated discharge for the year 2006 as obtained with the calibrated TOPMODEL parameters from the SCE (a), DDS (b) and WARPE method (c) (TOPLATS run in statistical mode). 93
8.3 The runoff routing functions for the Bellebeek catchment: the normalized width function (a), the normalized hillslope function (b) and the resulting unit hydrograph (c). 96
8.4 Time series of the observed discharge and the simulated discharge with the calibrated TOPMODEL and routing parameter values for the year 2007 (TOPLATS run in distributed mode). 97
8.5 Scatterplot between the observed and simulated discharge for the validation period (year 2007) as obtained with the calibrated TOPMODEL and routing parameters from the WARPE method (TOPLATS run in distributed mode). 97

9.1 A schematic of the different uncertain situations w.r.t. land cover information to model the water fluxes for a catchment. In situation 1, a land cover map of the basin is readily available from different sources: a general land cover map LCM-GEN and a detailed land cover map LCM-CROP, in situation 2 the land cover map needs to be derived through classification of remotely sensed imagery. . . . 102

10.1 Different levels of generalization in a land cover classification scheme (modified from Foody and Atkinson (2002)). 108
10.2 Fuzzy agreement map (left) and thematic uncertainty map (right). 114
10.3 Simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b for the Bellebeek catchment during 2006, using two different land cover maps LCM-GEN and LCM-CROP. 116
10.4 Uncertainty map of the simulated evapotranspiration as a result of generalization in the land cover information by replacing land cover map LCM-CROP with LCM-GEN. 117

11.1 Aerial photograph of the Foulum test site in Denmark. 123
11.2 Crop development stage according to the BBCH scale for the individual crop types as a function of time. The first digit of the BBCH scale refers to the principal growth stage, the second digit to the secondary growth stage. BBCH values were linearly interpolated between the dates of field survey. The principal growth stages may be summarized as: (0) germination, sprouting, bud development, (1) leaf development, (2) formation of side shoots/tillering, (3) stem elongation or rosette growth, shoot development, (4) development of harvestable vegetative plant parts or vegetatively propagated organs, booting, (5) inflorescence emergence, heading, (6) flowering, (7) development of fruit, (8) ripening or maturity of fruit and seed, and (9) senescence, beginning of dormancy. For detailed information on the scaling system, the reader is referred to Hack et al. (1992) and Meier (2001). 125
11.3 Schematic overview of the construction of a Random Forest (full arrows) and of the calculation of the model error (dotted arrows) based on respectively the bootstrap subsets and the out-of-bag subsets sampled from data set X. c_l indicates the number of correctly classified pixels by the l-th tree. 128
11.4 Triangle plot of uncertainty showing (a) $U = 1 - p_{\max}$ and (b) entropy H of pixel u as a function of probabilities $p(1)$, $p(2)$ and $p(3)$ (three-class example) from the probability vector $\mathbf{p}_u$. The linear behaviour of U contrasts with the nonlinear behaviour of H. 134
11.5 Grid representation of the Spearman correlation coefficient between the 20 polarimetric SAR features of the L-band acquisition on May 20. The feature numbers corresponds to the features as listed in Table 11.3 138
11.6 Cumulative relative frequency distribution of $U = 1 - p_{\max}$ and entropy H for correctly classified and incorrectly classified pixels by the different classification models. Results for classification models applying C-band imagery are given in dashed lines, results for models applying L-band imagery are given in solid lines. 144

11.7 (a) Uncertainty map of the multi-date classification model in L-band using entropy (H). (b) Scatterplot of entropy values for inner-field pixels and boundary pixels for the 15 different classification models averaged over the different reference fields for each crop separately. Different crops have different scatter symbols, whereas different models have different colours (for the colour legend, *cfr.* Figure 11.6). (c) The within-field variability of uncertainty values reflects the agricultural practices for some crops, e.g. winter wheat. 152
11.8 Importance score (calculated following Algorithm 3) of the 20 features derived from L-band single-date SAR imagery (acquisitions on April 17, May 20, June 15 and July 16 in 1998) used to classify the land cover at the Foulum test site (Denmark), averaged over 10 model runs (with standard deviation). The numbers on the abscis correspond to specific features from Table 11.3. 154
11.9 Overall accuracy, Cohen's kappa and $\mathrm{AUC}_{\mathrm{avg}}$ as a function of the number of features removed from the single-date data sets. Feature removal is based on the feature importance ranking per date, in which the removed feature is not necessarily the same on all dates. 155
11.10 Importance score (averaged over 10 model runs with indication of standard deviation, and calculated following Algorithm 3) for the classification on the 8 most important features per date (April 17, May 20, June 15 and July 16). The numbers on the abscis correspond to specific features from Table 11.3. 157
11.11 Average importance score $C_{j,\mathrm{avg}}$ (error bars indicate the standard deviation) of the input variables over all acquisition dates (April 17, May 20, June 15, July 16). The numbers on the abscis correspond to specific features from Table 11.3. 158
11.12 Overall accuracy, Cohen's kappa and $\mathrm{AUC}_{\mathrm{avg}}$ as a function of the number of features removed from the single-date data sets. The feature removal is based on the average importance score $C_{j,\mathrm{avg}}$ over all dates (*cfr.* Figure 11.11) with the removed feature being the same on each acquisition date. The numbers on the abscis correspond to specific features from Table 11.3. 159
11.13 ROC curve for the prediction of peas and winter wheat with the complex model (based on 20 features) and the simple model (based on 1 feature), the 0-1 line indicates the curve corresponding with random guessing. 163
11.14 Frequency distribution of $U = 1 - p_{\max}$ (a) and the Shannon entropy H (b) for the correct predictions as resulted from the complex, the optimal, the practical and the simple model. 164
11.15 Frequency distribution of $U = 1 - p_{\max}$ (a) and the Shannon entropy H (b) for the incorrect predictions as resulted from the complex, the optimal, the pratical and the simple model. 165

12.1 Schematic of the uncertainty analysis, including an uncertainty modelling step and an uncertainty propagation step. 173
12.2 Pixel-based (spatially independent) simulation approach (a) and field-based simulation approach (b) to generate error-corrupted versions of the land cover map (modified from Canters et al. (2002)). 176
12.3 Hard classified land cover map of the Bellebeek (LCM-RF$_{\text{hard}}$) as obtained through Random Forests (a) and map of the entropy (uncertainty) in the predicted land cover (b). 178
12.4 Spaghetti plot of the simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b for the Bellebeek catchment during 2006, as resulted from 100 model applications with error-corrupted versions of LCM-RF. The black line indicates the result obtained with LCM-RF$_{\text{hard}}$. 181
12.5 Temporal variability of uncertainty in the simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b (for the Bellebeek catchment during the year 2006) due to classification uncertainty. . 182
12.6 Mean [mm] (a) and standard deviation [mm] (b) of the simulated annual evapotranspiration after propagation of the classification uncertainty. 183
12.7 Evaluation of the predicted ET with the hard classified LCM (ET_{RF}) and the entropy H in the land cover classification as an indicator for respectively the mean (left) and the standard deviation (right) on ET from the Monte Carlo simulation (ET_{MC}). 185
12.8 Mean (left) and standard deviation (right) of the annual evapotranspiration [mm] as resulted from the fast indicative uncertainty assessment based on uncertainty proxies. 186
12.9 Predicted evapotranspiration (ET_{RF}) with the hard classified land cover map as a function of the topindex λ_T. 187
12.10 Evaluation of the FI-UA based on uncertainty proxies (ET_{proxy}) to approximate the mean (left) and standard deviation (right) of the simulated evapotranspiration (ET_{MC}) that resulted from the Monte Carlo-based uncertainty analysis (MC-UA). 189
12.11 Sensitivity of the simulated evapotranspiration, runoff and baseflow to the time-varying vegetation parameters as a function of the month of the year. 190

13.1 Flowchart for simulating the soil moisture in case the soil hydraulic parameters (SHPs) need to be derived from soil textural information. 197
13.2 A schematic of the different uncertain situations w.r.t. soil information to model the soil moisture for a catchment. In situation 1, soil texture information is available from a soil type map, in situation 2 soil texture information is available from a soil type map and some texture measurements. 198

14.1 USDA texture triangle with indication of the boundary conditions on the PTFs of Rawls and Brakensiek (1985, 1989) (grey full line) and on the PTFs of Vereecken et al. (1989, 1990) (grey dashed line). The area of the particle size distributions considered in the study is shaded and the central point per soil class is indicated with a grey dot. 203

14.2 Top panel: contour plot of the logarithm ($\log_{10}$) of the saturated hydraulic conductivity K_s [$\mathrm{m} \cdot \mathrm{day}^{-1}$] predicted with PTF A (a) and PTF B (b). Bottom panel: contour plot of the bubbling pressure ψ_c [m] predicted with PTF A (c) and PTF B (d). The soil class clay loam is highlighted in grey. 210

14.3 Probability distribution of (a) the saturated hydraulic conductivity, K_s [$\mathrm{m} \cdot \mathrm{day}^{-1}$] and (b) the bubbling pressure ψ_c [m] predicted with PTF A (grey) and PTF B (white) for the soil class clay loam. . . . 211

14.4 Spaghetti plots of the daily simulated soil moisture content (2006), θ_{daily} [$\mathrm{m}^3 \cdot \mathrm{m}^{-3}$], for sandy loam (SL) and silty clay (SiCl) with the SHPs predicted from PTF A and PTF B. The ensemble mean is indicated with black dots. 215

14.5 Left panel: ensemble (328 members) of soil moisture retention curves (a) and hydraulic conductivity curves (c) for the soil class sandy loam resulting from application of PTF A (dark grey) and PTF B (light grey). Right panel: ensemble (106 members) of soil moisture retention curves (b) and hydraulic conductivity curves (d) for the soil class silty clay resulting from application of PTF A (dark grey) and PTF B (light grey). The ensemble mean is indicated with black dots. 216

14.6 Contour plot of the average simulated soil moisture content, θ_{avg} [$\mathrm{m}^3 \cdot \mathrm{m}^{-3}$], using the SHPs as predicted from PTF A (a) and PTF B (b). The soil class clay loam is highlighted in grey. 217

14.7 Time series of the ensemble mean, ensemble spread, ensemble skewness and ensemble kurtosis for the simulated soil moisture content (2006), θ_{daily} [$\mathrm{m}^3 \cdot \mathrm{m}^{-3}$]. The results are given for the soil classes sandy loam (black) and silty clay (grey) using the SHPs as predicted from PTF A (solid line) and PTF B (dotted line). 218

14.8 Probability distribution of the average simulated soil moisture content, θ_{avg} [$\mathrm{m}^3 \cdot \mathrm{m}^{-3}$], using the SHPs as predicted from PTF A (grey) and PTF B (white) for the USDA class 'clay loam'. 221

14.9 Probability distribution of the simulated soil moisture content, θ_{daily} [$\mathrm{m}^3 \cdot \mathrm{m}^{-3}$] using the SHPs predicted from PTF A (grey) and PTF B (white), in a silty clay soil on a wet day (DOY 215) (a) and on a dry day (DOY 260) (b). The dotted black line shows the corresponding normal distribution. 222

14.10 Range of the SHPs and θ_{avg} within each USDA class as predicted from PTF A (solid line) and PTF B (dotted line) with indication of the SHP for the central point (CP). The results are compared to standard SHPs as reported by Rawls et al. (1982) and Meyer et al. (1997). 225

14.11 Time series of the ensemble mean of the simulated soil moisture content, θ_{daily} [$m^3 \cdot m^{-3}$] resulting from PTF A (a) and PTF B (b) for the soil class clay loam with indication of the range in model output for comparison with the model run for the central point and with the SHPs taken from by Rawls et al. (1982) and Meyer et al. (1997). 226

15.1 Representation of a sample $\mathbf{x} = (C, Z, L)$ in the texture triangle with indication of the bisectors B_1, B_2 and B_3, the vertices $\mathbf{p_1}$, $\mathbf{p_2}$, $\mathbf{p_3}$ and the boundary conditions on the PTFs of Rawls and Brakensiek (1985, 1989). 235

15.2 Rescaling a sampled soil texture $\mathbf{x_0}$ (indicated with a triangle) with a factor ω, varying between -10 and 10, following the operation $\omega \odot \mathbf{x_0}$. 235

15.3 Adding to a sampled soil texture $\mathbf{x_0}$ (indicated with a triangle) (a) a set of other textures situated on the bisectors of the texture triangle and obtained after scaling texture $\mathbf{x}'$ (indicated with squares), and (b) a set of other textures situated in the neigbourhood of the sampled texture with the same direction of $\mathbf{x}'$ and obtained after scaling $\mathbf{x_0}$ (indicated with squares). 236

15.4 (a) Representation of the baricenter $\mathbf{p_0}$, the sampled composition $\mathbf{x}$ and the bisectors B_1, B_2 and B_3 after ILR transformation and (b) illustration of perturbation in the 2D Euclidean space with indication of the perturbed compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ and the translated bisectors B_1', B_2' and B_3'. 238

15.5 Perturbation in the simplex of a composition $\mathbf{x}$ sampled from the texture triangle with indication of the directional vectors $\{\mathbf{v_{i+}}, \mathbf{v_{i-}}\}$, the perturbed compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$, the bisectors B_1, B_2 and B_3 and the translated bisectors B_1', B_2' and B_3'. Compositional circles at different locations in the texture triangle are shown for illustration. 240

15.6 The mean (full line), minimum and maximum (dotted lines) values of $C_{\max}$ as a function of the perturbation factor ξ (a) and the frequency distribution of the optimal perturbation factor (b), for 100 sampled compositions. 247

15.7 Distribution of the optimal perturbation factor ξ (100 samples) in the texture triangle with delineation of the USDA soil class (the corresponding class labels can be found in Figure 7.2). 248

15.8 Contour plot of the sensitivity index across the texture triangle for the estimated soil hydraulic parameters (a)-(e) θ_s (mean is 0.18 $m^3 \cdot m^{-3}$), θ_r (mean is 0.03 $m^3 \cdot m^{-3}$), ψ_c (0.16 m), λ (mean is 0.49), and $\log_{10} K_s$ (mean is $-6.11\,m \cdot s^{-1}$) and for the simulated soil moisture content θ (f). Results outside the validity zone of the PTFs (indicated by a white line) and near the borders of the triangle should be interpreted with care. 250
15.9 Average sensitivity index $\overline{S}_\theta$ for the 12 USDA soil classes with indication of the standard deviation and the sensitivity classes low ($0 \leq \overline{S}_\theta < 0.04$), medium ($0.04 \leq \overline{S}_\theta < 0.08$), high ($0.08 \leq \overline{S}_\theta < 0.12$) and very high ($0.12 \leq \overline{S}_\theta$). 253
15.10 Soil moisture retention curve (SMRC) for which the sensitivity of the simulated soil moisture to textural changes is the highest and the lowest, respectively. 253
15.11 Scatterplot between S_θ and the soil hydraulic parameters θ_s, θ_r, ψ_c, λ, K_s (a)-(e) and the average simulated soil moisture content θ_{avg} (f) (5000 sampled textures). 254

List of Tables

6.1 Meteorological inputs (time-varying) in the hydrological model. . . 73
6.2 Soil parameters (time-invariant) from the soil type look-up table (ST-LUT) in the hydrological model. 73
6.3 TOPMODEL parameters (time-invariant) from the catchment look-up table (CA-LUT) in the hydrological model. 74
6.4 Vegetation parameters from the land cover look-up table (LC-LUT) in the hydrological model. 74

7.1 Overview of the different soil types in the soil type maps (STMs) for the Bellebeek catchment. 80
7.2 Overview of the different land cover types in the available land cover maps (LCMs) for the Bellebeek catchment. 81

8.1 Initial, minimum, maximum and calibrated value for the TOPMODEL parameters as obtained with the different calibration algorithms. 94
8.2 Statistics of the calibrated discharge as obtained with the different calibration algorithms. 94
8.3 Initial, minimum, maximum and calibrated value for the routing parameters as obtained with the WARPE method. 95

10.1 Cross-tabulation between land cover maps LCM-GEN and LCM-CROP. Explanation of the class labels can be found in Table 7.2. . 110
10.2 Change in entropy of LCM-CROP (map X_2) given the class label information in LCM-GEN (map X_1). 110
10.3 Legend lookup table to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2. 111
10.4 Overlap matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2. . . . 111
10.5 Importance of disagreement matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2. 112
10.6 Fuzzy agreement matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2. 112
10.7 Number of pixels showing no, partial and complete uncertainty together with the average uncertainty in function of the map type. 114
10.8 Number of pixels showing no, partial and full agreement together with the average agreement in function of the map type. 114

10.9 Comparison between LCM-GEN and LCM-CROP in terms of water flux predictions and uncertainty ratio. 115

11.1 Overview of the number of fields (n) surveyed per crop type and the average observed crop height (h) per crop type together with the standard deviation. 126

11.2 Confusion matrix with indication of the true positive (TP), false positive (FP), false negative (FN) and true negative (TN) elements. 132

11.3 Overview of the 20 polarimetric SAR features used for land cover classification. 137

11.4 Spearman correlation coefficients r between γ_{hv} and γ_{ll}. Values of r lower than 0.5 are in italic and values of r higher than 0.8 are in boldface. 139

11.5 Overall classification accuracy (OA), model user's accuracy (MA), producer's accuracy (PA) and Cohen's κ for single configuration and multi-configuration classifications. The MA and PA are summarized by their range over the different land cover classes. The last column shows the results from a pairwise classification comparison by the McNemar test at the 0.05 significance level. Different numbers indicate significantly different performances, and models are ranked from the best performing classification model (1) to the worst performing model (11). 142

11.6 Class-specific user's accuracies (CA) resulted from the single-configuration classifications. The CA is indicated with -, + or ++ if its value exceeded 0.25, 0.50 or 0.75, respectively. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Be), peas (Pe), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww). 143

11.7 Median values of $U = 1 - p_{\max}$ for the single-configuration and the multi-configuration classifications for each of the land cover classes separately. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Be), peas (Pe), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww). For each classification model the minimum value of U among the agricultural crops is in boldface, the maximum value of U is underlined. 146

11.8 Median entropy values for the single-configuration and the multi-configuration classifications for each of the land cover classes separately. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Br), peas (Pr), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww). For each classification model the minimum entropy value among the agricultural crops is in boldface, the maximum entropy value is underlined (excluding classes Wa, Fo and Gr). 147

11.9 Overview of the eight SAR features with the highest permutation importance C_j (ranked from 1 to 8 with decreasing C_j) on each of the four acquisition dates. 156

11.10 Accuracy measures of four multi-date models, representing a different degree of dimensionality (the complex, optimal, practical and simple model are based on 20, 8, 3 and 1 features, respectively). 160

11.11 Class-specific accuracies (CA) of four multi-date models, representing a different degree of dimensionality (the complex, optimal, practical and simple model are based on 20, 8, 3 and 1 features, respectively). 160

11.12 Confusion matrix as resulted from the simple model (based on γ_{hv}). The main confusions are indicated in bold. 161

11.13 Comparison of the complex model (based on 20 features) with the optimal (based on 8 features), the practical (based on 3 features) and the simple model (based on 1 feature) by means of the McNemar test: y, significant difference in performance between the models, n, no significant difference, both at the 0.05 significance level; n_{01} and n_{10} are the error rates from the contingency table. 162

11.14 Mean uncertainty ($\overline{U}$) and mean entropy ($\overline{H}$) of the correct and incorrect predictions, as resulted from the complex (based on 20 features), optimal (based on 8 features), practical (based on 3 features) and simple (based on 1 feature) multi-date classification model. . . 162

12.1 Cover-dependent and time-varying (updated at a monthly time step) vegetation parameters for the different vegetation types in the TOPLATS simulations. For the leaf area index LAI, albedo α, momentum roughness $z_{0,m}$, heat roughness $z_{0,h}$ and zero plane displacement d_0, the minimum and maximum value within the growing season are listed. 172

12.2 Entropy ($\overline{H}$) and areal cover percentage (a) per land cover class in the classified land cover map LCM-RF. The land cover classes include urban area (Ur), water area (Wa), broad-leaved forest (Bf), needle-leaved forest (Nf), pasture (Pa) and 7 specific classes: grass (Gr), potatoes (Po), beets (Be), maize (Ma), wheat (Wh), barley (Ba) and other crop species (Oc). 179

12.3 Average membership probability vector $\overline{p}$ for the different land covers in the study area, as obtained from the soft classified LCM-RF. The maximum probability is indicated in bold. The land cover classes include urban area (Ur), water area (Wa), broad-leaved forest (Bf), needle-leaved forest (Nf), pasture (Pa) and 7 specific classes: grass (Gr), potatoes (Po), beets (Be), maize (Ma), wheat (Wh), barley (Ba) and other crop species (Oc). 179
12.4 Evaluation of uncertainty in the water flux predictions due to uncertainty in the classified land cover map. 182
12.5 Probabilistic reference evapotranspiration ET_{ref} derived from ET_{RF}, obtained after simulation with the hard classified LCM-RF_{hard}, for different user-defined topindex (λ_T) classes. 188
12.6 Sensitivity index S of the area-averaged annual evapotranspiration, runoff and baseflow rates to small changes in the monthly values of the leaf area index (LAI), albedo (α), momentum roughness ($z_{0,m}$), heat roughness ($z_{0,h}$) and zero displacement height (d_0). 191

14.1 Model parameters and regression equations of the region specific (PTF A) and non-region specific pedotransfer functions (PTF B). . 204
14.2 Number of synthetically generated samples in the 9 truncated USDA soil classes considered in this study. 207
14.3 Mean (μ) and standard deviation per unit area on the texture triangle (σ_{unit}) of the SHPs as predicted from PTF A for the 9 selected soil classes of the USDA texture triangle. 212
14.4 Mean (μ) and standard deviation per unit area on the texture triangle (σ_{unit}) of the SHPs as predicted from PTF B for the 9 selected soil classes of the USDA texture triangle. 212
14.5 Skewness ($Skew$) and kurtosis ($Kurt$) of the SHPs as predicted from PTF A for the 9 selected soil classes of the USDA texture triangle. 213
14.6 Skewness ($Skew$) and kurtosis ($Kurt$) of the SHPs as predicted from PTF B for the 9 selected soil classes of the USDA texture triangle. 213
14.7 Mean (μ), standard deviation per unit area on the texture triangle (σ_{unit}), skewness ($Skew$) and kurtosis ($Kurt$) of the average simulated soil moisture content θ_{avg} (within the upper 5 cm of the soil during the year 2006) for the 9 selected USDA classes using the SHPs as predicted from PTF A and B. 220

15.1 Coordinates of the directional vectors. 241
15.2 Average sensitivity index within the USDA soil classes, for the different soil hydraulic parameters. The class with the highest average sensitivity is indicated in boldface, whereas the class with the lowest average sensitivity is underlined. 251

List of Abbreviations and Acronyms

CDF	Cumulative Distribution Function
CP	Central Point
CPU	Central Processing Unit
DOY	Day Of the Year
FF	Factors Fixing
FI – UA	Fast Indicative Uncertainty Assessment
FM	Factors Mapping
FP	Factors Prioritisation
HCC	Hydraulic Conductivity Curve
OAT	One-At-a-Time
PDF	Probability Density Function
PTF	Pedotransfer function
SA	Sensitivity Analysis
SHP	Soil Hydraulic Parameter
SMRC	Soil Moisture Retention Curve
UA	Uncertainty Analysis

Institutions

BBCH	Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie
DTU	Danish Technical University
FAO	Food and Agriculture Organization
FGIA	Flanders Geographical Information Agency
RCF	Research Centre Foulum
USDA	United States Department of Agriculture

Quantitative measures

AMI	Average Mutual Information
AUC	Area under the ROC-curve
CA	Class-specific Accuracy
FPR	False Positive Rate
GLCM	Grey-Level Co-occurrence Matrix
LAI	Leaf Area Index
MA	Model user Accuracy
MRE	Maximum Relative Error
OA	Overall Accuracy
PA	Producer's Accuracy

R	Ratio
RMSE	Root Mean Squared Error
ROC	Receiver Operating Characteristics
SAE	Sum of Absolute Errors
SSE	Sum of Squared Errors
TPR	True Positive Rate
UR	Uncertainty Ratio

Techniques and methods

BMA	Bayesian Model Averaging
CCE	Competitive Complex Evolution
DDS	Dynamically Dimensioned Search
eFAST	Extended Fourier Amplitude Sensitivity Test
FAST	Fourier Amplitude Sensitivity Test
FOSM	First-Order Second Moment
GLUE	Generalized Likelihood Uncertainty Estimation
ILR	Isometric logratio
LHS	Latin Hypercube Sampling
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
RF	Random Forests
RSM	Response Surface Method
SCE	Shuffled Complex Evolution
WARPE	Weight-Adaptive Recursive Parameter Estimation

Models and tools

LSM	Land Surface Model
MIKESHE	integrated hydrological modelling system that emerged from System Hydrologique European
RHEM	Rangeland Hydrology and Erosion Model
SVATS	Soil-Vegetation-Atmosphere Transfer Scheme
SWAT	Soil and Water Assessment Tool
TOPLATS	TOPMODEL-based Land Surface-Atmosphere Transfer Scheme
TOPMODEL	Topographic index model

Class labels

Ba	Barley
Be	Beets
Bf	Broad-leaved forest
Cr	Cropland

Cl	Clay
ClL	Clay loam
Fo	Forest
Gr	Grassland
L	Loam
LS	Loamy sand
Ma	Maize
Nf	Needle-leaved forest
Oa	Oats
Oc	Other crop species
Pa	Pasture
Pe	Peas
Po	Potatoes
S	Sand
Sb	Spring barley
SCl	Sandy clay
SClL	Sandy clay loam
Si	Silt
SiCl	Silty clay
SiClL	Silty clay loam
SiL	Silt loam
SL	Sandy loam
Ry	Rye
Ur	Urban area
Wa	Water bodies
Wb	Winter barley
Wh	Wheat
Ww	Winter wheat

Model input

CA – LUT	Catchment Look-Up Table
DEM	Digital Elevation Model
LC – LUT	Land Cover Look-Up Table
LCM – CROP	Land cover map of the Bellebeek with indication of the specific crop type
LCM – GEN	Land cover map of the Bellebeek with indication of the general land cover
LCM – RF	Land cover map of the Bellebeek obtained through classification of remote sensing imagery with Random Forests
SHPs – A	Soil hydraulic parameters as predicted with the PTFs of Vereecken et al. (1989, 1990)
SHPs – B	Soil hydraulic parameters as predicted with the PTFs of Rawls and Brakensiek (1985, 1989)

ST – LUT	Soil Type Look-Up Table
STM – BEL	Soil type map of the Bellebeek provided by FGIA
STM – FAO	Soil type map of the Bellebeek provided by FAO
PTFA	Regression equations of Vereecken et al. (1989, 1990)
PTFB	Regression equations of Rawls and Brakensiek (1985, 1989)

Remote sensing

ALOS	Advanced Land Observing Satellite
EMISAR	Electromagnetics Institute Synthetic Aperture Radar
HRV	High Resolution Visible
HRVIR	High-Resolution Visible and Infrared
RADARSAT	Satellite of the Canadian Space Agency
SAR	Synthetic Aperture Radar
SENTINEL	Satellite constellation of the European Space Agency
SPOT	Système Pour l'Observation de la Terre

List of Symbols and Notations

α	albedo [-]
θ	soil moisture content [$m^3.m^{-3}$]
θ_r	residual soil moisture content [$m^3.m^{-3}$]
θ_s	saturated soil moisture content [$m^3.m^{-3}$]
κ	kappa coefficient of agreement [-]
λ	pore size distribution index [-]
μ	mean value
$\pi(.)$	possibility function
ρ_b	soil bulk density [$g.cm^{-3}$]
ρ_s	particle density [$g.cm^{-3}$]
σ	standard deviation
ϕ	value of model parameter
ξ	perturbation factor
ψ_c	bubbling pressure [m]
$p(.)$	probability function
$E[.]$	expected value operator
$V[.]$	variance operator
∂	partial derivative operator
sup	supremum operator
inf	infimum operator
min	minimum operator
max	maximum operator
$\bar{.}$	average operator
$.^-$	a priori estimate
$.^+$	a posteriori estimate
$\langle\cdot\rangle$	ensemble averaging operation
$.*$	complex conjugation operation
C	clay content [%]
d_0	zero plane displacement height [m]
ET	evapotranspiration rate [$m^3.m^{-3}$] or [mm]
$f(.)$	model function
H	Shannon entropy
K_s	saturated hydraulic conductivity [$m.s^{-1}$]
L	silt content [%]
LAI	leaf area index [-]
N	number of time steps in the simulation
Q_b	baseflow [$m^3.s^{-1}$] or [mm]
Q_s	runoff at the outlet of the catchment [$m^3.s^{-1}$] or [mm]

r	correlation coefficient
R^2	model coefficient of determination
S	sensitivity index
t	time [s]
U	uncertainty measure equal to $1 - p_{max}$
X	model input variable
x	value of model input X
Y	model output variable
y	value of model output Y
Z	sand content [%]
$z_{0,m}$	roughness length for momentum transfer [m]
$z_{0,h}$	roughness length for heat transfer [m]

Part I

β	regression coefficient
Δx	the change in the value of the input variable X
ϵ	error term
$\Pi(E)$	possibility measure for event E
a	confidence level
$A(x)$	membership degree of x in A
A_α	α-cut of fuzzy set A
B	background information
$core(A)$	core of fuzzy set A
CSS	composite scaled sensitivity index
D	additional evidence
E	an event, a subset of the domain of X
E^c	the complement of event E
$f_X(.)$	probability distribution function of X
$F_X(.)$	cumulative distribution function of X
I	unit interval [0,1]
k	number of model structures
$L(.)$	likelihood function
l	number of model parameters
m	number of output variables
n	number of input variables
n_α	number of α-cuts in the extension principle propagation
$N(E)$	necessity measure for event E
n_{MC}	number of realization in the MC analysis
$P(E)$	probability measure of event E
SSE	sum of squared errors
$supp(A)$	support of fuzzy set A
u	value of model input X

x^0	baseline (or nominal) value of variable X
y^0	model output obtained with x^0
$\hat{y}$	output from the regression model

Part II

β	local slope angle [radians]
γ	psychrometric constant [Pa.K^{-1}]
ΔS	change in water storage per unit of surface area [m.s^{-1}]
ϵ	emissivity [-]
η	transmission factor [-]
θ_c	amount of water stored on the leaves in the canopy [m]
θ_l	soil moisture content of the transmission zone [m^3.m^{-3}]
θ_u	soil moisture content of the surface zone [m^3.m^{-3}]
ζ	soil thermal conductivity [W.m^{-1}.K^{-1}]
λ_T	topographic index
λ_w	latent heat of vaporization of water [J.kg^{-1}]
ρ_a	density of air [kg.m^{-3}]
ρ_w	density of water [kg.m^{-3}]
ν	Stefan-Boltzmann constant (5.6676W.m^{-2}K^{-4})
τ	time lag [s]
ξ	psychrometric constant [Pa.K^{-1}]
ψ	matric head [m]
ω_{wc}	areal fraction of wet canopy
A	area of the catchment [m^2]
a	area draining into the given location [m]
a_0	an empirical constant [-]
b	diffusion coefficient [m^2.s^{-1}]
c_p	specific heat of air [J.kg^{-1}.K^{-1}]
C_s	dry soil heat capacity [J.K^{-1}.kg^{-1}]
D	diffusivity [m^2.s^{-1}]
d_l	diffusive flux from the lower zone to the saturated zone [m.s^{-1}]
d_u	diffusive flux from the upper zone to the lower zone [m.s^{-1}]
e_a	vapor pressure in the air [Pa]
E_{bs}	actual rate of bare soil evaporation ([m.s^{-1}])
e_s	saturated vapor pressure [Pa]
E_{vg}	actual rate of canopy transpiration ([m.s^{-1}])
E_{wc}	evaporation rate from the wet canopy [m.s^{-1}]
F	Froude's number [-]
$f_b(t)$	unit hydrograph
f_{bs}	fraction of bare soil surface [-]
$f_c(t)$	response function for open channel flow
$f_h(t)$	response function for overland flow

f_K	hydraulic conductivity decay coefficient [m^{-1}]
f_{rt}	root fraction in a given zone [-]
f_{vg}	fraction of vegetated land surface [-]
G	soil heat flux [$W.m^{-2}$]
g	acceleration of gravity [9.81 $m.s^{-2}$]
g_l	gravitational drainage towards the ground water table [$m.s^{-1}$]
g_u	gravitational drainage towards the lower zone [$m.s^{-1}$]
H_s	sensible heat flux [$W.m^{-2}$]
h	distance along the channel [m]
$H(h)$	normalized hillslope function
H_h	water height above the bed in the hillslopes [m]
$\mathbf{H}_k$	Jacobian matrix
H_{rel}	relative humidity [-]
H_w	water height above the bed in the rivers [m]
$\mathbf{I}$	unity matrix (elements on diagonal are one)
I_{bs}	actual infiltration into the bare soil [$m.s^{-1}$]
I_{cap}	infiltration capacity [$m.s^{-1}$]
I_{vg}	actual infiltration into the vegetated soil [$m.s^{-1}$]
K_0	hydraulic conductivity at the surface [$m.s^{-1}$]
$\mathbf{K}_k$	Kalman gain factor
N_{eval}	number of objective function evaluations
$o(.)$	objective function
P	precipitation rate [$m.s^{-1}$]
p_a	air pressure [Pa]
$\mathbf{P_k}$	error covariance matrix at the k-th iteration
P_n	net precipitation rate [$m.s^{-1}$]
q	quartz content [-]
Q_0	critical baseflow [$m^3.s^{-1}$]
q_{bs}	runoff flux for the bare soil [$m.s^{-1}$]
q_c	discharge rate per unit width of channel caused by the channel flow [$m^2.s^{-1}$]
q_h	discharge rate per unit width of channel caused by the overland flow [$m^2.s^{-1}$]
$\mathbf{Q_k}$	model noise covariance matrix at the i-th iteration step
q_s	local surface runoff [$m.s^{-1}$]
q_{vg}	vegetated runoff flux [$m.s^{-1}$]
r_p	perturbation size parameter
r_{ah}	aerodynamic resistance for heat transport [$s.m^{-1}$]
r_{av}	aerodynamic resistance for vapor transport [$s.m^{-1}$]
r_c	canopy resistance [$s.m^{-1}$]
$\mathbf{R_k}$	observation error correlation matrix
R_l	incoming longwave radiation [$W.m^{-2}$]
R_n	net radiation [$W.m^{-2}$]

R_s	incoming shortwave radiation [W.m^{-2}]
$r_{s,min}$	minimum soil resistance [s.m^{-1}]
$r_{st,min}$	minimum stomatal resistance [s.m^{-1}]
$r_{st,max}$	maximum stomatal resistance [s.m^{-1}]
s	number of sampled points in the SCE
S_0	channel bed slope [-]
T	soil temperature [K]
T_a	the air temperature [K]
T_b	dry bulb temperature [K]
T^m	transmissivity [m^2.s^{-1}]
T_s	skin temperature [K]
u	wind speed [m.s^{-1}]
v	drift velocity [m.s^{-1}]
V_h	flow velocity in the hillslopes [m.s^{-1}]
V_w	flow velocity in the channels [m.s^{-1}]
w	capillary rise [m.s^{-1}]
$W(h)$	normalized width function
z_0	initial water table depth [m]
z_a	height of the thermodynamic measurements [m]
z_i	water table depth [m] at location i
z_l	depth of the lower zone [m]
z_u	depth of the upper zone [m]
z_w	height of the wind speed measurements [m]

Part III

β	local incidence angle
$\Delta c_{l,j}$	raw importance score for variable j and tree l
γ_T	topographic index
γ^0_{tr}	angle-corrected gamma naught backscatter coefficient
σ^0_{tr}	sigma naught backscatter coefficient
$\overline{\alpha}$	angle in the Cloude-Pottier decomposition
ϵ	eigenvalue
ρ_{hhvv}	correlation coefficient between the co-polarized linear responses
ϕ_{hhvv}	phase difference between the co-polarized responses
ψ_{tr}	phase of the complex number S_{tr}
A	anisotropy in the Cloude-Pottier decomposition
a	areal cover percentage per land cover [-]
AE	percentage entropy w.r.t a chosen reference map
AMI	average mutual information
AUC_{avg}	area under the ROC-curve, averaged over the different classes
$\mathbf{C}$	covariance matrix
C_j	mean importance score for variable j

c_l	number of correct classifications by the l-th tree
$c_{l,\tilde{j}}$	number of correct classifications by the l-th tree when variable j is perturbed
ET_{ref}	reference evapotranspiration for a given land cover class
ET_{RF}	evapotranspiration for the simulation with LCM − RF_{hard}
ET_{proxy}	approximation of the Monte Carlo-based evapotranspiration as obtained with the FI-UA
ET_{MC}	evapotranspiraton from the Monte Carlo simulation
h	canopy height [m]
H_{e}	entropy in the Cloude-Pottier decomposition
$k_u(i)$	number of trees classifying the land cover of pixel u as cover type i
M	McNemar test statistic
m	number of variables used to split the nodes in the tree
n_{class}	number of land cover classes
n_{hill}	number of hillslope classes
n_{oob}	number of out-of-bag elements
n_{tot}	total number of elements in the data set
n_{tree}	number of trees in the Random Forest
n_{var}	number of predictor variables
p_e	proportional agreement expected by chance
$p_{\max}$	maximum probability in the probability vector
p_o	proportional observed agreement
$\mathbf{p}_u$	probability vector for pixel u
q	quantile
r	Spearman correlation coefficient
S_{tr}	complex scattering amplitude
UR_{classif}	uncertainty ratio due to classification uncertainty
UR_{map}	uncertainty ratio due to land cover generalization
$w(i)$	weighing factor associated with class i
$y^{a\%}$	model output at the a-th quantile
$y^{\text{LCM}}(t)$	model output at time step t obtained with the given land cover map
$z-\text{score}$	scaled importance score

Part IV

$\boldsymbol{\alpha}$	parameter vector in the Dirichlet distribution
θ_t	soil moisture at time t [$\text{m}^3.\text{m}^{-3}$]
θ_{daily}	daily simulated soil moisture [$\text{m}^3.\text{m}^{-3}$]
θ_{avg}	annually-averaged simulated soil moisture [$\text{m}^3.\text{m}^{-3}$]
κ_C	closure constant [-]
ψ	matric head [m]
ω	scalar

$\nabla_{\mathbf{v}}$	derivative in direction v
$\|.\|_A$	norm operator
a	van Genuchten parameter $[m^{-1}]$
C_i	sum of squared absolute difference for direction i
D	number of dimensions
d	distance
$d_A(.)$	Aitchison distance
$f_D(.)$	density function of the Dirichlet distribution
$Kurt_t$	kurtosis of the PDF at time t
M	number of directions to perturb
n	van Genuchten parameter [-]
n_{class}	number of soil classes
n_{e}	number of ensemble members
OC	organic carbon content [%]
P	porosity [%]
$\mathbf{p_0}$	baricenter of the 3D-simplex sample space
S	sensitivity index
S^D	simplex sample space with D dimensions
$Skew_t$	skewness of the PDF at time t
$\mathbf{T}$	texture sample
$\mathbf{v_{i+}}$	directional vector with direction i

Dutch summary - Samenvatting

Gedurende de voorbij decennia zijn hydrologische modellen sterk toegenomen in complexiteit omwille van een toegenomen detail in de beschrijving van de hydrologische processen. Nu hydrologen beschikken over modellen die een goeie weergave zijn van de werkelijkheid, worden ze geconfronteerd met een nieuw probleem: de onzekerheid in het toegenomen aantal model inputs en parameters dat nodig is om het model aan te sturen. Hoewel de modelstructuur (gewoonlijk) heel betrouwbaar is, zijn de bekomen modelresultaten heel onzeker omwille van onzekerheid in de talrijke inputs (veronderstellend dat de onzekerheid in de modelstructuur van ondergeschikt belang is hoewel dit niet noodzakelijk altijd het geval is). Het is belangrijk dat de onzekerheid op de modelresultaten geïntegreerd wordt in de besluitvorming door deze te communiceren naar de beleidsvormers. Hierdoor beschikken deze laatste over extra informatie inzake het belang dat mag worden gehecht aan de modelvoorspellingen en hoe sterk deze mogen doorwegen in de besluitvorming. Vooral wanneer hydrologische voorspellingen gebruikt worden in waarschuwingssystemen, bvb. bij het voorspellen van overstromingen, kan de onzekerheid onmogelijk genegeerd worden omwille van de maatschappelijke belangen die ermee gepaard gaan. Ondanks de toegenomen vraag naar onzekerheidsmodellering is dit nog niet opgenomen als standaard procedure in hydrologische modellering. Dit is voornamelijk te wijten aan het gebrek aan richtlijnen om onzekerheid te begroten en aan de lange rekentijd.

Dit werk is ontsproten aan de vele problemen met betrekking tot onzekerheid in hydrologische modellen, en meer specifiek in ruimtelijk verdeelde modellen. Het accent in dit werk ligt op (i) onzekerheid in de karakterisering van het landoppervlak, dit betreft de inputs en parameters die enerzijds de landbedekking en anderzijds de bodem beschrijven, en (ii) simulaties in slecht of niet-bemonsterde stroomgebieden. In dergelijke omstandigheden is er slechts een beperkte hoeveelheid informatie over de karakteristieken van het landoppervlak beschikbaar. Gewoonlijk bestaat deze informatie uit reeds beschikbaar kaartmateriaal (aangeleverd door de lokale overheid of door internationale instellingen), teledetectie data (tegenwoordig is er een enorm aanbod via verschillende bronnen), en data die via de literatuur aangeleverd wordt. De verschillende bronnen van informatie impliceren verschillende types onzekerheid, die op een verschillende manier moeten behandeld en gemodelleerd worden. Hoewel heel wat onzekerheid in hydrologische modellering epistemisch is (d.w.z. te wijten aan een gebrek aan kennis), behandelen we in dit werk voornamelijk aleatorische onzekerheid (d.w.z. te wijten aan natuurlijke variabiliteit), welke op een probabilistische manier kan gemodelleerd worden.

Deel I: Basisprincipes van onzekerheids- en gevoeligheidsanalyse
Dit deel start met een overzicht van de state-of-the-art technieken voor onzekerheids- en gevoeligheidsanalyse. Bij elke techniek worden enkele voorbeelden van hydrologische toepassing gegeven en wordt stil gestaan bij de typische valkuilen die gepaard gaan met de toepassing ervan. Het literatuuroverzicht laat toe om voor de specifieke problemen in de volgende delen van dit werk een structuur voor de onzekerheids- of gevoeligheidsanalyse op te stellen. Het is duidelijk dat een succesvolle onzekerheidsanalyse begint bij een correcte identificatie van het type onzekerheid omdat dit de keuze van de techniek zal bepalen. Aleatorische onzekerheid wordt het best weergegeven d.m.v. een probabiliteitsverdeling, terwijl epistemische onzekerheid beter wordt weergegeven d.m.v. probabiliteitsboxen, possibiliteitsverdeling of gelooffuncties. Bij een gevoeligheidsanalyse speelt het doel van de analyse een cruciale rol in de selectie van een geschikte techniek.

Deel II: Ruimtelijk verdeeld modelleren van de hydrologische processen
In een volgend deel wordt de ruimtelijke modellering van de hydrologische processen besproken. Dit omvat een gedetailleerde beschrijving van de structuur van het hydrologische TOPLATS model, een beschrijving van het stroomgebied van de Bellebeek, waarvoor de modellering wordt uitgevoerd, en een beschrijving van de beschikbare data. Het deel wordt afgesloten met een calibratie van het hydrologische model waarbij verschillende technieken voor automatische calibratie met elkaar vergeleken worden. De methoden die gebaseerd zijn op het minimaliseren van een objectieffunctie werden niet geschikt bevonden om het TOPLATS model te calibreren, hoewel de performantie statistieken duidden op een optimilisatie van de modelparameters. De Kalman filter-gebaseerde methode daarentegen (welke een minimalisatie van de onzekerheid op de geschatte parameters beoogt) was wel in staat om simulaties te produceren die een sterke overeenkomst tonen met de geobserveerde afvoer. Hieruit blijkt dat het belangrijk is om voldoende aandacht te besteden aan de selectie van een gepaste techniek voor automatische calibratie van ruimtelijk verdeelde modellen en dat evaluatie van het calibratieresultaat op basis van performantie statistieken best gepaard gaat met een visuele inspectie van de gesimuleerde output.

Deel III: Impact van onzekere input m.b.t. de landbedekking op de hydrologische modelvoorspellingen
In dit deel van de thesis worden verschillende aspecten van onzekerheid die betrekking hebben op de landbedekkingsinput behandeld. Doorgaans wordt de landbedekking in ruimtelijk verdeelde hydrologische modellen beschreven a.d.h.v. een landbedekkingskaart en een set van parameters die de biofysische karakteristieken van elk type landbedekking beschrijft. Het effect van landbedekking op de hydrologische modelrespons wordt dus uiteindelijk gestuurd door deze biofysische parameters. Ze spelen een belangrijke rol in het model omdat ze de uitwisseling van energie en water tussen het landoppervlak en de atmosfeer bepalen. Hierdoor zullen veranderingen in de parameterwaarden het gemodelleerde energiebudget

in het gekoppelde land-atmosfeer-transfer schema wijzigen en een verandering in verdamping, runoff en infiltratie teweeg brengen. De landbedekkingskaart op zijn beurt bepaalt het aandeel van de verschillende types landbedekking in het stroomgebied en tevens het detail waarmee de informatie beschikbaar is (zoals vastgelegd door de legende).

Hoewel de meeste vegetatieparameters, bvb. LAI en gewashoogte, kunnen gemeten worden in het veld, worden hydrologische simulaties vaak uitgevoerd voor stroomgebieden waarvoor weinig data inzake vegetatie beschikbaar is. Informatie omtrent de landbedekking dient daarom indirect te worden afgeleid, bvb. via expertkennis, literatuurwaarden of teledetectie data. Dit zorgt ervoor dat de verkregen informatie (heel) onzeker is waardoor de betrouwbaarheid van de modelvoorspelling gehypothekeerd wordt vermits de onzekerheid op de landbedekking propageert doorheen het hydrologische model. De onzekerheids- en gevoeligheidsanalyses in dit deel van de thesis passen in het kader van voorspellingen in slecht of niet-bemonsterde gebieden, wat betekent dat er geen of nauwelijks veldmetingen van de landbedekkingskarakteristieken beschikbaar zijn. Afhankelijk van het beschouwde studiegebied, zullen andere types aan informatie voorhanden zijn zodat ook de onzekerheid omtrent de landbedekking sterk variabel is. In deze thesis worden twee verschillende bronnen van onzekerheid m.b.t. landbedekking behandeld:

1. Een landbedekkingskaart van het stroomgebied is onmiddellijk beschikbaar via meerdere bronnen. Afhankelijk van de bron kunnen verschillen in ruimtelijke resolutie, legende en informatie-inhoud aanwezig zijn. Deze laatste kan variëren van gedetailleerde informatie over het geteelde gewas, bvb. bieten of aardappelen, naar sterk algemene informatie over het vegetatietype, bvb. akkerland of bos. De onzekerheid die geïntroduceerd wordt in de modelvoorspelling door het verschil aan informatie in de beschikbare kaarten wordt in dit deel begroot en geanalyseerd. Er wordt nagegaan of en in welk geval meer gedetailleerde informatie over de landbedekking relevant is voor de hydrologische simulaties. De resultaten tonen aan dat het effect van het detail van de informatie (algemene landbedekking versus gewas type) afhangt van het doel van de simulatie: indien gebiedsgemiddelde voorspellingen van de water fluxen afdoende zijn, is het niet zinvol om informatie over het geteelde gewas in te brengen; indien lokale voorspellingen van de fluxen wel van belang zijn, is het aangewezen om meer gedetailleerde informatie te gebruiken vermits een toename in de heterogeniteit van de landbedekking resulteert in een meer heterogeen modelrespons en alsdanig de onzekerheid op de modelvoorspelling reduceert.

2. Een landbedekkingskaart van het stroomgebied is niet onmiddellijk beschikbaar of heeft niet het vereiste detail. In dergelijke situatie dient een landbedekkingskaart te worden afgeleid van teledetectie data d.m.v. classificatie. Zowel optische als radarbeelden kunnen hiervoor gebruikt worden, hoewel

radarbeelden het voordeel hebben dat ze minder beïnvloed worden door de aanwezigheid van wolken. Daartegenover staat dat radarbeelden een hogere dimensionaliteit (bvb. temporeel of frequentieel) vereisen om te resulteren in een afdoende classificatie. In dit werk wordt de onzekerheid in het geclassificeerde beeld begroot en geëvalueerd m.b.t. de dimensionaliteit van de data. Er blijkt dat er een duidelijk verband bestaat tussen de accuraatheid van de classificatie en de ontwikkelingsstadia van het gewas en dat multi-temporele data beter geschikt zijn dan multi-frequentie data voor de classificatie van landbedekking. Door het reduceren van de input data tot een set van de meest belangrijke inputs kon de onzekerheid op het classificatie resultaat gereduceerd worden. Op basis hiervan wordt aangegeven hoe de verwerving van teledetectie data kan geoptimaliseerd worden om landbedekkingskaarten met lage onzekerheid te bekomen. Daaropvolgend wordt het effect van de onzekerheid in het geclassificeerde beeld op de hydrologische modelvoorspelling besproken. Hieruit blijkt dat propagatie van de classificatie onzekerheid niet nuttig is indien stroomgebiedsgemiddelde voorspellingen beoogd worden en dat een betrouwbare voorspelling kan bekomen worden door het toepassen van het hydrologische model met de meest waarschijnlijke landbedekking als input. In andere gevallen kan propagatie van de classificatie onzekerheid eenvoudig ingeschat worden d.m.v. de FI-UA methode die werd ontwikkeld in dit deel van het proefschrift.

Deel IV: Impact van onzekere input m.b.t. de bodem op de hydrologische modelvoorspellingen

In dit deel van de thesis worden verschillende aspecten van onzekerheid die betrekking hebben op de bodem behandeld. Doorgaans wordt in ruimtelijk verdeelde modellen de bodem beschreven a.d.h.v. een bodemkaart en een set van bodemparameters, die de fysische eigenschappen van elk bodemtype beschrijft. De bodemparameters spelen een cruciale rol in hydrologische modellen vermits ze het transport van water en de waterbalans rechtstreeks beïnvloeden. De bodemkaart op zijn beurt bepaalt de ruimtelijke verdeling van de verschillende bodemtypes en het detail waarmee de informatie beschikbaar is (zoals vastgelegd door het bodemclassificatieschema).

Hoewel de meeste bodemparameters zoals bulkdichtheid en hydraulische geleidbaarheid kunnen opgemeten worden in het veld, worden hydrologische simulaties vaak uitgevoerd voor stroomgebieden waarvoor weinig bodemkundige data beschikbaar is. Dit is te wijten aan het tijdrovende en dure karakter van veldmetingen op regionale schaal. De bodemparameters dienen dan indirect bepaald te worden, bvb. op basis van de bodemtextuur. Vaak wordt gebruik gemaakt van zogenaamde pedotransferfuncties die de bodemtextuur vertalen naar hydraulische parameters. Bijgevolg zal de onzekerheid op de bodemtextuur propageren doorheen het hydrologische model en onzekerheid op het gesimuleerde bodemvocht introduceren.

In deze thesis wordt de onzekerheid op de modelpredictie geanalyseerd voor twee verschillende situaties m.b.t. de beschikbare bodemkundige informatie:

1. Een bodemkaart van het stroomgebied is onmiddellijk beschikbaar maar er zijn geen bijkomende metingen van de bodemtextuur of de fysische bodemparameters uitgevoerd. Gezien de bodemkaart de enige bron van informatie is, is de exacte bodemtextuur (uitgedrukt als een percentage zand, klei en leem) onzeker omwille van de variabiliteit van de textuur binnen een gegeven bodemtype. De variabiliteit in textuur wordt in deze thesis begroot en de resulterende onzekerheid op het gesimuleerd bodemvocht wordt ingeschat. De resultaten tonen aan dat de onzekerheid t.g.v. texturele variabiliteit niet kan verwaarloosd worden wanneer een betrouwbare predictie vereist is. Bij het inschatten van de onzekerheid is de keuze van de pedotransferfunctie van cruciaal belang gezien verschillen in de pedotransferfuncties zich manifesteren in de modelresultaten na propagatie doorheen het model. Regio specifieke pedotransferfuncties moeten verkozen worden boven niet-regio specifieke pedotransferfuncties gezien verondersteld mag worden dat de eerstgenoemde meer betrouwbaar zijn. Eveneens blijkt dat een representatieve karakterisatie van de onzekerheid vereist dat de hogere-orde momenten van de probabiliteitsverdeling in rekening worden gebracht, vooral wanneer droge bodemcondities van toepassing zijn.

2. Een bodemkaart van het stroomgebied is onmiddellijk beschikbaar alsook een beperkte set van textuurmetingen (typisch betreft dit één meting per bodemtype). Tijdens het afleiden van de bodemparameters uit de bodemtextuur wordt verondersteld dat de textuurmeting representatief is voor de volledige bodemeenheid zoals aangegeven op de bodemkaart. De afwijking tussen de schaal waarop de metingen worden uitgevoerd en de schaal waarop het model wordt toegepast, kan echter aanleiding geven tot een hoge onzekerheid in het gesimuleerde bodemvocht. In deze thesis wordt een gevoeligheidsanalyse uitgevoerd die de gevoeligheid van het gesimuleerde bodemvocht voor kleine veranderingen in de bodemtextuur moet nagaan. De methode maakt gebruik van de eindige differentie techniek maar is aangepast voor het gebruik van textuurgegevens, die gekenmerkt worden door een compositioneel (of gesloten) karakter. De resultaten tonen aan dat kleiige bodems meer gevoelig zijn voor variaties in textuur dan zandige bodems zodat voor het afleiden van een representatieve textuur per bodemeenheid meer bodemmonsters zullen vereist zijn in geval het eerste type bodem van toepassing is.

Dit werk besluit met een samenvatting van de antwoorden op de onderzoeksvragen en enkele suggesties voor verder onderzoek.

Summary

During the last decades, the complexity of hydrological models has considerably increased in order to better represent the hydrological processes. Since hydrologists have obtained models that are good representations of reality, they are being confronted with a new type of problem: uncertainty in the increasing number of model inputs and parameters that are needed to drive the complex model. Although the model structure itself is (generally) highly reliable, the resulting model predictions are highly uncertain due to uncertainty in the numerous model inputs (assuming that uncertainty in the model structure is of minor importance although this is not necessarily the case). This prediction uncertainty should be integrated in the decision-making process by communicating the uncertainty to the decision makers such that they dispose of additional information concerning the weight that should be given to the model predictions. Especially when the hydrological model predictions are used in early-warning systems, e.g. for flooding, uncertainty cannot be disregarded from a societal point of view. Despite the growing demand for uncertainty analysis, it has not yet become a standard practice in hydrological modelling, mainly due to the lack of guidance and due to the high computational cost.

This work has emerged from the problems regarding uncertainty in hydrological models, and more specifically in spatially distributed hydrological models. The focus is on (i) uncertainty related to land surface characterization, *i.e.* land cover and soil inputs and parameters and (ii) simulations in poorly gauged or ungauged basins. Under the latter condition, there is a limited amount of information about the land surface characteristics available, usually consisting of readily available maps (provided by the local government or by an international organisation), remote sensing data (nowadays there is a variety of easily accessible sources) and literature-based data. These different types of information imply different types of uncertainty that should be dealt with and modelled in a different way. Although a lot of uncertainty in hydrological modelling is epistemic (due to a lack of knowledge), this dissertation mainly deals with aleatoric uncertainty (due to natural variability) that is generally modelled in a probabilistic way.

Part I: General concepts of uncertainty and sensitivity analysis
The dissertation starts with an overview of the state-of-the-art techniques for uncertainty and sensitivity analysis. For each technique, some examples of applications in hydrological modelling are given and special attention goes to the typical pitfalls that are involved. The literature review allows to set-up an appropriate framework for uncertainty and sensitivity analysis for the specific problems that are dealt with in the subsequent parts. It pointed out that a successful uncertainty analysis starts

with a correct identification of the type of uncertainty because this determines the selection of an appropriate technique for uncertainty modelling and propagation. Aleatoric uncertainty is best represented using probability distributions, whereas epistemic uncertainty is more faithfully represented using probability boxes, possibility distributions or belief functions. For sensitivity analysis the objective of the analysis is the crucial factor when selecting an appropriate technique. Moreover, special attention should be given to the correlation between the model inputs because the effect of simultaneously varying model inputs might not be the same as the superposition of their individual effects.

Part II: Distributed hydrological modelling

In a next part, the distributed modelling of the catchment hydrology is discussed and includes a detailed description of the hydrological model structure of TOPLATS, a description of the Bellebeek catchment for which the modelling is performed and a description of the available data. It ties up with a calibration of the hydrological model for the study area under consideration in which different algorithms for automatic calibration are evaluated. We found that the methods that minimize an objective function were unsuitable for calibrating the TOPLATS model although the performance statistics indicated that the parameter values have been optimized. On the contrary, the Kalman filter-based method which minimized the uncertainty in the estimated parameter values was able to produce simulations that closely resembled the observed discharge. As such, it is important to carefully select an appropriate algorithm for the automatic calibration of a spatially distributed hydrological model and to evaluate the calibration result based on a combination of the model performance statistics and a visual inspection of the model simulation.

Part III: Impact of uncertain land cover information on hydrological model predictions

In this part of the dissertation, different aspects of uncertainty modelling w.r.t. land cover (vegetation) input are investigated. Usually, the land cover in spatially distributed hydrological models is described by a land cover map and a set of parameters that describe the biophysical characteristics for each of the land cover classes. As such, the effect of land cover on the hydrological model response is ultimately controlled by the biophysical parameters. The latter play a crucial role within the model because they determine the energy and moisture exchange between the land surface and the atmosphere, such that changes in the land cover parameters alter the modelled energy budget in the coupled land-atmosphere-transfer scheme and directly affect evapotranspiration, runoff and infiltration. The land cover map determines the areal proportions of the different land covers in the watershed and the level of detail at which the land cover information is present (as given by the land cover labels).

Although most of the vegetation parameters, such as LAI and crop height, can be measured in the field, hydrological model simulations are typically performed under conditions of data scarcity concerning the vegetation characteristics (including predictions in ungauged basins). Land cover information hence needs to be inferred by indirect means, e.g. from expert-knowledge, from values reported in literature or from remotely sensed data. This causes the land cover information to be (highly) uncertain and affects the reliability of the simulation result as the uncertainty propagates through the hydrological model. The uncertainty and sensitivity analyses within this part of the dissertation fit within the scope of water flux predictions in ungauged basins, *i.e.* no field-based measurements of the land cover characteristics are available. Depending on the study area under consideration, different types and levels of information concerning the land cover are available, such that also the uncertainty in the land cover input is highly variable. In this part, two different sources of uncertainty are distinguished based on the status of the land cover information:

1. A land cover map of the basin is readily available from different data sources. Depending on the data source, differences in spatial resolution, class labeling, and information content may arise, ranging from detailed information about the crop type, e.g. beets or potatoes, to general information about the vegetation type, e.g. cropland or forest. The uncertainty introduced by two different land cover data sets in the prediction of water fluxes is quantified. It is investigated whether and when detailed land cover (crop) information is relevant for hydrological model simulations. Results show that the effect of the level of land cover information (general land cover classes versus crop type classes) depends on the purpose of the hydrological model application: if modellers are interested in area-averaged predictions of the water fluxes, adding information about the crop type is not very useful; if local predictions of the water fluxes are of interest, it is meaningful to increase the detail of land cover information because an increased heterogeneity in the land cover map results in an increased heterogeneity in the hydrological model response and as such reduces the uncertainty in the water flux predictions.

2. A land cover map of the basin is not readily available or does not have the required level of detail. In this case, a land cover map should be derived from remotely sensed data through classification analysis. Both optical and radar-based imagery can be used for this purpose. The latter has the advantage over the former of being less affected by cloud cover, but requires a higher dimensionality (e.g. frequency or temporal) to result in a good classification. The uncertainty in the land cover classification is quantified and is evaluated w.r.t. the dimensionality of the data. Results show that there is a clear relationship between classification performance and crop development and that multi-date information is more useful than multi-frequency data. Through importance analysis, the input data were reduced to a set of most

important remote sensing features for land cover classification, which caused a decline in the classification uncertainty. Based on these results, it is indicated how the acquisition of remote sensing imagery can be optimized to produce land cover maps with a low uncertainty. Following this, a discussion on the effect of classification uncertainty on the prediction of water fluxes is given, which pointed out that propagating the classification uncertainty is not very useful when area-averaged predictions of the water fluxes are intended and that reliable predictions can be easily obtained by applying the hydrological with the hard classified (most probable) land cover map. In other cases, propagation of the classification uncertainty can be performed in a computationally efficient way using the fast indicative uncertainty assessment method that was developed in this part of the thesis.

Part IV: Impact of uncertain soil information on hydrological model predictions

In this part of the dissertation, different aspects of uncertainty modelling w.r.t. soil related inputs are investigated. Usually, the soil information in spatially distributed hydrological models is given by a soil type map and a set of soil parameters that describe the physical soil properties for each of the soil types on the soil map. The soil parameters play a crucial role in hydrological models as they directly control the movement of water and the water balance partitioning. The soil map determines the areal proportions of the different soil types in the watershed and the level of detail (as given by the soil classification scheme) at which the soil information is present.

Although most soil parameters, such as bulk density and hydraulic conductivity, can be measured in the field, hydrological model simulations are typically performed under conditions of data scarcity concerning the soil characteristics because of the time consuming and costly nature of soil data collection at the regional scale (predictions in ungauged basins). Therefore, the soil parameters need to be determined indirectly, e.g. from textural information. Modellers often rely on relationships between the soil hydraulic parameters and the soil texture by using pedotransfer functions, which translate soil related information into the hydrological parameters needed. As such, the uncertainty in the textural information propagates through the hydrological model and induces uncertainty in the simulated soil moisture.

For two different situations w.r.t. the availability of soil information, the uncertainty is analyzed:

1. A soil map of the basin is readily available, but measurements of the soil texture or the soil physical parameters are lacking. As the soil type is the only source of information, the exact soil texture (in terms of the particle sizes clay, sand and silt) is prone to uncertainty due to variability of the textural composition within a given soil type. The within-soil-class variability in the

soil parameters is quantified and the resulting uncertainty in the simulated soil moisture is assessed. Results point out that the uncertainty arising from textural variability within soil classes cannot be disregarded when a reliable model prediction is required. In the uncertainty assessment, the choice of the pedotransfer function is a crucial step since intrinsic differences in the pedotransfer function remain present after propagating the estimated soil parameters through the hydrological model. Region specific pedotransfer functions are to be preferred over non-region specific pedotransfer functions as the former are assumed to be more reliable. Further, a representative characterization of the uncertainty requires that higher-order moments of the probability distribution are taken into account, especially when dealing with dry soil conditions.

2. A soil map of the basin is readily available together with a set of texture measurements scarcely distributed over the basin (usually one measurement per soil unit on the map). When deriving the soil parameters from the textural information, it is assumed that the texture measurement is representative for the entire soil unit. However, the discrepancy between the scale of measurement and the scale of model application might give rise to large uncertainties in the simulated soil moisture. A sensitivity analysis of the simulated soil moisture w.r.t. soil texture is developed by modifying the finite difference technique to deal with the closed character of texture data. Results point out that clayey soils are more sensitive to textural variations than sandy soils, such that in order to derive a representative texture, a higher sampling density should be applied for the former type of soils as compared to the latter type of soils.

The work concludes with a summary of the answers to the research questions and some suggestions for future research.

1

Introduction

1.1. History of hydrological modelling

Hydrological modelling in the early years (second half of the 19th century) was undertaken to better understand the relationship between rainfall and runoff (*i.e.* descriptive modelling) or to estimate the discharge of a river (*i.e.* predictive modelling). Depending on the goal of the model, approaches to the modelling differed, but the model evaluation was performed objectively through comparison of the model predictions with the observations. Many of the first models were based on empirical equations to predict peak flow, some of them using the rational method (Mulvany, 1850) in which the peak discharge is a function of the rainfall intensity. In the early 1930's, the concept of the unit hydrograph was introduced by Sherman (1932). It was based on the principle of superposition of effects and enabled modellers to estimate the complete shape of the hydrograph rather than simply estimating the peak flows. The unit hydrograph dominated hydrological modelling for more than 25 years (although it is still frequently used today) such that it lasted until the 1960's before modellers began to develop conceptual models, e.g. the Stanford Watershed Model (Crawford and Linsley, 1966). In these newly developed models, the responses of and interconnections between the various components in the hydrological cycle were represented by means of conceptual elements. Because the conceptual models needed to be parameterized, model parameterisation was introduced at that time (Dawdy and O'Donnell, 1965). At the end of the 1970's, lumped models emerged, e.g. PDM (Moore and Clarke, 1981), which were based on the idea that the rainfall-runoff process is mainly dominated by the dynamics of the saturated areas. Although lumped models were characterized by few significant model parameters, parameterisation of the processes was based on empirical parameters that were derived from the available data. This was an incentive for modellers to introduce more physically meaningful parameters, e.g. in the TOPMODEL (Beven and Kirkby, 1979). Following this, they also introduced distributed physical knowledge of the surface and subsurface processes in the hydrological models, which led to the development of spatially distributed models. In these models, flow and transport processes are represented by finite difference representations of partial differential equations. In the subsequent decades, the use of spatially distributed models was encouraged because of (i) the increasing

availability of spatially distributed data sets, e.g. remote sensing data, and (ii) the increasing availability of powerful computer resources (Pechlivanidis et al., 2011). Although distributed modelling has been argued to be more realistic than simpler models (Beven, 2001), there does not seem to be a clear trend in the literature supporting the superiority of distributed against lumped models in terms of model accuracy (Pechlivanidis et al., 2011). Moreover, lumped models have been found to perform better than their distributed counterparts in case there is a spatial inconsistency in the model parameters (Ajami et al., 2004). The discussion about the superiority of different model types was tackled by the introduction of the concept of equifinality, which states that models with a different complexity might result in equally good reproductions of the system response (Beven, 1993).

Irrespective of the approach and complexity of the hydrological model, uncertainty in the model output variables is inevitable as a model is always a simplification of the real system. Even if all assumptions regarding the model structure and input data are believed to be true, uncertainties present in the description of the model will influence the model predictions. From the history of hydrological modelling, it is clear that there is a natural tendency for models to grow in complexity in order to better represent the processes of the system. Paradoxally, the building of higher levels of complexity into hydrological models has increased the need for methods to deal with the increasing uncertainty in the models itself and in the data that drive the model (Liu and Gupta, 2007). In case of spatially distributed models, the number of model parameters increases with the degree of spatial discretization such that uncertainty in the model parameters is a common problem in this type of models (Madsen et al., 2002).

Hydrological simulations are often used in real time prediction systems for natural hazards or for assessing the long term effects of climate change or water resource infrastructures. This makes quantification of uncertainty in the hydrological model response in such situations extremely important from a societal point of view (Montanari et al., 2009). Moreover, there is an increasing attention for hydrological prediction in ungauged or poorly gauged basins, which represent the majority of practical applications, since prediction under these conditions is highly uncertain. This led to a growing interest of the model users for uncertainty analysis, which is the means of calculating and representing (un)certainty with which the model predictions represent reality (Singh, 1995). The resulting measures of confidence in the model prediction are essential for supporting water management decisions, as they provide guidance about the weight that should be given to the model (Beven and Binley, 1992), and for indicating where the model needs to be improved. Consideration of uncertainty in the decision process attempts to recognize the lack of knowledge and provides the decision maker with additional information on which to base the decision. Therefore, water managers need to communicate these uncertainties clearly so that decisions can be made with this additional knowledge and understanding.

1.2. Uncertainty issues in hydrological modelling

In practice, hydrologists are interested in either the uncertainty in the design variable or in the uncertainty in the forecasted variable, which are both related to the model output. Additionally, hydrologists are also interested in the reliability of the model parameters and the observed variables. Montanari (2007) suggests to keep uncertainty assessment in the observed variables and model parameters on one hand well distuinguished from uncertainty assessment in the model output. Although uncertainty assessment has received more attention over the last years, a significant part of the community is still not convinced to include uncertainty estimation in hydrological modelling because (i) mature guidance on methods and applications does not exist (Pappenberger and Beven, 2006) and (ii) evaluation of uncertainty in the prediction of environmental models is computationally very expensive (Beven, 2006a). This has prevented uncertainty assessment from becoming a standard practice in hydrological modelling such that it remains common to show model results without uncertainty bounds to decision makers (Pappenberger and Beven, 2006). Moreover, uncertainty assessment is usually conceived as an end-of-pipe analysis whereas it should be integrated in the modelling process (starting with problem definition and identification of modelling objectives) (Refsgaard et al., 2005).

It is clear that dealing with uncertainty is not evident and requires additional research. The main bottlenecks in uncertainty analysis are:

- Uncertainty is traditionally dealt with by using probability theory. However, there is a growing awareness among modellers that statistical information is one of the possible forms of information and that uncertainty should be dealt with in a more broader perspective through non-statistical approaches (Zadeh, 2005). The latter approaches are very useful for hydrological modelling because much human reasoning about hydrological systems is possibilistic rather than strictly probabilistic: we reason about whether a given scenario could happen without attaching probabilities to the likelihood that the scenario will happen (Montanari, 2007). Conventional uncertainty analyses also focus on quantifiable uncertainties, while many uncertainties cannot be quantified or remain difficult to quantify (Ayyub, 2001), e.g. when only a limited amount of information is available (*cfr.* prediction in ungauged basins), or when the input data consist of single non-repeatable values. In such situations, expression of uncertainty in terms of probability might not be possible (Montanari, 2007).
- There is a lack of coherent terminology and a lack of guidelines in identifying the best method for uncertainty propagation, because systematic testing of the high number of available methods is practically impossible (Montanari, 2007). The choice of the uncertainty propagation method is case-dependent and

affects the obtained prediction uncertainty. However, the lack of guidelines on the choice of an appropriate method makes it difficult to compare and interpret results in an objective manner. This supports an urgent need to clearly classify the practical problem to be solved and the method to be used in which selection of the method is based on the purpose of the application and on the type of information available (Montanari, 2007). In this perspective, Pappenberger et al. (2006a) have developed a decision tree that guides the user to a small set of possibly suitable methods, a number of competing methods within different philosophical frameworks that might be more or less appropriate in the given situation with its model dimensionalities and sources of uncertainty (Pappenberger and Beven, 2006).

- The concept of uncertainty is conceived and understood in a variety of ways by different communities and by different people (Pappenberger and Beven, 2006). This is an important point for the modeller since misunderstanding of the uncertainty in the modelling results might lead to a loss of credibility and confidence of stakeholders in the modelling process (Beven, 2006b). In some cases, it is not clear how to cope with the prediction uncertainty e.g. when it leads to wide ranges of predictions such that the uncertainty for outcomes of different scenarios is larger than the differences between the expected values of those scenarios. Research on the communication of uncertainty to decision makers has resulted in a guidance for uncertainty assessment and communication in hydrological applications, which has been published by the Netherlands Evironmental Assessment Agency (RIVM/MNP) (van der Sluijs et al., 2003). Although uncertainty analysis is receiving increasing attention in hydrological modelling, efficiently transferring the knowledge about uncertainty from the scientist to the end-user is still difficult (Montanari, 2007).

A logical step following the additional knowledge on uncertainty in the model output is to constrain this uncertainty. Beven (2006b) states that the solution for this goal is very obvious but that the practical achievement is very difficult. Furthermore, he argues that hydrologists should deal with this uncertainty to advance science. Reducing the uncertainty and enhancing the predictability can be achieved by (i) improving the understanding of physical processes built into model structures and improving their representation, (ii) improving the characterisation of the particularities of places (iii) reducing the uncertainty in model inputs and boundary conditions, and (iv) improving the value of conditioning data (Beven, 2006a). Most of these aspects imply the acquisition of more informative and higher quality data, both in space and time (Liu and Gupta, 2007).

With this dissertation we try to contribute to some of the listed issues concerning uncertainty analysis with a focus on the development of methods for uncertainty assessment, strategies for uncertainty reduction and ways to reduce the computational

effort. The specific research objectives and questions are formulated below.

1.3. Research objectives, questions and overview

This dissertation has emerged from the problems regarding uncertainty in hydrological models, as discussed in the previous section, and more specifically in spatially distributed hydrological models. The focus is on (i) uncertainty related to land surface characterization, *i.e.* land cover and soil inputs and parameters and (ii) simulations in poorly gauged or ungauged basins. Under the latter condition, there is limited information about the land surface characteristics available, usually consisting of readily available maps (provided by the local government or by an international organisation), remote sensing data (nowadays there is a variety of easily accessible sources) and literature-based data. These different types of information imply different types of uncertainty that should be dealt with and modelled in a different way. Although a lot of uncertainty in hydrological modelling is epistemic (due to a lack of knowledge), this dissertation mainly deals with aleatoric uncertainty (due to natural variability) that can be modelled in a probabilistic way.

The spatially distributed and physically-based hydrological TOPLATS model (Famiglietti and Wood, 1994, Peters-Lidard et al., 1997, Pauwels and Wood, 1999) is used in this dissertation in order to analyze the impact of uncertainty in the land surface characterization. The study area for which the model is applied is a small basin that is mainly covered by agricultural area and is situated in the northern part of Belgium. For this given scenario, some research objectives (Ob) and research questions (Q) are formulated:

- Ob1: Overview of the basic concepts and techniques for uncertainty and sensitivity analysis.
  - Q1: Which aspects are important when setting up a framework for uncertainty or sensitivity analysis?
- Ob2: Calibration of the spatially distributed hydrological model.
  - Q2: Does the selection of the calibration method have an influence on the calibration result?
- Ob3: Evaluation of the impact of uncertainty in the land cover input on the hydrological model predictions.
  - Q3: Is more detailed information about the land cover valuable for the prediction of water fluxes?
  - Q4: How can classification uncertainty efficiently be assessed?

  - Q5: How can the dimensionality of remote sensing data be optimized w.r.t. the classification uncertainty?
  - Q6: Is propagation of the classification uncertainty relevant for the predictability of water fluxes and how can it be performed in a computationally efficient way?

- Ob4: Evaluation of the impact of uncertainty in the soil input on the hydrological model predictions.
  - Q7: How can the within-soil-class variability of the soil hydraulic parameters be modelled and propagated through the hydrological model?
  - Q8: How can a sensitivity analysis on soil texture be performed?
  - Q9: Which strategies for sampling soil texture can be applied to reduce the uncertainty in the soil moisture predictions?

Throughout this dissertation, the research questions are addressed and answers are formulated. The first research question (Q1) is addressed in Part I, which outlines the general concepts of uncertainty (Chapter 3) and sensitivity analysis (Chapter 4) and gives a brief overview of the available techniques together with some applications in hydrological modelling (Ob1). The literature review allows to set-up an appropriate framework for uncertainty and sensitivity analysis for the given problem(s) in the subsequent parts.

Part II describes how the distributed modelling of the catchment hydrology is performed in order to deal with the second research question (Q2). It starts with a detailed description of the hydrological model structure (Chapter 6), followed by a description of the catchment for which the modelling is performed and a description of the available data (Chapter 7). In Chapter 8, a calibration of the hydrological model for the study area under consideration is performed (Ob2) in which it is investigated whether different calibration methods perform equally well.

Research questions Q3-Q6 are addressed in Part III that discusses on the impact of uncertain land cover input on the hydrological model output (Ob3). In Chapter 10, it is questioned whether it is useful to collect information on the crop types for the simulation of the water fluxes (Q3). Chapters 11 and 12 deal with uncertainty in the land cover map obtained through classification of remotely sensed data of which the former chapter focusses on the assessment of the classification uncertainty (Q4-Q5) and the latter chapter focusses on the propagation of the classification uncertainty through the hydrological model (Q6).

Research questions Q7-Q9 are addressed in Part IV that deals with the impact of uncertain soil information on the prediction of soil moisture (Ob4). If a soil map of the basin is available, but measurements of the soil parameters are lacking, the latter need to be estimated based on the texture information as provided by the soil map. In this case, the soil parameters are uncertain due to variability

within the soil classes and induce uncertainty in the predicted soil moisture. In Chapter 14, a methodology is presented to assess these uncertainties with special attention for some practical aspects (Q7). Because the soil parameters are usually derived from soil texture, a sensitivity analysis method is developed in Chapter 15 in order to evaluate how small variations in the soil texture affect the soil moisture predictions (Q8). Based on the results of the sensitivity analysis, some guidelines about texture sampling are formulated with the intention to efficiently reduce the prediction uncertainty (Q9).

Part I

General concepts of uncertainty and sensitivity analysis

2

Introduction

Models are always simplifications of the real world and necessarily involve a series of assumptions and approximations to compensate for our incomplete understanding of the real world. In addition, precise model input data are rarely, if ever, available. This implies that the reliability of the model output is ultimately uncertain (*cfr.* Figure 2.1). Knowledge about the model output uncertainty is crucial for decision-making and for a meaningful interpretation of the model result. If the uncertainty is costly or has a high impact on the decision to be made, it is desirable to reduce the uncertainty in the model output by reducing the uncertainty in the model inputs that are considered as most important. Prior to the uncertainty reduction, an uncertainty and sensitivity analysis is carried out in order to formulate an answer to the questions 'How uncertain is the model output?' and 'Where is the uncertainty coming from?'.

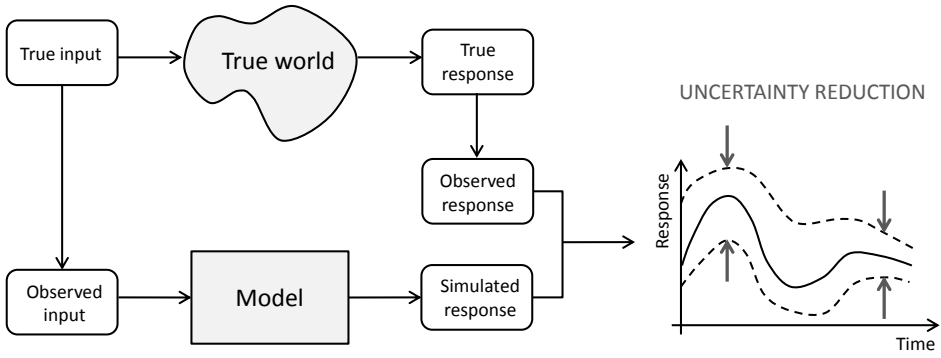


Figure 2.1: General scheme that illustrates the establishment of model prediction uncertainty arising from uncertainty in the description of the real world and in the model input data (modified from Wagener and Gupta (2005)).

Uncertainty analysis (UA) studies how the input uncertainty affects the model output by quantifying the uncertainty in the model output arising from uncertainty in the model inputs. The main objectives of UA are (i) to explore possibilities for model simplification, (ii) to explore ways to improve the model predictions and (iii) to judge if the estimated output uncertainty is acceptable for the given model application (Jansen, 1998). The basic principles and techniques for uncertainty

analysis are outlined in Chapter 3.

Sensitivity analysis (SA) studies how the uncertainty in the model output can be allocated to uncertainty in the model inputs by quantifying the sensitivity of the model output to systematic changes in the model input. The sensitivity is usually expressed by means of a sensitivity measure and indicates the importance of the model input. The main objectives of SA concern the desire to determine (i) which input contributes the most to the output variability, (ii) if and which input interactions are present, (iii) which inputs are insignificant (these can then be eliminated or held constant) and (iv) which regions within the parameter space are optimal for use in subsequent calibration studies (Ascough et al., 2005). SA is complementary to UA as it can be used to prioritize additional data collection or research, based on the identification of the most important uncertainties (Frey and Patil, 2002). Techniques for sensitivity analysis are discussed in Chapter 4.

The computational model that expresses the relationship between the input variables X_i and the output variable Y, is represented by a mathematical function f, such that:

$$y = f(\mathbf{x}) = f(x_1, x_2, ..., x_n) \tag{2.1}$$

with y the value of the model output, x_i (with $i \in \{1, 2, ..., n\}$) the value of the i-th model input and n the number of input variables. A general scheme of a UA and SA for model f is shown in Figure 2.2. This shows that UA and SA can be performed jointly by executing the model repeatedly for different combinations $\mathbf{x}$ of the input variables. Through UA, an empirical distribution of the model output Y is obtained, whereas through SA the variance of the model output is partitioned to the various model inputs X_i. In the next chapters, the techniques for UA and SA are presented with a focus on the model input, although the techniques are also applicable to other sources of uncertainty, such as parameter uncertainty.

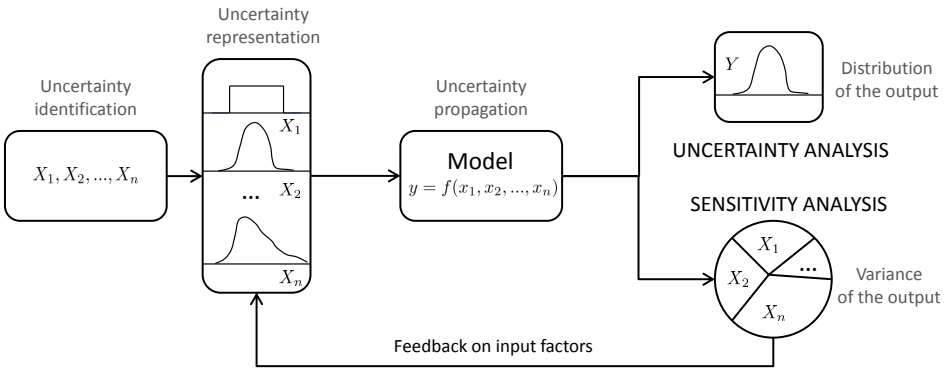


Figure 2.2: General scheme of uncertainty analysis and sensitivity analysis with a focus on input uncertainty (modified from Crosetto et al. (2000)).

3

Uncertainty analysis

Usually, uncertainty analysis involves three main steps (Crosetto et al., 2001b):

1. *Uncertainty identification*: identifying all sources and types of uncertainty present in the modelling study (discussed in Section 3.1).
2. *Uncertainty representation*: selecting an appropriate tool to represent the uncertainty in function of the type and source of uncertainty (discussed in Section 3.2).
3. *Uncertainty propagation*: propagating the input uncertainty through the model in order to quantify the uncertainty in the model output (discussed in Section 3.3).

A schematic of the different steps in the uncertainty analysis of the model output uncertainty w.r.t. the uncertainty in the input variables is given in Figure 2.2.

3.1. Uncertainty identification

The term uncertainty seems a commonly and well understood everyday word. Nevertheless, a unique definition of uncertainty is hard to find. Uncertainty is mostly considered as the antonym of certainty, such that uncertain situations are defined as situations that do not satisfy the definition of certainty: 'Certainty implies that a person has quantitatively and qualitatively the appropriate information to describe, prescribe or predict deterministically and numerically a system, its behaviour or other phenomena' (Zimmermann, 1997). Definitions of uncertainty range from 'a situation of inadequate information, which is due to either inexactness, unreliability or border with ignorance' (Funtowicz and Ravetz, 1990) to 'the deviation from the unachievable ideal of completely deterministic knowledge of the relevant system' (Walker et al., 2003). In this dissertation we follow the definition of uncertainty as given by Brown (2004), because this definition specifically aims at applications in model-based water management. Here, uncertainty is defined as 'a degree of confidence about the model outcome(s)' (Brown, 2004). This is consistent with the view of Zadeh (2005) who considers uncertainty as an attribute of information.

Identification of all uncertainties present in the modelling study is a crucial step in the uncertainty analysis (Refsgaard et al., 2007) as it simplifies the uncertainty representation (*cfr.* Section 3.2) that follows the uncertainty identification. Different types of uncertainty warrant different approaches of uncertainty representation. A wrong assumption regarding the nature of the uncertainty may seriously bias the results of a risk analysis (Baudrit et al., 2007). Ideally, the uncertainty identification leads to a comprehensive list of unique and complementary uncertainties. Over the past years, several frameworks have been proposed to identify and classify uncertainties (e.g. Walker et al., 2003, Brown, 2004, van Asselt and Rotmans, 1996, Funtowicz and Ravetz, 1990, Klir and Folger, 1988, Reddy and Haldar, 1992, Plate and Duckstein, 1987, Bernier, 1987). In this dissertation, uncertainty is characterized based on (i) the type of uncertainty and (ii) the source of uncertainty. The type of uncertainty is related to the nature, the origin of the uncertainty, e.g. natural variability. The source of uncertainty indicates the factor which is causing the uncertainty in the model output, e.g. parameter uncertainty. The source and type of uncertainty can be represented as a two-dimensional concept in an uncertainty matrix, which is a useful tool to get an overview of the uncertainties present in the modelling study.

3.1.1. Types of uncertainty

In engineering modelling, the type of uncertainty is usually classified based on the nature (the origin) of the uncertainty, *i.e.* whether the uncertainty is due to imperfect knowledge or random variability (Jacquin, 2010, Ferson and Ginzburg, 1996):

- *Epistemic uncertainty*: uncertainty associated with quantities that are totally deterministic but anyway ill-known because they pertain to the future or because of imperfection of our knowledge (Dubois, 2006). This type of uncertainty is related to many aspects of modelling, e.g. limited and inaccurate data, imperfect representation of processes in a model. It is associated with a subjective view of uncertainty as the properties of the uncertainty are linked to the knowledge of the model user (which might differ from the knowledge of another model user) about the exact nature of the real world and are derived from the information that is available to the model user (Smets, 1999). By improving the knowledge about the system, collecting additional data or by improving the data analysis, the epistemic uncertainty can be reduced (Jacquin, 2010).
- *Aleatoric uncertainty*: uncertainty arising from randomness due to natural variability of observations resulting from heterogeneity (spatial variability) or from fluctuations of a quantity in time (temporal variability) (Baudrit and Dubois, 2006). This type of uncertainty is associated with an objective view

> of uncertainty because of the fact that the likelihood of an event is considered to be independent of the modeller's opinion about the occurrence of the event. The randomness is then an objective property of the experimental set up that generates the event (Smets, 1999). The fact that variability is treated as an inherent property of the real system implies that aleatoric uncertainty cannot be reduced through improved insights about the underlying natural processes (Jacquin, 2010).

The distinction between epistemic and aleatoric uncertainty is not always clear and both types of uncertainty do not form exclusive sets. There exists an overlapping region where variability can also be considered as epistemic when the information to describe the variability of a deterministic quantity is incomplete, e.g. a model parameter which is known to vary within a certain range, but with unknown frequency distribution. One may also have incomplete knowledge about the variability of a non-deterministic quantity if the observations were poor or if only expert knowledge is available (Dubois, 2006). Moreover, measured quantities can include both epistemic and aleatoric uncertainty as inappropriate measurement methods introduce epistemic uncertainty while repeated measurements by other methods or persons introduce aleatoric uncertainty due to randomness in the measuring. Another example of ambiguity in the uncertainty is the prediction of a 100 year flood event on basis of existing flow data. The epistemic uncertainty in the prediction can be reduced by additional monitoring or improved data analysis, but there will always be aleatoric uncertainty due to the chaotic nature of the weather (Refsgaard et al., 2007).

An entire spectrum of different levels of knowledge exist, ranging from complete deterministic understanding at one end of the scale and total ignorance at the other (Walker et al., 2003). Determinism is the state of knowledge in which we know everything precisely. Another extreme state of knowledge is the state of total ignorance, which means that one does not even know that he/she doesn't know. In this grey zone between what is well known and what is not known, uncertainty (imperfect knowledge) can be situated and is reflected by the situation where the likelihood of events is ill-known, except maybe for a lower and an upper bound (Walley, 1991). This implies that, paradoxically, to be uncertain one has to know something because uncertainty can only be ascribed to known processes and known state variables (Hoffmann-Riem and Wynne, 2002). Likewise, it is impossible to be uncertain about something of which one is completely ignorant because ignorance is related to unknown procesess and unknown state variables.

3.1.2. Sources of uncertainty

According to Scholten et al. (2007), the modelling process can be decomposed into five major steps: (i) model study plan (identify the problem, define requirements),

(ii) data and conceptualisation (collect and process data, develop conceptual model, select model code), (iii) model set-up (construct model, assess performance criteria), (iv) calibration and validation, (v) simulation and evaluation (model predictions, review and dialogue). Each step (or most of them) is associated with an amount of uncertainty, on the basis of which the source of uncertainty can be classified into five possible categories (Walker et al., 2003, Refsgaard et al., 2007):

- *Context uncertainty*: uncertainty associated with the boundaries and the representation of the system to be modelled. It is related to the assumptions and choices underlying the model, e.g. the identification of the boundaries of the system to be modelled, external social and economic circumstances, the choice of a particular kind of model, choices about the range of model application.

- *Input uncertainty*: uncertainty associated with the system data (geographical data and time-dependent driving forces) that drive the model and typically give a description of the system to be modelled (e.g. soil map, landcover map) and uncertainty associated with the external forces that have an influence on the system and its performance (e.g. scenario and policy variables). Input uncertainty results from e.g. measurement uncertainties, uncertainties due to scaling, and uncertainties in outcomes of preceding models.

- *Model structure uncertainty*: uncertainty that arises from the model itself. It arises from the fact that the model is a more or less crude simplification of the reality in nature, and from an incomplete understanding of the modelled processes. The relations between the model components inside the model are regarded as model structure, whereas the decisions on which processes to include in the model are regarded as model context.

- *Parameter uncertainty*: uncertainty associated with the parameter values, e.g. methods and data used to calibrate the model parameters. Parameters are supposed to be invariant in the chosen context and algorithmic representation, and do not change as an effect of causal relations within the model. They differ from input because they do not refer to the geographical location and the period to be modelled.

- *Model technical uncertainty*: uncertainty arising from software implementation of the model and numerical implementation of the algorithms, e.g. numerical approximations and software errors. Constants that control the software implementation of the model such as time step and grid size are regarded as model technical aspects.

By propagating the above sources of uncertainty through the model (*cfr.* Section 3.3), the model outcome uncertainty, *i.e.* the total uncertainty associated with the model outcomes of interest, can be assessed. The complexity of the model determines how the model outcome uncertainty is related to the model structure

uncertainty and the parameter uncertainty. In a simple model with few parameters, the outcome is dominated by model structure uncertainty. In a complex model with many parameters, the outcome is dominated by parameter uncertainty in case the calibration data do not contain sufficient information to adequately calibrate the model parameters. As such, increasing the complexity of the model may increase the uncertainty in the model outcomes. An optimal model complexity can be identified as a function of the data available to calibrate the model parameters (Harremoes and Madsen, 1999). Although there is a natural tendency for models to grow in complexity, in order to better represent the processes of the system, models are mathematically ill-posed because the available information does not allow a single unambigious solution to the problem (Beven, 2006a). This leads to the concept of equifinality that focuses on the fact that there are many acceptable representations of reality, such that models with different complexity might result in good reproductions of the system response (Beven, 1993). For techniques to deal with equifinality, we refer to Section 3.3.1.

3.2. Uncertainty representation

Altough several approaches to represent uncertainty are available, the probability theory is the most popular. It was the primary tool for uncertainty representation in mathematical models and relates the uncertainty of an event to its frequency. However, it is not always meaningful to relate uncertainty to frequency, because some types of events can be rare, unrepeatable, or statistical data may simply be unavailable (Smets, 1997). In some occasions it is necessary to appeal to another class of information, e.g. experience, intuition, ideology, beliefs, feelings, which all have a degree of subjectivity, but not of arbitrariness (Buchanan et al., 1998). Alternative non-probabilistic approaches to represent uncertainty were therefore developed. One approach was initiated by Zadeh (1965) who proposed the idea of fuzzy sets and later also introduced the possibility theory. Other non-probabilistic models are based on the Dempster-Shafer theory of evidence (Dempster, 1967, Shafer, 1976), in which uncertainty is expressed as a degree of belief.

The choice of an appropriate theory for uncertainty representation depends on the type of uncertainty, as identified in Section 3.1. Aleatoric uncertainty is best represented using probability distributions, whereas epistemic uncertainty is more faithfully represented using families of probability distributions (Walley, 1991). In practice, families of probability distributions are encoded either by probability boxes (Ferson et al., 2003), *i.e.* upper and lower cumulative distribution functions, by possibility distributions (Dubois et al., 2000) or by belief functions (Shafer, 1976). In the remainder of this chapter, the basic principles and properties of the probability theory, fuzzy set theory and possibility theory are briefly discussed.

3.2.1. Probability theory

In probability theory, uncertainty is represented as a probability density function (PDF). Based on the PDF, the probability that the value of variable X belongs to the interval $[u_1, u_2]$ is determined as:

$$P(u_1 \leq X \leq u_2) = \int_{u_1}^{u_2} f_X(x)dx \tag{3.1}$$

with x the value of variable X and f_X the PDF of variable X. A cumulative form of the PDF is called a cumulative distribution function (CDF) and expresses the probability that the variable X takes a value less than or equal to u:

$$F_X(u) = P(X \leq u) = \int_{-\infty}^{u} f_X(x)dx \tag{3.2}$$

An example of a PDF and CDF is shown in Figure 3.1. From the CDF, confidence intervals can be derived in order to give an estimated range of values that is likely to include the unknown true value of variable X. Usually, the confidence level a is chosen at 0.90, 0.95 or 0.99 and the confidence interval is placed symmetrically around the mean. The $a \cdot 100\%$ confidence interval is the interval $[u_1, u_2]$ in which the value of variable X is included with a probability a, such that:

$$a = \int_{u_1}^{u_2} f_X(x)dx \tag{3.3}$$

The lower and upper confidence limits are derived from the CDF by using the quantile function:

$$\begin{cases} u_1 = \inf\{x \in X : (1-a)/2 \leq F_X(x)\} \\ u_2 = \inf\{x \in X : 1-(1-a)/2 \leq F_X(x)\} \end{cases} \tag{3.4}$$

The construction of a confidence interval with confidence level a is illustrated in Figure 3.1 for variable X.

The probability measure satisfies following conditions: (i) the probability of an event is always non-negative, (ii) the probability of the entire sample space equals one, (iii) if two events are mutually exclusive, then the probability of their union is equal to the sum of their probability. The latter condition refers to the self-dual character of probability, which means that the negation of an event E (a subset of the domain of X) is improbable if and only if the event is probable, or thus (Dubois et al., 2000):

$$P(E) = 1 - P(E^c) \tag{3.5}$$

where E^c is the complement of E. Concerning the interpretation of probability, there are two main views: the frequentist view and the Bayesian view.

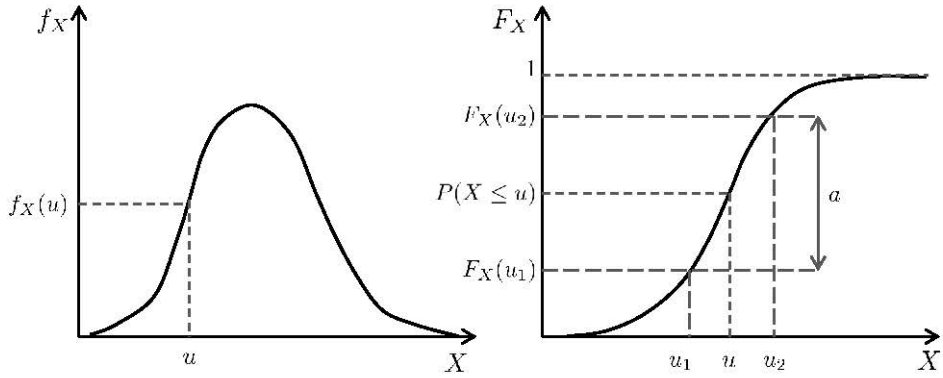


Figure 3.1: Example of a probability density function (PDF) (left) and a cumulative distribution function (CDF) (right), with indication of a $a \cdot 100\%$ confidence interval $[u_1, u_2]$.

Frequentist view of probability

The frequentist approach relates to the situation where an experiment can be repeated indefinitely under identical conditions, but where the observed outcome is random. It defines probability as the relative frequency of a particular observation, *i.e.* the number of times an event occurs divided by the total number of observed events. This frequency definition implies that statistical inference is based on data that are collected from a defined population, and asserts that if the number of repetitions becomes infinite, the frequency represents the true value of the probability (Maskey, 2004).

Bayesian view of probability

In the Bayesian approach, the probability of a particular event E is subjective as it is conceived as the individual's degree of belief that the event will occur (Maskey, 2004). This means that the probability is conditional on the past experience, it is assigned on basis of the current state of knowledge and is updated when new information becomes available. The procedure for updating the a priori probability is based on the Bayes rule (Bayes, 1763):

$$P(E|D,B) = \frac{P(E|B)P(D|E,B)}{P(D|B)} \tag{3.6}$$

The term $P(E|D,B)$ is the a posteriori probability, which gives the probability of the event E after considering the additional evidence D in the given context (background information) B. The term $P(E|B)$ is the a priori probability, which gives the probability of the event E given B alone, *i.e.* before the evidence is considered. The term $P(D|E,B)$ is the likelihood, which gives the probability of the evidence assuming that the event occurs and the context is true. The denominator $P(D|B)$ is simply a normalizing constant.

Application of the probability theory as a quantitative model for randomness and indecisiveness (Dubois, 2006) should be well-considered. For example, when it is only known that the value x of variable X lies in an interval $[u_1, u_2]$, adopting a uniform distribution to express ignorance about X introduces information that is in fact not available (Ferson and Ginzburg, 1996). The uniform distribution is just one option among an infinity of probability distributions with support $[u_1, u_2]$. In such case, it is better to express the knowledge on x by a family of probability distributions (Baudrit and Dubois, 2006).

3.2.2. Fuzzy set theory

In fuzzy set theory, as introduced by Zadeh (1965), uncertainty is represented by means of fuzzy sets. In classical (crisp) set theory, the transition between membership and non-membership is abrupt and well-defined with a membership 1 (certainly belongs to the set) or 0 (certainly does not belong to the set). On the contrary, fuzzy sets include boundaries that are imprecise and vague such that they allow for a transition between membership and non-membership. Therefore, fuzzy sets are useful to describe subjective knowledge which makes them adequate representations of epistemic uncertainty.

A fuzzy set A is defined as a function from the universe I to the unit interval $[0, 1]$ that maps an element x to $A(x)$ (Zimmerman, 1996):

$$A : I \to [0, 1] : x \to A(x) \tag{3.7}$$

where $A(x)$ is the membership degree of x in A, and is conceived as a degree of belief to which the element belongs to the set. Although in principle any function of the form $A(x)$ describes a membership function, fuzzy sets are mostly represented by parameterized functions. Commonly used are trapezoidal and triangular membership functions, of which an example is given in Figure 3.2. A trapezoidal membership function is characterized by four parameters and is defined as:

$$A(x) = \begin{cases} 0 & \text{if } x < u_1 \\ \dfrac{x - u_1}{u_2 - u_1} & \text{if } x \in [u_1, u_2] \\ 1 & \text{if } x \in [u_2, u_3] \\ \dfrac{u_4 - x}{u_4 - u_3} & \text{if } x \in [u_3, u_4] \\ 0 & \text{if } x > u_4 \end{cases}$$

A triangular membership function is obtained by adapting the above equation such that $u_2 = u_3$. Alternatively, a fuzzy set A can be described by referring to its features:

- The *support* of a fuzzy set contains all elements of X that belong to the fuzzy set A with a membership level different from zero:

$$supp(A) = \{x \in X | A(x) > 0\} \tag{3.8}$$

- The *core* of a fuzzy set contains all elements of X that completely belong to the fuzzy set A:

$$core(A) = \{x \in X | A(x) = 1\} \tag{3.9}$$

- The α-*cut* of a fuzzy set contains all elements of X for which the level of membership exceeds the threshold level α:

$$A_\alpha = \{x \in X | A(x) \geq \alpha\} \tag{3.10}$$

These features are depicted in Figure 3.2 for a trapezoidal fuzzy set.

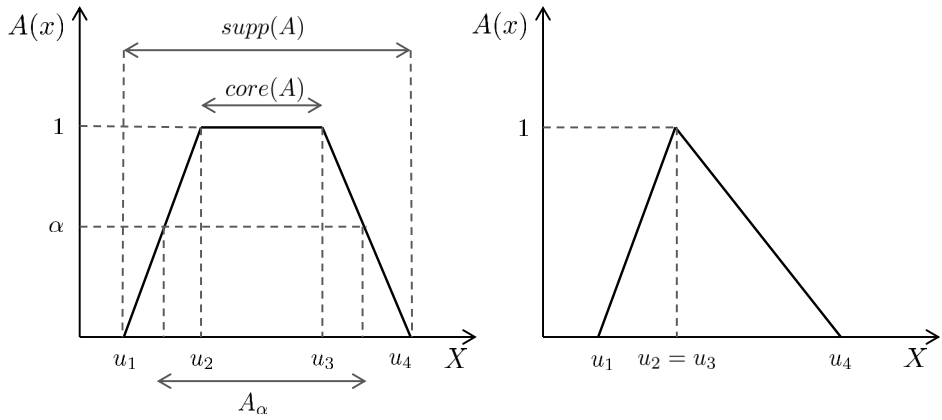


Figure 3.2: Example of a trapezoidal (left) and triangular (right) membership function $A(x)$ that maps every element of the universe to the interval [0,1].

3.2.3. Possibility theory

Possibility theory was developed as an approach to describe uncertainty based on vague linguistic information using fuzzy sets (Zadeh, 1999). It offers a qualitative model of incomplete knowledge (Dubois, 2006) and stands at the crossroad between probability and fuzzy set theory. As such, there is a parallellism between the different theories, especially between probability and possibility. For a detailed overview of the differences and similarities between probability, fuzzy set and possibility theory, we refer to Dubois et al. (2000).

In possibility theory, knowledge about a variable X is expressed in a possibility

distribution π_X which is a mapping of X to the unit interval [0,1]. Formally a possibility distribution coincides with the membership function of a fuzzy set of X (Zadeh, 1965), but in which the possibilistic entities have different semantics than the fuzzy entities (Dubois et al., 2000). The possibility distribution π_X has following properties: (i) it is characterized by its core and support (Figure 3.3), respectively the set of values of X for which $\pi_X(x) = 1$ and $\pi_X(x) > 0$, (ii) it does not have to sum to unity (as it is context-dependent) and (iii) it constraints the possible values of X with the following conventions:

- $\pi_X(x) = 0$ if it is definitely impossible that $X = x$ because X cannot take the value x.
- $\pi_X(x) = 1$ if it is definitely possible that $X = x$ because nothing prevents X to take the value x.
- $\pi_X(u_1) > \pi_X(u_2)$ if u_1 is a more plausible value for X than u_2.

In the possibility framework, situations of partial ignorance can be described: it is natural to consider that both an event and its opposite are possible, whereas in a probabilistic setting, if an event is highly probable, then the opposite one must be highly improbable. Possibility is therefore not self-dual, which implies that the statement 'it is possible that an event occurs' is not equivalent to 'it is not possible that the opposite event occurs'. Moreover, extreme forms of partial knowledge can be captured: complete knowledge (for state u, $\pi_X(u) = 1$ whereas for all other states x, $\pi_X(x) = 0$), complete ignorance ($\pi_X(x)$=1 for all states x).

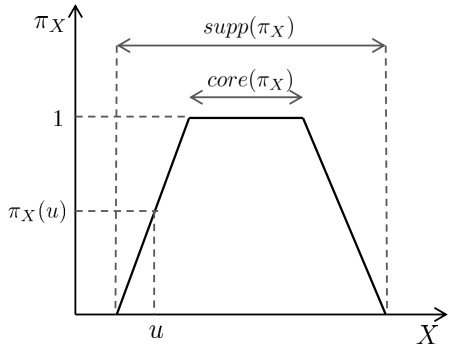


Figure 3.3: Example of a possibility distribution with indication of the core and the support of the distribution.

In order to characterize the state of knowledge (partial belief) about an event E, possibility measures (as limited cases of confidence measures) are derived from the possibility distribution. These measures are the possibility Π and the necessity N,

respectively defined as:

$$\Pi(E) = \sup\left(\pi_X(x)\right),\ x \in E \tag{3.11}$$

$$N(E) = 1 - \Pi(E^c) = \inf(1 - \pi_X(x)),\ x \notin E \tag{3.12}$$

where E^c is the complement of E. The possibility evaluates to what extent event E is consistent with the knowledge expressed by π_X, while the necessity evaluates to what extent event E is certainly implied by the knowledge. The possibility measures for the event $X \leq u$ are depicted in Figure 3.4. A high possibility does not necessarily imply a high probability. However, if an event is not possible, it is also not probable. Possibility is therefore considered as the upper bound for probability (Zadeh, 1999). This weak connection between both theories is called the possibility-probability consistency principle. Based on this principle, a possibility measure can be transformed into a probability measure and vice versa. Turning a probability measure into a possibility measure may be useful in the presence of other weak sources of information, or when computing with possibilities is simpler than computing with probabilities (Jamison and Lodwick, 2002). For more details on possibility-probability transformations, we refer to Dubois et al. (2004).

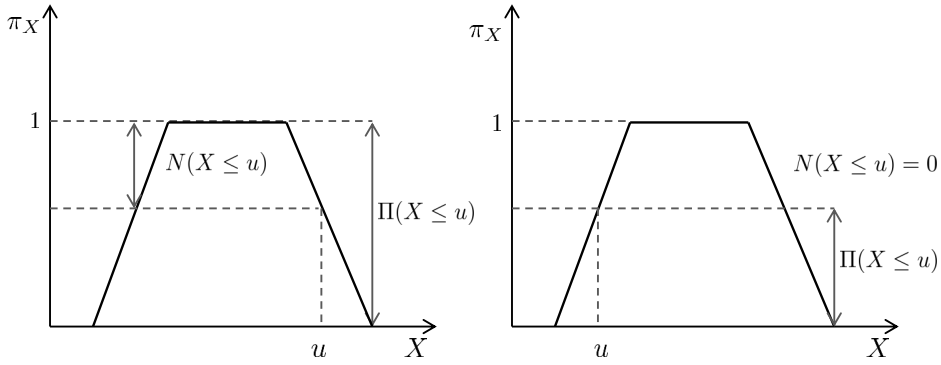


Figure 3.4: Illustration of the possibility Π and the necessity N for the event $X \leq u$.

Similarly to probability, possibility can be interpreted in different ways. There are four ideas that can be conveyed by the word 'possibility' (Dubois et al., 2000):

- **Possibility as feasibility**
  In this interpretation, possibility is conceived as an achievement, a solution to a problem. The semantic is expressed by means of sentences as 'it is possible to solve the problem'.

- **Possibility as plausibility**
  Here, possibility refers to the propensity of an event to occur. At the linguistic level, this meaning is expressed by sentences as 'it is possible that the train arrives on time'.

- **Possibility as consistency**
  In this interpretation, possibility means that the occurrence of an event does not contradict with the available information. E.g. 'it is possible that the man left earlier because he didn't feel well'.

- **Possibility as deontic standard**
  Following this semantic, an event that is possible means that it is allowed, it is permitted by the law that the event takes place. E.g. 'it is possible to put charge against the thief'.

In the different semantics, possibility can be understood as an objective or as a subjective notion, either within a qualitative or quantitative setting (Dubois et al., 2000). From an objective point of view, possibility is understood as the physical ability of an event to occur, whereas in the subjective view, possibility is conceived as the degree of potential surprize that an event occurs. In contrast with the probability theory, possibility can be described both by means of numbers in the unit interval (quantitative setting), and by means of a suitable relation between events (qualitative setting), e.g. represented by possibility measures ranging on any totally ordered set.

3.3. Uncertainty propagation

Various methods for uncertainty propagation exist. The methods are usually grouped into two main classes w.r.t. the theory used to represent the uncertainty (*cfr.* Section 3.2): probabilistic propagation methods and non-probabilistic propagation methods. Among the probabilistic methods are (i) analytical methods, e.g. central limit theorem, (ii) sampling methods, e.g. Monte Carlo simulation, and (iii) approximation methods, e.g. First-Order Second Moment (FOSM). The analytical methods transform the PDF of the input variables into the PDF of the output variable using an explicit input-output relationship. As these methods are limited to very simple problems (e.g. linear functions), sampling or approximation methods are preferred for practical applications. Sampling methods provide an estimation of the PDF of the output variable, whereas the approximation methods only provide the moments (e.g. mean and variance) of the PDF. Propagation methods based on fuzzy set theory or possibility theory are less popular and are based on the extension principle. In this section, we will outline the basic principles of Monte Carlo simulation, the FOSM method and the extension principle. For a discussion on other techniques, we refer to Helton (1993).

3.3.1. Sampling methods

Monte Carlo simulation

Monte Carlo (MC) simulation is a sampling method for error propagation that is not built on assumptions upon the model structure. The technique is based on multiple evaluations of the model output Y with error-corrupted versions (or realizations) of the model input X_i, generated from the probability distribution of X_i. A Monte Carlo analysis usually consists of following steps:

1. Assign a stochastic error model (e.g. probability distribution) to each input variable X_i (with $i \in \{1, 2, ..., n\}$) (*cfr.* Section 3.2.1).

2. Generate a random number between 0 and 1. This value corresponds to $P(X_i \leq x_i)$.

3. Determine the realization of the model input variables X_i from their stochastic error model corresponding to the random number as generated in step 2. The result is a sample of the form $\mathbf{x_r} = (x_{1r}, x_{2r}, ..., x_{nr})$.

4. Evaluate the model output Y for realization r of the set of input values $\mathbf{x_r}$. The result is of the form $y_r = f(\mathbf{x_r}) = f(x_{1r}, x_{2r}, ..., x_{nr})$.

5. Repeat steps 2 to 4 for n_{MC} (defined by the model user) different random numbers such that $r \in \{1, 2, ..., n_{\mathrm{MC}}\}$.

6. Compute the empirical probability distribution of the model output from the set of n_{MC} model output values y_r (with $r \in \{1, .., n_{\mathrm{MC}}\}$). The PDF of the model output can be summarized in terms of its statistics, usually the expected value $E[Y]$ and the variance $V[Y]$:

$$E[Y] = \sum_{r=1}^{n_{\mathrm{MC}}} \frac{y_r}{n_{\mathrm{MC}}} \tag{3.13}$$

$$V[Y] = \sum_{r=1}^{n_{\mathrm{MC}}} \frac{(y_r - E[Y])^2}{n_{\mathrm{MC}} - 1} \tag{3.14}$$

   Note that a characterization of the uncertainty by means of $E[Y]$ and $V[Y]$ reduces all information about the variability in Y in these two statistics.

The procedure for MC analysis is illustrated in Figure 3.5 for a two-dimensional problem. The generation of the samples in step 2 from the analysis can be performed using different procedures. A random sampling technique is commonly used although it may not be efficient to summarize the probability distribution. The sampling can result in a clustering of sampled values due to repeated sampling in the same region of the distribution and a lack of sampling in other regions of the distribution. On the contrary, a stratified sampling approach ensures a more even coverage of the range of values with the same number of simulation runs. Examples

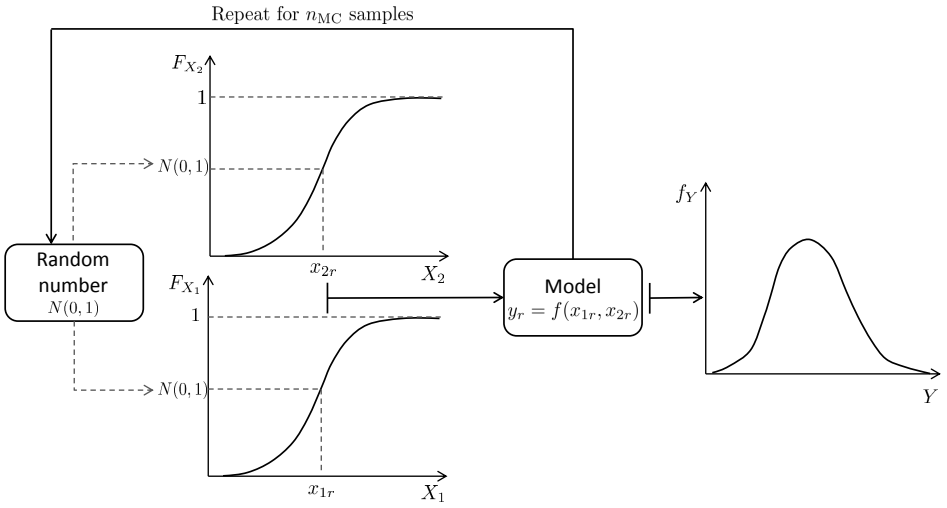


Figure 3.5: Schematic of a Monte Carlo analysis of a model with two input variables X_1 and X_2.

of such sampling schemes are Latin Hypercube Sampling (LHS), which divides the entire sampling space into ranges of equal probability and takes one sample within each range (McKay et al., 1979, Helton and Davis, 2003), and Markov chain sampling, which is based on a random process to express the transition from one state to another (*i.e.* a Markov chain) (Hastings, 1970). Selecting the sampling pattern is a crucial step in the MC analysis as a good choice of the sample locations ensures an efficient uncertainty propagation.

Ultimately, the accuracy of the estimated PDF depends on the quality of the stochastic error model and on the number of Monte Carlo realizations. The larger n_{MC}, the less dispersed the estimates are. MC analysis therefore requires a high number of model runs to accurately describe the tails of the output PDF. The latter is important as the tails of the distribution are used to characterize the risk of extreme events and scenarios. The high number of model runs is the main disadvantage of the Monte Carlo technique (as compared to the moment propagation techniques, *cfr.* Section 3.3.2), however, the computational time can be considerably reduced through the use of an efficient sampling method.

Generalized Likelihood Uncertainty Estimation

The generalized likelihood uncertainty estimation (GLUE) is a simulation technique that was proposed by Beven and Binley (1992). The basic idea of GLUE is to reject the concept of one single optimal solution and adopts the idea of equifinality of models, parameters and inputs (Beven, 1993). Equifinality originates from the imperfect knowledge of the system to be modelled and assumes that all model structures, input sets and parameter sets have an equal likelihood of being accept-

able. With GLUE, the prior set of models, parameters and inputs is divided into a set of behavioural (acceptable) and non-behavioural (non-acceptable) solutions based on a subjective likelihood function. The GLUE methodology consists of following steps:

1. A set of k plausible model structures $\mathbf{f} = \{f_1, f_2, .., f_k\}$, a set of l parameter vectors $\mathbf{\Phi} = \{\boldsymbol{\phi_1}, \boldsymbol{\phi_2}, ..., \boldsymbol{\phi_l}\}$ and a set of n input variable vectors $\mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\}$ are considered. The statistics for the models, parameters and variables are defined based on all available knowledge. Typically, a uniform distribution is chosen to characterize the prior distributions in case there is little information about the uncertainties available.

2. A stochastic simulation based on the prior distributions as defined in step 1 is performed. Following the Monte Carlo procedure, a high number of system representations (simulators) is generated and the model is run for each realization to estimate the output $\mathbf{Y} = \{Y_1, Y_2, ..., Y_m\}$ (with m the number of output variables).

3. An evaluation of each simulation performed in step 2 is carried out. Based on a chosen likelihood measure L, the simulators are rated according to their performance, *i.e.* the degree to which their simulation result fits the observed data. The performance of the simulation with model structure f_k, parameterized with parameter vector $\theta_\mathbf{l}$ and forced with the input vector $\mathbf{x_n}$ is given by $L(f_k, \phi_\mathbf{l}, \mathbf{x_n}|\mathbf{Y}_{\text{obs}})$, in which $\mathbf{Y}_{\text{obs}}$ contains the observations of the output variables. The simulators that perform below a rejection criterion L_{crit} are discarded from the analysis and are considered as non-behavioural. Usually, the likelihood function is a model efficiency function, e.g. the Nash and Sutcliffe efficiency (Nash and Sutcliffe, 1970), but also fuzzy likelihood functions can be used. For a discussion on the choice of an appropriate likelihood measure, we refer to Beven and Freer (2001).

4. The likelihood measures of the retained, and thus bevioural simulators are rescaled in order to obtain a cumulative likelihood equal to one. Finally, a cumulative distribution function of the predicted output variables $\mathbf{Y}$ is constructed from which uncertainty bounds corresponding to a given confidence level are derived in addition to a median simulation (*cfr.* Section 3.2.1).

Bayesian Model Averaging

An alternative methodology that rejects the idea of a unique optimal representation of the system is Bayesian Model Averaging (BMA) (Draper, 1995, Hoeting et al., 1999). BMA combines predictions from multiple competing models to provide a more realistic description of the total prediction uncertainty by weighting the predictions of the competing models by their corresponding posterior probability. The latter represents the model's skill to reproduce the system reponse in the

observation period and is computed using Bayes' rule:

$$p(f_i|\mathbf{Y}_{\text{obs}}) = \frac{p(\mathbf{Y}_{\text{obs}}|f_i)p(f_i)}{\sum_{j=1}^{k} p(\mathbf{Y}_{\text{obs}}|f_j)p(f_j)} \tag{3.15}$$

where k is the number of competing model structures, $p(f_i)$ is the prior probability of the i-th model structure f_i and $p(\mathbf{Y}_{\text{obs}}|f_i)$ is the integrated likelihood of model structure f_i (parameterized with parameter vector ϕ and input vector $\mathbf{x}$), which provides a measure of how well the observations $\mathbf{Y}_{\text{obs}}$ support the model. The predictive distribution of the output variable Y is then given by (Hoeting et al., 1999):

$$p(Y|\mathbf{Y}_{\text{obs}}) = \sum_{i=1}^{k} p(Y|\mathbf{Y}_{\text{obs}}, f_i)p(f_i|\mathbf{Y}_{\text{obs}}) \tag{3.16}$$

An approximation to BMA is the Maximum Likelihood BMA, proposed by Neuman (2003), that relies on a maximum likelihood parameter estimation around which a Monte Carlo simulation is expanded and thus requires prior information about the model parameters.

Applications in hydrology

Monte Carlo simulation is a very popular technique to propagate parameter and input uncertainty in hydrological studies. Applications include for example, analysis of the impact of parameter uncertainties on the streamflow predictions by a water-balance hydrological model (Benke et al., 2008), quantification of the uncertainty in the hydrological response to land use change scenarios (Breuer et al., 2006, Eckhardt et al., 2003) and treatment of precipitation error during the calibration of a hydrological model (Vrugt et al., 2008). Also the GLUE methodology is popular among hydrological modellers. The reason for the popularity is that GLUE is easy to use and can account for all sources of uncertainty. It was used for example to assess the prediction limits of daily streamflow with the MIKE SHE model under conditions of data uncertainties (Vazquez et al., 2009), to evaluate the uncertainty in the prediction of soil moisture arising from uncertainty in the soil parameters (Hossain and Anagnostou, 2005) and to estimate the uncertainty in the inundation extent arising from uncertainty in the observed spatial information using a fuzzy likelihood measure (Pappenberger et al., 2007, Christiaens and Feyen, 2002). However, many modellers are reluctant to adopt the idea of equifinality in hydrological modelling because it is assumed that hydrological science is supposed to be an attempt to work towards a single correct description of reality although in realistic cases it is very difficult to assume that the model structure is correct or that a single parameter set is dominant in fitting the available observations. Arguments to give more attention to the potential of multiple acceptable models are given by Beven (2006a). Applications of BMA in hydrology are rather limited (e.g. Rojas et al., 2008, Ye et al., 2004, Renard et al., 2010).

3.3.2. Approximation methods

The First-Order Second Moment (FOSM) analysis, also known as the Taylor series method, is an approximation method that replaces the complex model f (Eq. (2.1)) with a truncated Taylor series. The first and second moment (mean and variance) of the input variables are then propagated through the surrogate model instead of the original model. FOSM analysis usually consists of following steps:

1. Assign a probability distribution and baseline value to each input variable X_i (with $i \in \{1, 2, ..., n\}$). The baseline values are represented by $\mathbf{x^0} = (x_1^0, x_2^0, ..., x_n^0)$ and form the point around which the Taylor series will be developed.

2. Develop a (first-order) Taylor series approximation of the model f, such that

$$y(\mathbf{x}) = y(\mathbf{x^0}) + \sum_{i=1}^{n} \left(\frac{\partial y(\mathbf{x^0})}{\partial x_i} \right) (x_i - x_i^0) \tag{3.17}$$

3. Estimate the expected value $E[Y]$ and variance $V[Y]$ as:

$$E[Y] = y(\mathbf{x^0}) \tag{3.18}$$

$$V[Y] = \sum_{i=1}^{n} \left(\frac{\partial y(\mathbf{x^0})}{\partial x_i} \right)^2 V[X_i] \tag{3.19}$$

   while assuming that the input variables X_i are uncorrelated and that $E[X_i]$ equals x_i^0.

The principle of FOSM analysis is illustrated in Figure 3.6. Although FOSM is a simple and widely used method for uncertainty propagation, it suffers from some typical limitations. The technique is only applicable to models that are differentiable and can therefore not be applied to models with a discrete or discontinuous behaviour. Moreover, the method only propagates the central moments of the input distributions such that information regarding the tail of the distribution is not considered. This can be a problematic limitation in risk assessment studies where the shape of the tail is critical to assess the risk of extreme events (Cullen and Frey, 1999).

Applications in hydrology

The Taylor series method is easy to implement and has low computational requirements. In hydrological studies FOSM has been used, for example to quantify the prediction uncertainty of river flow models arising from uncertainty in the model structure (Kasiviswanathan and Sudheer, 2013), to evaluate the risk of flooding when taking into account hydrologic, hydraulic, and geomorphologic uncertainty factors (Wu et al., 2011), to estimate uncertainties in the water balances due to uncertainties in field measurements (Schmadel et al., 2010), to assess the prediction

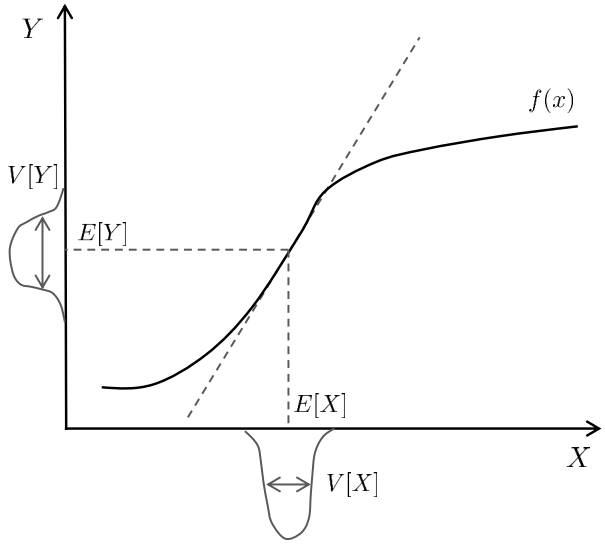


Figure 3.6: Illustration of the FOSM method for propagating uncertainty in a one-dimensional problem.

limits of groundwater flow estimates (Dettinger and Wilson, 1981) and to evaluate the uncertainty in solute transport arising from parameter uncertainty (Kuczera, 1988).

3.3.3. Extension principle

Suppose the model f (Eq. (2.1)) has two input variables X_1 and X_2, each described by their possibility distribution π_{X_1} and π_{X_2}, respectively. The possibility distribution of the model output variable Y is then obtained based on the extension principle of Zadeh (Zadeh, 1975a,b,c), as follows:

$$\pi_Y(v) = \sup_{f(u_1,u_2)=v} (\min(\pi_{X_1}(u_1), \pi_{X_2}(u_2))) \tag{3.20}$$

The above equation implies that the possibility that Y takes the value of v is determined by the best value (*i.e.* the supremum) of the joint possibility degree (*i.e.* the minimum) of all input couples (u_1, u_2) that are mapped to the value v using the model. The joint possibility degree is determined as the minimum of $\pi_{X_1}(u_1)$ and $\pi_{X_2}(u_2)$. In case of continuous functions and upper semi-continuous possibility distributions, application of the extension principle can be performed by the alpha-cut approach (Nguyen, 1978) and consists of following steps:

1. Construct the possibility distributions π_{X_1} and π_{X_2} for the respective input variables X_1 and X_2 (*cfr.* Section 3.2.3).

2. Define an α-cut.
3. Calculate $y = f(u_1, u_2)$ for all elements u_1 of X_1 and u_2 of X_2 that belong to the corresponding α-cuts $(\pi_{X_1})_\alpha$ and $(\pi_{X_2})_\alpha$.
4. Determine the corresponding α-cut for the model output Y based on the performed model evaluations:

$$(\pi_Y)_\alpha = \left[\min_{\substack{u_1 \in (\pi_{X_1})_\alpha \\ u_2 \in (\pi_{X_2})_\alpha}} (f(u_1, u_2)), \max_{\substack{u_1 \in (\pi_{X_1})_\alpha \\ u_2 \in (\pi_{X_2})_\alpha}} (f(u_1, u_2)) \right] \tag{3.21}$$

5. Repeat steps 2 to 4 for n_α (defined by the model user) different α-cuts.
6. Construct the possibility distribution π_Y based on the resulting α-cuts.

The extension principle is illustrated in Figure 3.7 for a two-dimensional problem.

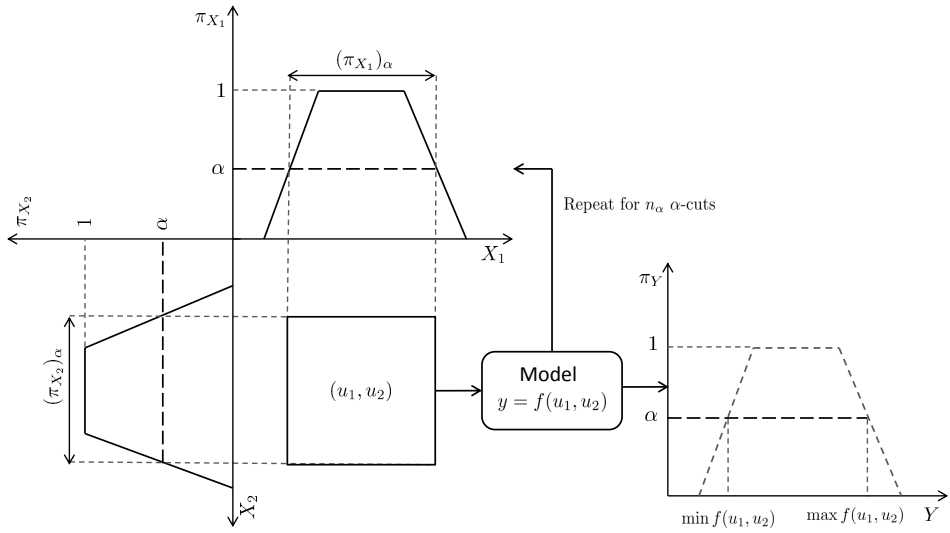


Figure 3.7: Illustration of the extension principle to propagate uncertainty in a two-dimensional problem.

Applications in hydrology

At present, applications of the extension principle in hydrological modelling studies are rather limited because it requires a high number of model applications to evaluate the different input combinations (u_1, u_2) for each α-cut. Examples include the assessment of the prediction uncertainty by a conceptual snowmelt runoff model arising from parameter uncertainties (Jacquin, 2010), and the evaluation of the uncertainty in soil moisture retrieved from radar backscattering as a consequence of uncertainty in the soil roughness characterization (Verhoest et al., 2007, Vernieuwe

et al., 2011). Other applications are the propagation of precipitation uncertainty (using a fuzzy set approach) through a deterministic rainfall-runoff model for flood forecasting (Maskey et al., 2004) and the uncertainty quantification of soil hydrological properties from uncertain information in a regional soil database (Martin-Clouaire et al., 2000). Recently, Scheerlinck et al. (2012, 2013) have developed a computationally efficient Fuzzy Calculator to propagate epistemic uncertainty regarding non-interactive as well as interactive fuzzy input variables. It is based on parallel computing of an optimization algorithm to determine the minimum and maximum of the function value for different α-cuts.

4

Sensitivity analysis

Prior to sensitivity analysis, the purpose of the analysis should be specified unambigiously in order to enlarge its effectiveness (Saltelli et al., 2004). It is crucial to first define the question that needs to be answered by the SA because the research question defines the setting (context) of the analysis and has an influence on the interpretation of the resulting sensitivity indices. The main settings are:

- *Factors fixing (FF) setting*: the objective of this setting is to identify the input variable(s) that can be fixed at any given value within its possible range without significantly reducing the output variance. This setting is sensitive to interactions between the input variables (Saltelli et al., 2004).
- *Factors prioritisation (FP) setting*: the objective here is to obtain a ranking of the input variables w.r.t. their importance, *i.e.* their ability to reduce the variance of the model output. Based on the ranking, efforts to reduce the model output uncertainty are guided by investigating resources in the factor that has the largest sensitivity. This approach is blind at model interaction as it is assumed that factors are fixed one-at-a-time (Saltelli et al., 2004).
- *Factors mapping (FM) setting*: the purpose of this setting is to analyse the input variables w.r.t. their capacity to produce realizations of the output variable within a given region (specified by bounds or thresholds).

The setting of the SA allows for the selection of an appopriate strategy. In addition, a number of problem constraints, e.g. computational cost, nonlinearity, correlated inputs, influence the choice of the method. Numerous SA techniques have been developed and are usually classified into screening, global and local methods.

- *Screening methods*: methods that aim at identifying the model inputs to which the model output is most sensitive. These methods are designed for application to models with a large number of input factors. They only provide a qualitative measure for importance and rank the input factors in order of importance. Screening methods are often relatively simple and are a particular instance of sampling-based methods. One of the most commonly used screening methods is the elementary effect method (Campolongo et al., 2007).
- *Global methods*: methods that calculate the total effect of a model input

on the model output across the entire input space. Commonly used global methods are the Sobol method (e.g. Sobol, 1993, Saltelli et al., 2008a), the Fourier amplitude sensitivity test (FAST) (e.g. Saltelli et al., 1999, McRae et al., 1982), the response surface method (RSM) (e.g. Cryer and Havens, 1999, Kleijnen et al., 1992) and Monte-Carlo-based methods (Hofer, 1999, Gwo et al., 1996). Most of them are variance-based, which means that the resulting sensitivity reflects the contribution of the model input to the total variance in the model output.

- *Local methods*: methods that investigate the sensitivity of the model output for a specific input scenario, *i.e.* at a fixed set of points from the model input domain. They are especially important for complex, nonlinear models as the effect of a model input on the model output may be highly localized, which makes the assessment of a global effect inappropriate in this case. Local methods are often based on first-order second-moment approximations (FOSM) in which it is assumed that the first two moments are sufficient to characterize a variable (Dettinger and Wilson, 1981). Examples of local methods are the Green's function method (e.g. Hwang et al., 1978) and the finite difference method (e.g. Lenhart et al., 2002, Foglia et al., 2009).

The sensitivity is generally expressed in terms of a quantitative sensitivity measure. The procedure to calculate the sensitivity measure is either derivative-based, variance-based, regression-based or sampling-based. Interpretation of the resulting sensitivity measure is determined by the setting of the SA such that 'sensitivity' (not per se a mathematical concept) should be defined at the stage of framing the analysis. Variance-based and regression-based methods lend itself to an interpretation of the sensitivity in terms of the FP setting. The FF setting can be studied by a regression- or variance-based method. FP and FF settings are linked to calibration and can be used to identify the parameters that can be effectively calibrated and which cannot. The FM setting can only be dealt with sampling-based methods. In this chapter, the basic principles of derivative-based, variance-based and regression-based methods to calculate sensitivity measures will be discussed. For the basic principles of sampling-based methods, we refer to Section 3.3.1. For a thorough review on methods for sensitivity analysis, the reader is referred to Saltelli et al. (2006), Frey and Patil (2002) and Helton and Davis (2003).

4.1. Derivative-based methods

Differential sensitivity analysis is a method for local SA that is based on partial differentiation of the model. The aim is to measure the effect of perturbing an input variable X_i (with $i \in \{1, 2, ..., n\}$), *i.e.* inducing small relative changes around its baseline value x_i^0, on the model output Y. This effect can be expressed by approximating the model $y = f(\mathbf{x})$ by a first-order Taylor series around the baseline

(or nominal) values $\mathbf{x^0} = (x_1^0, x_2^0, ..., x_n^0)$ (*cfr.* Eq.(3.17)). The partial derivatives $\partial y/\partial x_i$ in the Taylor series, further denoted as S_i, describe the sensitivity of Y with respect to X_i and are called first-order local sensitivity functions. Various methods to solve the partial derivatives have been developed: (i) the finite difference method, (ii) the direct differential method, (iii) the Green's function method, (iv) the polynomial approximation and (v) the automatic differentiation. In the finite difference method, the partial derivative is evaluated at one particular value of X_i by using the finite difference approximation (forward (+), backward (-) or central difference (c)):

$$\begin{cases} \left(\dfrac{\partial y}{\partial x_i}\right)_+ \approx \dfrac{y(x_i + \Delta x_i) - y(x_i)}{\Delta x_i} \\ \left(\dfrac{\partial y}{\partial x_i}\right)_- \approx \dfrac{y(x_i) - y(x_i - \Delta x_i)}{\Delta x_i} \\ \left(\dfrac{\partial y}{\partial x_i}\right)_c \approx \dfrac{y(x_i + \Delta x_i) - y(x_i - \Delta x_i)}{2\Delta x_i} \end{cases} \tag{4.1}$$

where Δx_i is the change in value of the input variable X_i, practically implemented as the particular value x_i of variable X_i multiplied by a user-defined perturbation factor ξ (for a discussion on the role of the perturbation factor in the SA, we refer to Section 4.1.1). Note that only one input variable is perturbed at a time, while all other variables are kept at their nominal value. This approach is called one-at-a-time SA (discussed in Section 4.1.2). Above sensitivity functions reflect the absolute sensitivity and are useful for calculating errors due to variations in input variables and for identifying the time steps on which the input variable has the largest effect in case Y is time-dependent. Alternatively, relative sensitivity functions can be used to compare the effect of different input variables (Saltelli et al., 2000):

$$S_i = \frac{x_i^0}{y^0}\frac{\partial y}{\partial x_i} \tag{4.2}$$

$$S_i = \frac{\sigma_{X_i}}{\sigma_Y}\frac{\partial y}{\partial x_i} \tag{4.3}$$

where y^0 is obtained as $f(\mathbf{x^0})$ and σ_Y and σ_{X_i} are the standard deviation on respectively the model output and the model input. The first relative sensitivity function (Eq. (4.2)) measures the effect on the output value y of perturbing the input value x_i by a fixed fraction of its nominal value, whereas the second function (Eq. (4.3)) measures the effect on y of perturbing x_i by a fixed fraction of its standard deviation. The sensitivity function (either absolute or relative) is then evaluated as follows: the model output Y is said to be sensitive to model input X_i if small changes in the input value x_i produce large changes in the output value y (high value of S_i). On the contrary, Y is called insensitive to X_i if small changes

in the input value x_i have almost no effect on the output value y (low value of S_i).

In case the model output is a function of time, the sensitivity function is a time series of S_i, such that $S_i(t) = \partial y(t)/\partial x_i$ with $t \in \{1, 2, ..., N\}$ in which N is the number of time steps in the simulation. Usually, $S_i(t)$ is then summarized into a sensitivity index, e.g. the composite scaled sensitivity index of Hill and Tiedeman (2007):

$$CSS_i = \sqrt{\sum_{t=1}^{N} \frac{S_i(t)^2}{N}} \tag{4.4}$$

with N the total number of time steps in the simulation and $S_i(t)$ a scaled sensitivity function, either scaled by the input variable or by the input standard deviation (Saltelli et al., 2008b, Hill and Tiedeman, 2007).

Despite its problems and recommendations against its use (Turanyi, 1990), differential sensitivity analysis is still frequently used because it is easy to implement and, unlike more sophisticated methods, the finite difference method does not require modification to the model code (De Pauw and Vanrolleghem, 2006). However, the method is computationally intensive in case of a large number n of input variables (De Pauw and Vanrolleghem, 2006). The required number of model runs is $n + 1$ when using the forward or backward difference approximation and raises to $2n$ when using the central difference approximation.

Applications in hydrology
Differential sensitivity analysis was used in hydrological modelling e.g. to assess the relative contributions of parameter uncertainties to the uncertainties in the unsaturated flow (Pan et al., 2011), to identify the sensitive parameters in the physically-based hydrological model SWAT (Lenhart et al., 2002), and to quantify the local sensitivity of the erosion response to the input parameters for the Rangeland Hydrology and Erosion Model (RHEM) (Wei et al., 2007).

4.1.1. Selecting the perturbation factor

The choice of the perturbation factor ξ to induce systematical changes in the input values determines the quality of the sensitivity function: the smaller its value, the better the finite difference scheme approximates the derivative. However, if ξ is taken too small it might give rise to numerical errors and if ξ is taken too large, errors due to model nonlinearities might be introduced in the analysis. De Pauw and Vanrolleghem (2006) proposed a methodology to optimize the perturbation factor by evaluating the quality of the sensitivity function. The basic idea is to make both types of error as small as possible by minimizing the difference in model sensitivity when inducing opposite changes in the model input, *i.e.* the difference between the sensitivity functions $\partial y/\partial x_{i+}$ and $\partial y/\partial x_{i-}$ should be as small as

possible. Several criteria can be used to quantify this difference such as the sum of squared errors (SSE), the sum of absolute errors (SAE), the maximum relative error (MRE) and the ratio (R) between the sensitivity functions. Although the MRE and R were found to be useful to assess the quality of the sensitivity function (De Pauw and Vanrolleghem, 2006), they should be interpreted with care as the criterion returns 0 (optimal value) in case $\partial y/\partial x_{i+}$ becomes 0. Generally, the SSE is used as optimization criterion:

$$SSE_i = \frac{1}{N}\sum_{t=1}^{N}\left(\left(\frac{\partial y(t)}{\partial x_i}\right)_+ - \left(\frac{\partial y(t)}{\partial x_i}\right)_-\right)^2 \tag{4.5}$$

with $y(t)$ the value of the model output at time step t and N the total number of time steps in the simulation. In literature, applying a fixed perturbation factor ξ to all input variables is frequently done, however, it is not advisable because the optimal perturbation factor depends on the input variable under consideration (De Pauw and Vanrolleghem, 2006). Commonly used values for ξ range between 10^{-2} and 10^{-4}.

4.1.2. One-at-a-time paradox

In case an SA on multiple model inputs is intended, the inputs can be varied simultaneously based on their underlying probability distribution (e.g. Gwo et al., 1996), or they can be varied individually around a baseline value while keeping the value of the other model inputs constant (e.g. Ferreira et al., 1995). The latter strategy is known as one-at-a-time sensitivity analysis (OAT-SA).

OAT analysis has been the subject of discussion because it is built on assumptions of model linearity and cannot detect interactions between model inputs (Saltelli and Annoni, 2010). This implies that OAT does not indicate whether the effect of simultaneously perturbing two input variables is the same as the superposition of the individual effects obtained by perturbing the first variable, going back to the baseline value and then perturbing the second variable. At the baseline values, the input factors X_i (with $i \in \{1, 2, ..., n\}$) are at their reference, *i.e.* the best estimated value. The purpose of the SA is to explore the entire space of the input variables with uniformly distributed points. This sample space is given by an n-dimensional hypercube. However, as OAT explores the sample space starting from the baseline values in one direction at a time, it does not allow to leave outside the n-dimensional hypersphere that is included in the n-dimensional hypercube (Saltelli and Annoni, 2010). This is illustrated for a 2-dimensional problem in Figure 4.1, where the grey-shaded areas cannot be explored by OAT. When increasing the dimensionality of the SA problem, the volume of the explored hyperspace reduces to a fraction of the total sample space (Figure 4.2). This illustrates the so-called curse of dimensionality.

Despite the shortcomings of OAT-SA, a literature review by Saltelli et al. (2006) revealed that most published sensitivity analyses use OAT. In some cases (strong) input correlations were observed (e.g. Boateng and Cawlfield, 1999, Zhu et al., 2010) and the assumption of independent inputs was therefore incorrectly adopted. Only in a limited number of SA studies, correlation structures have been incorporated (e.g. Pan et al., 2011, Jacques et al., 2006, Gevrey et al., 2006). The reason why OAT is so popular is that the observed effect on the model output is solely due to the fact that one input has been changed, which is consistent with the modeller's way of thinking to systematically evaluate the effect of input variation. Possible alternatives to OAT have been proposed by Saltelli and Annoni (2010).

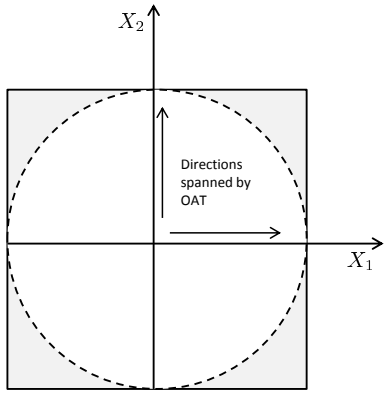


Figure 4.1: In two dimensions, the OAT partially explores a circle by moving along the abscis or the ordinate, while the entire sample space is given by the full square. Moving from the origin along the abscis, returning to the origin and moving along the ordinate does not allow to move outside the circle.

4.2. Regression-based methods

Regression-based methods are global methods for SA and are often used to replace a highly complex model $y = f(\mathbf{x})$ (Eq. (2.1)) with a simplified response surface (Iman and Helton, 1988), usually a simple regression equation. Regression analysis thus requires the identification of a relationship between the input variables X_i (with $i \in \{1, 2, ..., n\}$) and the output variable Y, with a general form:

$$\hat{y} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \epsilon \tag{4.6}$$

where $\hat{y}$ is the value of the model output as obtained with the simplified model (and is hence an approximation of the model output y obtained with the complex model), x_i is the value of the i-th model input X_i, β_i is the regression coefficient

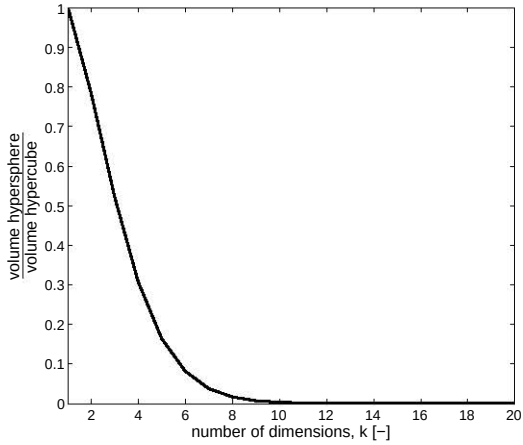


Figure 4.2: Illustration of the curse of dimensionality. The percentage of the explored space by OAT, relative to the entire sample space goes rapidly to zero as the number of dimensions increases.

for the i-th model input and ϵ is an error term. For a linear model, the regression coefficient β_i can be interpreted as the change in output value y when perturbing the input value x_i while keeping the value of the other inputs fixed. As such, it can be used to measure the sensitivity of Y w.r.t. X_i (Peck and Devore, 1996). In case the inputs are independent, β_i gives the fractional contribution of each input to the variance of Y such that $\sum_{i=1}^{n} \beta_i = V[Y]$. The regression coefficients are usually identified using standard least squares analysis: through a Monte Carlo simulation that covers the entire input space (*cfr.* Section 3.3.1), a set of model evaluations y_r (with $r \in \{1, 2, ..., n_{\mathrm{MC}}\}$) corresponding to n_{MC} sampled values of $\mathbf{x}$ are generated and a regression model is built.

The deviation of the prediction by the regression model from the actual model output can be measured using the model coefficient of determination R^2:

$$R^2 = \frac{\sum_{r=1}^{n_{\mathrm{MC}}} \left(\hat{y}_r - \overline{y}\right)^2}{\sum_{r=1}^{n_{\mathrm{MC}}} \left(y_r - \overline{y}\right)^2} \tag{4.7}$$

with $\hat{y_r}$ the prediction of the simple regression model using input $\mathbf{x_r}$ and $\overline{y}$ the average of the n_{MC} model evaluations obtained with the complex actual model using Eq. (2.1). If R^2 is close to 1, the regression model accounts for most of the variability in Y, while low values of R^2 indicate a nonlinear behaviour of the model. In the latter case, nonlinear regression approaches should be used, for example Neural Networks (Zeng and Yeung, 2003) or Random Forests (Pappenberger et al., 2006b).

Applications in hydrology

Examples of regression-based SA studies include for example the analysis of the

sensitivity of a water quality model to model input parameters (Ma et al., 2000) and the evaluation of parameter sensitivities in the calibration of the distributed SWAT model (Muleta and Nicklow, 2005).

4.3. Variance-based methods

Variance-based methods are global SA methods that decompose the variance of the model output into a sequence of variances. The total variance $V[Y]$ on the model output Y is conceived as a sum of terms with increasing dimensionality (Sobol, 2001):

$$V[Y] = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + ... + V_{12...n} \tag{4.8}$$

where $V_i = V[E[Y|X_i = x_i^0]]$, $V_{ij} = V[E[Y|X_i = x_i^0, X_j = x_j^0]] - V[E[Y|X_i = x_i^0]] - V[E[Y|X_j = x_j^0)]]$, etc. The total effect of an input X_i measures the residual variance of the model output that remains by removing the portion explained by all other inputs (Saltelli et al., 2000). It is defined as the sum of the main effect of X_i and all interaction effects of X_i with any other input variable. The main effects are given by the terms V_i and the interaction effects are given by the remaining terms V_{ij}, V_{ijm} and $V_{12...n}$. Quantifying the total effect allows to identify the factors that can be fixed (FF setting) because if the total effect of a factor is zero, it can be considered as irrelevant. The main effect V_i of an input X_i measures the portion of the variance of the model output that is explained by that input, by allowing all other inputs to be varied at the same time. As such, the main effect can be used to rank the factors w.r.t. their importance (FP setting).

Sensitivity indices are obtained by normalizing the main and interaction effects by the unconditional variance of Y, and are scaled between 0 and 1:

$$S_i = \frac{V_i}{V[Y]} \tag{4.9}$$

$$S_{1...n} = \frac{V_{1..n}}{V[Y]} \tag{4.10}$$

The total sensitivity indices, accounting for the interaction between inputs, are then given by (Sobol, 2001):

$$S_i^T = S_i + \sum_j S_{ij} + ... + S_{1...n} \tag{4.11}$$

The first-order sensitivity indices S_i can be conceived as the expected fractional reduction in output variance that would be achieved if the value of X_i is known, whereas S_i^T can be conceived as the expected fraction of the output variance that would be left if only the value of X_i is undetermined (Saltelli et al., 2000). For

additive models, $S_i^T = S_i$ and $\sum_{i=1}^{n} S_i = 1$ because no interaction effects exist. Several methods to estimate the variances are available, among which commonly used are the Sobol method and the extended FAST method. In the Sobol method, partial variances are estimated using multi-dimensional integrals computed via a Monte Carlo procedure (Sobol, 2001). It is an attractive approach as it is relatively easy to implement. The extended Fourier amplitude sensitivity test (eFAST) estimates the set of main and interaction effects using Fourier coefficients (Saltelli et al., 1999) and is numerically more efficient than the Sobol method.

Variance-based mehods have some useful properties in contrast with the derivative-based (*cfr.* Section 4.1) and regression-based SA methods (*cfr.* Section 4.2) (Saltelli et al., 2000):

1. The sensitivity measures take into account the whole range of input variation and the form of the PDF. The effect of factor X_i is then averaged over the whole range of input values x_i.
2. The sensitivity measures take into account interaction effects when determing the effect of X_i on Y.
3. The sensitivity measures are model independent, *i.e.* their performance is not conditional on the additivity or linearity of the model.
4. They can cope with the curse of dimensionality (*cfr.* Section 4.1.2).

Applications in hydrology

Variance-based SA methods are widely used in hydrological studies, e.g. in flood inundation modelling to assess the sensitivities to soil hydraulic model parameters (Cloke et al., 2008) or to identify the most sensitive and important parameters in the TOPMODEL (Ratto et al., 2007). Tang et al. (2007) compared the results of Sobol's method with three other SA methods to characterize the sensitivities of a lumped watershed model and showed that the resulting sensitivities are highly impacted by the choice of the SA method. Massmann and Holzmann (2012) evaluated the effect of the model parameters on the discharge of a conceptual rainfall-runoff model using Sobol's method and compared it with other SA methods

Part II

Distributed modelling of catchment hydrology

5

Introduction

In this part, it is described how the hydrological processes are being modelled to be used in the uncertainty and sensitivity analyses in Parts II and III. An inspection of the underlying processes is useful to reveal the conditions under which the analyses are performed. Chapter 6 gives a detailed description of the hydrological model structure, the model inputs and the model operation. The catchment for which the modelling is performed is described in Chapter 7, together with the available data for the catchment. Finally, Chapter 8 discusses the results of the hydrological model calibration.

An overview of the different types of hydrological models and a description of the general operation of land surface models, is given below.

5.1. Classification of hydrological models

Many hydrological models have been developed and refined during the last four decades. Although each of these models involves the interaction between soil, climate and land use, the behaviour of the model's temporal and spatial features is variable. A high variety in types of hydrological models therefore exist, but a classification of the models is not exact as different hydrologists may give different classifications (e.g. Shaw, 1983, Chow et al., 1988, Wheater et al., 1993, Singh and Frevert, 2006). In this dissertation, it is chosen to use the criteria-based classification of hydrological models according to Jajarmizadeh et al. (2012) (and references therein). A schematic of the classification is shown in Figure 5.1. Based on the five criteria (laws and assumption, equations, parameters, time and procedure of computation), the following types of models can be distinguished:

Empirical and physically-based models
Empirical models (black box) are models of which the model structure to characterize the system response, is built on observations of the model output variables. These models are a representation of a real system based on a mathematical description without using general physical laws. On the contrary, physically-based models (white box) represent different hydrological processes by considering the physical meaning of a hydrological system (usually described as partial differential

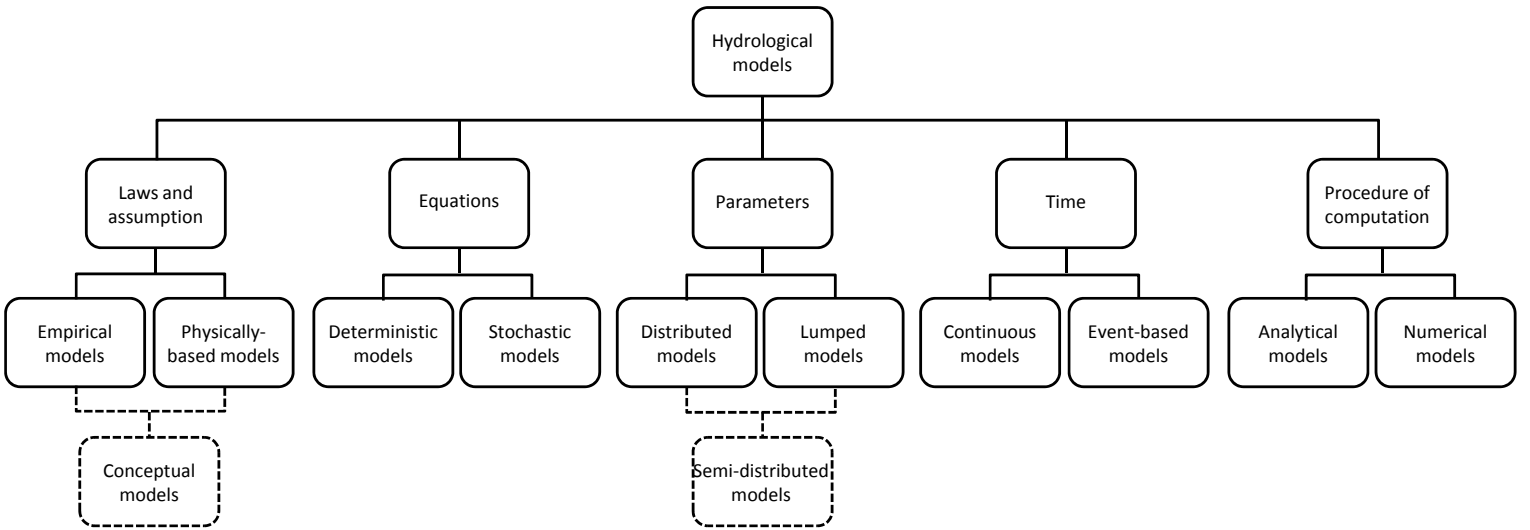


Figure 5.1: Criteria-based classification scheme for hydrological models (modified from Jajarmizadeh et al. (2012)). The five criteria used are (i) laws and assumption, (ii) equations, (iii) parameters, (iv) time and (v) procedure of computation.

equations). In between the empirical and physically-based models lie the conceptual models (grey box). These models represent the physical processes that are perceived to be of importance in the hydrological system and are generally based on schematic storages that are combined in a conceptual view of the most important processes.

Deterministic and stochastic models
Deterministic models always produce the same output for a given set of input and parameter values due to their disability to include randomness. Stochastic (or probabilistic) models, on the contrary, produce outputs that have a partial randomness and use probability equations to describe (partially or totally) the model inputs, parameters and/or processes. Therefore, stochastic models allow for assessing uncertainty in the model output due to uncertainty in the input variables, parameters or processes. This implies that with a given set of input and parameter values different outputs are generated.

Distributed and lumped models
Distributed models produce model outputs that are distributed in space, by discretizing the catchment into a number of grid elements (usually squares) and by calculating the state variables associated with each grid element. Figure 5.2 illustrates the application of a distributed model. As distributed models represent the main physical processes in detail, they offer a high degree of accuracy. However, parameter and input values are often averaged over many grid elements because of limited data availability. Moreover, distributed models deal with other problems such as nonlinearity, scale, uniqueness and uncertainty (Beven, 2001). Lumped models generate model outputs that are averaged over the catchment area by representing the complete hydrological system as a single homogeneous unit. Consequently, these models do not take into account the spatial variability of processes, inputs and catchment characteristics. This means that the number of parameters in a lumped model is often two or three orders of magnitude lower than in a distributed model of the same catchment, which generates different requirements w.r.t. parametrization, calibration and validation. In between distributed and lumped models lie the semi-distributed (or semi-lumped) models. These models simulate the average behaviour of small homogeneous units (sub-basins or representative hydrological units) in the catchment by discretizing the catchment to a feasible degree and by running a set of lumped models. As such, they partly allow for spatial distribution while at the same time keeping data requirements and computational costs low.

Continuous and event-based models
Continuous models take into account a time series of rainfall events whereas event-based models simulate a specific event for a short period of time. The time scale of a continuous simulation is usually chosen by the model user and is defined by the time intervals used for input and internal computations, or for output and

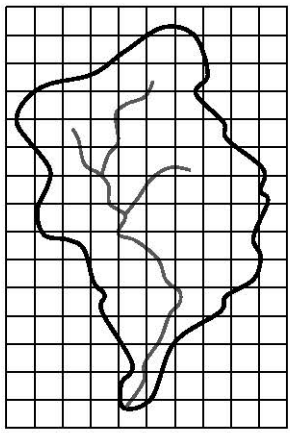


Figure 5.2: Illustration of spatial discretization of a catchment with regard to the application of a distributed model.

calibration of the model (Singh, 1995).

Analytical and numerical models
In analytical models, calculation of the model output is performed according to the analytical (closed-form) solution of the governing equations whereas numerical models use a numeric time-stepping procedure to obtain the solution of the governing equations. The latter implies that model outputs are only obtained at a finite number of points in both space and time.

5.2. Operation of Land Surface Models

The hydrological processes are only a part of the global processes controlling the earth's system, and interact with the processes in other system components (e.g. cloud formation in the atmosphere). If the hydrological model explicitly describes the interaction between land surface, vegetation and atmosphere, it is called a Soil-Vegetation-Atmosphere Transfer Scheme (SVATS) or a Land Surface Model (LSM). These models update the state of the land surface, *i.e.* prognostic variables that define the dynamics of a system (e.g. temperature and water content of the soil), through the influence of meteorological variables (e.g. precipitation, wind speed and longwave radiation) on the hydrological processes (Figure 5.3). The updated states form the initial conditions for the next time step, after which the SVATS is repeatedly applied.

The fundamental equations in a SVATS are the conservation of energy and the

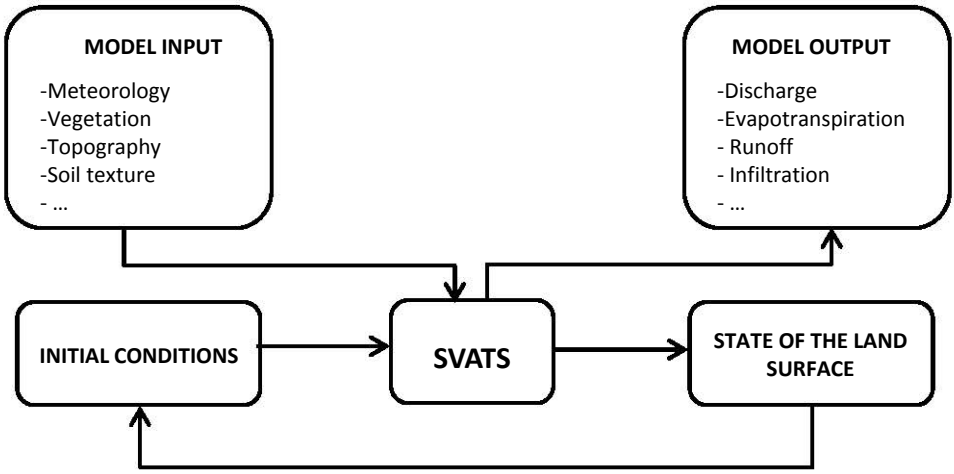


Figure 5.3: Schematic of the operation of a land surface model.

conservation of mass, respectively written as:

$$R_n = \rho_w \lambda_w E + H_s + G \tag{5.1}$$

$$P = E + Q + \Delta S \tag{5.2}$$

in which R_n is the net radiation [W.m^{-2}], ρ_w the density of water [kg.m^{-3}], λ_w the latent heat of vaporization of water [J.kg^{-1}], E the evapotranspiration rate [m.s^{-1}], H_s the sensible heat flux [W.m^{-2}], G the soil heat flux [W.m^{-2}], P the precipitation [m.s^{-1}], Q the discharge [m.s^{-1}] and ΔS the change in water storage per unit of surface area [m.s^{-1}].

The energy balance requires that the net radiation is partitioned into a latent heat flux ($\rho_w \lambda E$), a sensible heat flux and a soil heat flux, while at the same time the water balance equation requires that the precipitation is partitioned into evapotranspiration, discharge and change in water storage. Both balances are linked through the evapotranspiration term and solving the balances is the overall objective of the SVATS.

The fundamental processes in a SVATS are modelled as follows: the land surface receives an amount of shortwave radiation from the sun, which is partly reflected, and an amount of longwave radiation from the atmosphere. The incoming radiation induces an increase in the land surface temperature as a consequence of which the land surface emits an amount of longwave radiation. The sum of the incoming shortwave and longwave radiation minus the emitted longwave and the reflected shortwave radiation is the net radiation (R_n). The net radiation is used to evaporate water in the soil (latent heat flux), to warm up the soil (soil heat flux) and to warm up the air above the vegetation (sensible heat flux). At the same time, an amount of precipitation is received by the land surface and is partly stored in the

vegetation (interception). The precipitation that reaches the soil (net precipitation) partly infiltrates and further drains towards the water table (percolation) or gets evaporated or is used by plants for transpiration (where the latter two correspond to the latent heat flux from the energy balance). The percolated water can leave the catchment as baseflow while the part of the precipitation that does not infiltrate into the soil, leaves the catchment as surface runoff.

6

Hydrological model description

The hydrological model used in this study is TOPLATS, or the TOPMODEL-based Land Surface-Atmosphere Transfer Scheme that couples a simple local SVATS to a TOPMODEL-based scheme in order to aggregate spatially variable land surface fluxes to the regional scale. It is a physically-based, spatially distributed water and energy balance model based on a lateral redistribution of water (Famiglietti and Wood, 1994, Peters-Lidard et al., 1997, Pauwels and Wood, 1999), *i.e.* groundwater gradients induce spatial patterns of soil moisture and are estimated from the local topography and the soil transmissivity (Sivapalan et al., 1987). The model was originally developed to simulate water and energy fluxes for warm seasons (Famiglietti and Wood, 1994), but was later modified to expand the representation of the hydrological processes towards conditions in high latitudes (e.g. frozen ground and open water bodies) (Pauwels and Wood, 1999). Furthermore, the model was modified in 1997 to correct for deficiencies in the representation of the energy fluxes (e.g. soil evaporation and ground heat flux) (Peters-Lidard et al., 1997) and in 2001 to improve the procedure for solving the water balance (Crow et al., 2001).

For each computational unit (either a pixel or an interval), a separate local water and energy balance is solved at each time step in order to obtain a spatial distribution of soil moisture, water table depth, temperatures and fluxes. The basic concept in the model is that shallow groundwater gradients set up spatial patterns of soil moisture that influence infiltration and runoff during storm events and evaporation and drainage between storm events. The soil moisture content (assumed uniformly with depth) is initialized based on the local water table depth and the assumption of an equilibrium moisture profile after which the soil moisture content is updated using the local soil water balance equations as described in Peters-Lidard et al. (1997).

6.1. Model structure

In TOPLATS, the land surface is represented for each computational unit as follows: the soil column is partitioned into an upper surface zone (referenced as u) and a lower transmission zone (referenced as l). In both layers, roots (referenced as rt) can exist which means that water for vegetation transpiration can be supplied by

either layer depending on the distribution of the roots in the soil column (which is specified for each land cover class as a model parameter). Water for bare soil evaporation is only supplied by the surface zone. Beneath the transmission zone, the soil is saturated. Further, the land surface is partitioned into bare soil (referenced as bs) and vegetated components (referenced as vg). The vegetation is modelled as a single layer or as a two-layer model with an overstory and understory component following the big-leaf approach. A moss layer component can be added if necessary.

6.1.1. Local water balance equations

The general form of the water balance equation is given by Eq. (5.2). In the hydrological model, separate water balances are formulated for the different water reservoirs: the surface zone, the transmission zone, the saturated zone and the canopy. The vertical soil moisture fluxes are represented by the Richards equation (Richards, 1931) for unsaturated soil water flow and the water balances are solved numerically (Crow et al., 2001) for each computational unit at each time step in the simulation period. A schematic of the water budgets in the local SVATS within the hydrological model is given in Figure 6.1.

Surface zone water balance
The surface zone extends from the soil surface to the top of the transmission zone and functions as a reservoir for supplying water for vegetation transpiration and bare soil evaporation. The water balance for the surface zone is:

$$z_u \frac{d\theta_u}{dt} = f_{bs} I_{bs} + f_{vg} f_{rt,u} I_{vg,u} + f_{vg} f_{rt,l} I_{vg,l} - d_u - f_{bs} E_{bs} - f_{vg} f_{rt,u} E_{vg,u} - g_u \quad (6.1)$$

where θ_u is the soil moisture content of the surface zone [$m^3.m^{-3}$], t is the time [s], f_{rt} is the root fraction in a given zone (u or l) [-], f_{bs} is the fraction of bare soil surface [-], f_{vg} is the fraction of vegetated land surface [-] (and equals $1 - f_{bs}$), I is the infiltration rate [$m.s^{-1}$], E is the evaporation or transpiration rate [$m.s^{-1}$], z_u is the depth of the upper zone [m], d_u is the diffusive flux from the upper zone to the lower zone [$m.s^{-1}$] and g_u is the gravitational drainage from the lower zone towards the transmission zone [$m.s^{-1}$]. The diffusive flux is calculated using the approach as described by Mahrt and Pan (1984):

$$d_u = D \frac{\theta_u - (\theta_u + \theta_l)/2}{(z_u + z_l)/2} \quad (6.2)$$

where θ_l is the soil moisture content of the transmission zone [$m^3.m^{-3}$], z_l is the depth of the lower zone [m] and D is the diffusivity [$m^2.s^{-1}$], determined as a function of the Brooks and Corey (1964) parameters. Drainage from the surface zone occurs at gravity-driven rates and is calculated using the $K_s(\theta)$-relation of

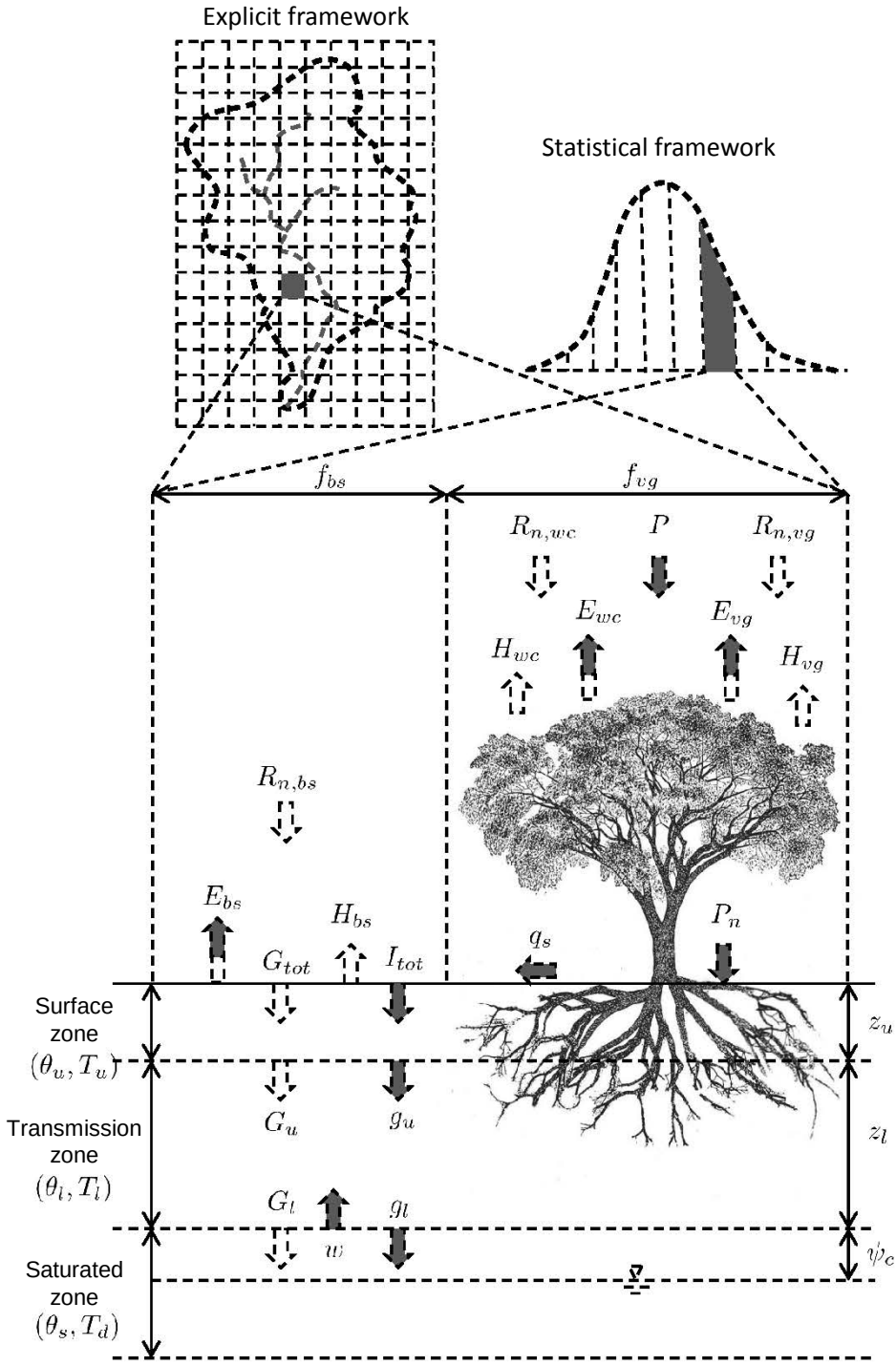


Figure 6.1: Schematic representation of the water budgets (indicated with filled arrows) and the energy budgets (indicated with open arrows) in the local SVATS (*i.e.* for one computational unit, either a pixel or an interval) within the hydrological model TOPLATS.

Brooks and Corey (1964) (*cfr.* Section 6.1.4):

$$g_u = K_s \left(\frac{\theta_u - \theta_r}{\theta_s - \theta_r} \right)^{\frac{2+3\lambda}{\lambda}} . \tag{6.3}$$

where K_s is the saturated hydraulic conductivity [m.s^{-1}], θ_s is the saturated soil moisture content [m^3.m^{-3}], θ_r is the residual soil moisture content [m^3.m^{-3}] and λ is the pore size distribution index [-]. The actual infiltration into the bare soil (I_{bs}[m.s^{-1}]) and into the vegetated soil (I_{vg}[m.s^{-1}]) is determined as the minimum of the infiltration capacity I_{cap}[m.s^{-1}] and respectively the total precipitation rate P[m.s^{-1}] and the net precipitation rate P_n[m.s^{-1}]:

$$I_{bs} = \min [I_{cap}, P] \tag{6.4}$$
$$I_{vg} = \min [I_{cap}, P_n] \tag{6.5}$$

The infiltration capacity is determined following Milly (1986) and is a function of the soil properties, the soil moisture content at the beginning of the storm event and the cumulative infiltration. The actual rate of bare soil evaporation (E_{bs}[m.s^{-1}]) and the actual rate of canopy transpiration (E_{vg}[m.s^{-1}]) are given by:

$$E_{bs} = \min [E_{bs,cap}, E_{bs,pot}] \tag{6.6}$$
$$E_{vg} = \min [E_{vg,cap}, E_{vg,pot}] \tag{6.7}$$

where $E_{bs,cap}$ is the soil-controlled evaporation rate [m.s^{-1}], calculated using a soil resistance formulation (Passerat de Silans et al., 1989), $E_{vg,cap}$ is the vegetation-controlled transpiration rate [m.s^{-1}], calculated using a linear interpolation for the canopy resistance (Jacquemin and Noilhan, 1990). $E_{bs,pot}$ and $E_{vg,pot}$ are the potential (atmospherically-controlled) rates of respectively bare soil evaporation and vegetation transpiration. The potential rates are obtained by solving the energy balance equations for both the dry canopy and the bare soil (*cfr.* Section 6.1.2).

The surface zone soil moisture content (θ_u[m^3.m^{-3}]) is compared with the saturated soil moisture content (θ_s[m^3.m^{-3}]) to determine whether runoff occurs as saturation excess runoff (if $\theta_u = \theta_s$) or as infiltration excess runoff (if $\theta_u < \theta_s$). Runoff fluxes for the bare soil are then obtained as:

$$\begin{cases} q_{bs} = P & \text{if } \theta_u = \theta_s \\ q_{bs} = P - I_{cap} & \text{if } \theta_u < \theta_s \text{ and } P > I_{cap} \end{cases} \tag{6.8}$$

The vegetated runoff fluxes q_{vg}[m.s^{-1}] are calculated by replacing P by P_n in the above equations.

Transmission zone water balance

The transmission zone extends from the bottom of the surface zone to the top

of the capillary fringe. The depth of the transmission zone varies in space and time because the local water table depth is calculated at each time step for each computational unit. This soil layer conducts water from the surface zone to the water table. As roots of the vegetation may penetrate the transmission zone, this soil layer also supplies water for transpiration. The water balance for the transmission zone is:

$$z_l \frac{d\theta_l}{dt} = g_u - g_l + d_u - d_l - f_{rt,l} f_{vg} E_{vg,l} \tag{6.9}$$

where z_l is the depth of the lower zone [m], θ_l is the soil moisture content of the lower zone [$m^3.m^{-3}$], g_l is the gravitational drainage towards the ground water table [$m.s^{-1}$] and d_l is the diffusive flux from the lower zone to the saturated zone [$m.s^{-1}$]. The diffusive flux is calculated by replacing θ_u by θ_l and θ_l by θ_s in Eq. (6.2). The drainage from the transmission zone (g_l) is calculated by replacing θ_u by θ_l in Eq. (6.3). Infiltration, evaporation and transpiration rates are obtained in a similar way as described for the surface zone.

Saturated zone water balance

From the top of the capillary fringe and deeper, the soil is saturated. The water table depth z_i [m] at location i is calculated from the average water table depth $\overline{z}$ [m] based on the TOPMODEL concept (*cfr.* Section 6.1.3, Eq. (6.18)). The water balance expresses the rate of change in local water table depth, which is assumed to be equivalent to the rate of change in the catchment-averaged water table depth:

$$\frac{dz_i \theta_s}{dt} = w - g_l - \frac{Q_b}{A} \tag{6.10}$$

where w is the capillary rise [$m.s^{-1}$], g_l the drainage from the transmission zone [$m.s^{-1}$] (calculated from Eq. (6.3)) Q_b the baseflow [$m^3.s^{-1}$] and A the area of the catchment [m^2]. The capillary rise is calculated following the formulation of Gardner (1958):

$$w = \frac{Ca}{(z_i - \psi_c)^b} \tag{6.11}$$

where C, a, and b are calculated from the Brooks and Corey (1964) parameters using the expressions of Eagleson (1978), and ψ_c is the bubbling pressure [m]. The baseflow is calculated from the average water table depth using Eq. (6.19) from the TOPMODEL (*cfr.* Section 6.1.3).

Canopy water balance

The change in the amount of water stored on the leaves in the canopy ($\theta_c[m]$), is expressed by the canopy water balance:

$$\frac{d\theta_c}{dt} = P - E_{wc} - P_n, \; 0 \leq \theta_c \leq \theta_{c,cap} \tag{6.12}$$

where P is the precipitation rate [$m.s^{-1}$], E_{wc} is the evaporation rate from the

wet canopy [m.s^{-1}], P_n is the net precipitation [m.s^{-1}] and $\theta_{c,cap}$ is the storage capacity [m], calculated as a function of the leaf area index (Dickinson, 1984). The evaporation from the wet canopy is derived from the potential evaporation $E_{wc,pot}$[m.s^{-1}] as follows:

$$E_{wc} = \omega_{wc} E_{wc,pot} \tag{6.13}$$

with ω_{wc} the areal fraction of wet canopy, determined following Deardorff (1978).

6.1.2. Local energy balance equations

The potential evaporation rate from the entire wet canopy ($E_{wc,pot}$), the potential (atmospheri-cally-controlled or unstressed) transpiration rate ($E_{vg,pot}$) and the potential evaporation rate for bare soils ($E_{bs,pot}$) are determined by solving the energy balance equations iteratively for the skin (or surface) temperature, using a Newton-Raphson iteration. The general form of the energy balance is given by Eq. (5.1). A schematic of the energy budgets in the local SVATS within the hydrological model is given in Figure 6.1.

The net radiation R_n[W.m^{-2}] from the energy balance is calculated as:

$$R_n = R_s(1-\alpha) + \epsilon R_l - \epsilon \nu T_s^4 \tag{6.14}$$

where R_s is the incoming shortwave radiation [W.m^{-2}], α is the albedo [-], ϵ is the emissivity [-], R_l is the incoming longwave radiation [W.m^{-2}], ν is the Stefan-Boltzmann constant (5.6676 W.m^{-2}K^{-4}) and T_s is the skin temperature [K]. The latent heat flux [W.m^{-2}] is calculated as (Milly, 1991):

$$\rho_w \lambda_w E = \frac{\rho_a c_p}{\gamma(r_c + r_{av})} \left(e_s(T_s) - e_a\right) \tag{6.15}$$

with ρ_w de density of water [kg.m^{-3}], λ_w the latent heat of vaporization of water [J.kg^{-1}], E the evapotranspiration rate [m.s^{-1}], ρ_a the density of air [kg.m^{-3}], γ the psychrometric constant [Pa.K^{-1}], c_p the specific heat of air [J.kg^{-1}.K^{-1}], r_c the canopy resistance [s.m^{-1}], r_{av} the aerodynamic resistance for vapor transport [s.m^{-1}], $e_s(T_s)$ the saturated vapor pressure [Pa] at temperature $T_s[K]$, e_a the vapor pressure in the air [Pa]. The formulation of the aerodynamic resistance (Ek and Mahrt, 1991) includes a stability correction factor because the stability of the atmospheric surface layer can have a significant effect on the heat fluxes. The sensible heat flux [W.m^{-2}] is given by:

$$H_s = \frac{\rho_a c_p}{r_{ah}} \left(T_s - T_a\right) \tag{6.16}$$

where r_{ah} is the aerodynamic resistance for heat transport [s.m^{-1}], calculated following Ek and Mahrt (1991), and T_a is the air temperature [K]. Calculation of

the ground heat flux $G(z)$[W.m^{-2}] at a given depth z [m] is based on discretization of the one-dimensional heat diffusion equation:

$$G(z) = \zeta \frac{\partial T(z)}{\partial z} \tag{6.17}$$

where ζ is the soil thermal conductivity [W.m^{-1}.K^{-1}] and T is the soil temperature [K]. In case the surface is vegetated, the ground heat flux should be adapted by multiplying Eq. (6.17) by a transmission factor η [-] (which is a function of the leaf area index).

6.1.3. The TOPMODEL

Water table dynamics are an important factor in estimating the water and energy budgets because the water table depth has an influence on the soil moisture content and hence on the hydrological fluxes. In the hydrological model, saturated subsurface flow and water table depth are estimated through the topographic index model (TOPMODEL). The TOPMODEL is a conceptual ground water model, originally developed by Beven and Kirkby (1979), in which the shape of the ground water table is estimated on basis of the topography and the soil properties. The spatial variability in topography and soil properties are parameterized by a soil-topographic index. Locations with the same value of the soil-topographic index are hence assumed to respond in a hydrologically similar way under the same meteorological and soil surface conditions. High values of the index imply a tendancy towards saturation of the soil surface, whereas low values imply a deeper water table.

The relationship between the local water table depth z_i [m] and the soil-topographic index ($\ln(a\overline{T^m}/(T_i^m \tan\beta))$) is expressed as (Sivapalan et al., 1987):

$$z_i = \overline{z} - \frac{1}{f_K}\left(\ln\left(\frac{a\overline{T^m}}{T^m \tan\beta}\right) - \overline{\lambda_T}\right) \tag{6.18}$$

where $\overline{z}$ is the catchment-averaged water table depth [m], f_K is a parameter that describes the exponential decay of the saturated hydraulic conductivity with depth, further called the hydraulic conductivity decay coefficient [m^{-1}], a is the area draining into location i [m], $\overline{T^m}$ is the catchment-averaged transmissivity of the soil [m^2.s^{-1}], T^m is the transmissivity at location i [m^2.s^{-1}], β is the local slope angle [radians] and $\overline{\lambda_T}$ is the catchment-averaged topographic index (with the local topographic index λ_T equal to $\ln a/\tan\beta$). Based on the above expression, a formula to calculate the baseflow Q_b[m^3.s^{-1}] from the average water table depth $\overline{z}$ can be derived:

$$Q_b = Q_0 e^{-f_K \overline{z}} \tag{6.19}$$

where A is the area of the catchment [m^2] and Q_0 is the critical baseflow [m^3.s^{-1}],

defined as:

$$Q_0 = A\overline{T^m}e^{-\overline{\lambda_T}} \tag{6.20}$$

The above equation is built on the assumption that (i) the precipitation is uniformly distributed over the catchment and is instantly discharged as groundwater (quasi-steady conditions), (ii) the hydraulic conductivity decreases exponentially with depth following the relationship $K_s(z) = K_0 e^{-f_K z}$ (where K_0 is the hydraulic conductivity at the surface [m.s^{-1}]) and (iii) the groundwater flow is driven by Darcy's law.

6.1.4. Soil description

In the hydrological model, the soil properties are described by the Brooks and Corey (1964) model. This hydraulic model includes a soil moisture retention curve (SMRC) to describe the relationship between the matric head ψ [m] and the soil moisture content θ [m^3.m^{-3}]:

$$\begin{cases} \theta(\psi) = \theta_r + (\theta_s - \theta_r)\left(\frac{\psi_c}{\psi}\right)^{\lambda} & \text{if } \psi > \psi_c \\ \theta(\psi) = \theta_s & \text{if } \psi \leq \psi_c \end{cases} \tag{6.21}$$

and a hydraulic conductivity curve (HCC) to describe the relationship between the matric head ψ [m] and the hydraulic conductivity K[m.s^{-1}]:

$$\begin{cases} K(\psi) = K_s\left(\frac{\psi_c}{\psi}\right)^{(2+3\lambda)} & \text{if } \psi > \psi_c \\ K(\psi) = K_s & \text{if } \psi \leq \psi_c \end{cases} \tag{6.22}$$

The parameters used in the Brooks and Corey (1964) model are the saturated soil moisture θ_s [m^3.m^{-3}], the residual soil moisture θ_r [m^3.m^{-3}], the bubbling pressure ψ_c [m], the pore size distribution index λ [-] and the saturated hydraulic conductivity K_s [m.s^{-1}].

6.2. Model operation

The states that describe the land surface in the local SVATS are θ_u, θ_l, θ_c, T_s and z_i. These states are updated in the simulation process following the equations as described in Sections 6.1.1 and 6.1.2. The model roughly operates as follows:

1. The soil moisture conditions in the surface and transmission zone (assuming gravitational equilibrium in the local soil profile) are initialized for the first time step in a storm event. Initialisation is based on the local water table depth and on the Brooks and Corey (1964) model for soil water retention by

substituting ψ by $\overline{z}$ in Eq. (6.21).

2. The soil evaporation and vegetation transpiration at potential rates are determined by solving the energy balance (Eq. (5.1)). Under potential conditions, evapotranspiration is controlled by the meteorological variables and is not restricted by soil moisture availability. The potential evaporation for the entire wet canopy $E_{wc,pot}$ is determined by solving the energy balance for T_s, representing the temperature of the wet vegetated surface, using Eqs. (6.14)-(6.17) in which r_c is set equal to zero. The potential transpiration $E_{vg,pot}$ is determined in the same manner as $E_{wc,pot}$ but with r_c equal to $r_{st,min}/LAI$ (where $r_{st,min}$ is a minimum value of stomatal resistance and LAI is the leaf area index). The potential evaporation for the bare soil $E_{bs,pot}$ is determined by solving the energy balance for the temperature of the bare soil surface T_s, and by setting r_c equal to zero.

3. The canopy water balance (Eq. (6.12)) is solved for the amount of water stored on the leaves (θ_c). If the obtained amount of water is less than zero, the actual evaporation rate E_{wc} is decreased until θ_c equals 0. If the water storage becomes larger than the storage capacity, θ_c is put equal to $\theta_{c,cap}$, and the net precipitation $P_n \Delta t$ is put equal to $(\theta_{c,cap} - \theta_c)$.

4. Infiltration, drainage and diffusion are calculated using Eqs. (6.5), (6.3) and (6.2), respectively. As such, the remaining terms of the surface and transmission zone water balance are obtained. The baseflow and the capillary rise from the saturated water zone balance are calculated using Eqs. (6.19) and (6.11), respectively.

5. The water balance equations (Eqs. (6.1), (6.9) and (6.10)) are solved for θ_u, θ_l and z_i, respectively. If the resulting surface zone soil moisture (θ_u) and/or the transmission zone soil moisture (θ_l) is lower than the residual soil moisture content (θ_r), the actual transpiration rate is reduced to ensure that the soil moisture content is never below θ_r. If the obtained soil moisture is larger than the saturated soil moisture content, it is put equal to θ_s. Based on the soil moisture content of the surface zone, the surface runoff is calculated following Eq. (6.8).

6. Given the actual evaporation and transpiration rates, the energy balance is resolved at actual rates, and the energy fluxes are recalculated with the correct skin temperature. To determine the correct bare soil temperature and the correct dry canopy temperature, E_{bs} and E_{vg} are respectively substituted for E in Eq. (6.15).

7. With the correct energy and water budgets for both the bare soil and vegetated components, the local evapotranspiration rate ET [m.s^{-1}] and the local surface runoff rate q_s [m.s^{-1}] are calculated as a weighted-average of the

component specific rates:

$$ET = f_{bs}E_{bs} + f_{vg}\left(E_{wc} + E_{vg}\right) \quad (6.23)$$

$$q_s = f_{bs}q_{bs} + f_{vg}q_{vg} \quad (6.24)$$

8. The states are updated by putting the initial states in the next time step equal to the states obtained in the previous time step. Steps 2 to 7 are then repeated with the updated states.

The hydrological model can be operated within two different frameworks to represent the spatial variability: either within an explicit framework or within a statistical framework (*cfr.* Figure 6.1). The way in which the soil water is redistributed in the catchment depends on the selected framework. In the explicit framework, the catchment is discretized into a finite number of grid cells (computational unit), each characterized by its land cover, soil type and topographic index. For each grid cell, the SVATS is applied and the water is redistributed based on the local water table depth. In the statistical framework, a statistical distribution of the soil-topographic index is employed for each land cover class. This distribution is discretized into a number of intervals and the SVATS is applied to each interval (computational unit). As such, water redistribution is based on the interval-specific water table depths per land cover. Running TOPLATS in statistical mode has the advantage that the computational cost is considerably reduced. However, this is at the expense of a reduced detail in the spatial information of both model inputs and outputs.

6.3. Inputs and parameters

TOPLATS requires a lot of inputs (model forcings) and parameters (quantities to characterize the system). Depending on the type of input or parameter, its value can be (i) spatially distributed or spatially constant and (ii) time-varying or time-invariant. Considering the high number of inputs and parameters, only the most important ones are discussed below and are grouped w.r.t. the component within the SVATS that they characterize: the atmosphere (Table 6.1), the soil (Table 6.2), the topography (Table 6.3) or the vegetation (Table 6.4). Additionally, some parameters need to be specified w.r.t. the model geometry (e.g. depth of the soil layers, the root zone fractions, the soil temperature depths) and the simulation control in space and time (e.g. the number of time steps in the simulation, the time step duration and the pixel resolution).

Meteorological data

A time series of a set of meteorological variables is required to force the model. The time stepping and duration of the measurements need to correspond to the

simulation timing control. Although the meteorological variables can be entered as spatially variable inputs (images) in TOPLATS, they are usually measured at a single location and are assumed to be representative for the entire catchment. An overview of the model inputs that describe the atmospheric conditions is given in Table 6.1.

Table 6.1: Meteorological inputs (time-varying) in the hydrological model.

Input/Parameter	Notation	Units
Longwave radiation	R_l	[W.m^{-2}]
Shortwave radiation	R_s	[W.m^{-2}]
Dry bulb temperature	T_b	[K]
Relative humidity	H_{rel}	[-]
Wind speed	u	[m.s^{-1}]
Air pressure	p_a	[Pa]
Precipitation	P	[m.s^{-1}]

Soil data

A soil map is needed to indicate the corresponding soil type for the different grid cells within the catchment. To each soil type, a number of physical soil properties are attributed and are summarized in a soil type look-up table (ST-LUT). The model then uses the corresponding soil properties (*cfr.* Section 6.1.4) to run the local SVATS at a given location. Values for the soil properties can be taken from literature or from field measurements. An overview of the soil parameters from the ST-LUT is given in Table 6.2. In case TOPLATS is run in statistical mode, the soil properties are assumed constant for the entire catchment.

Table 6.2: Soil parameters (time-invariant) from the soil type look-up table (ST-LUT) in the hydrological model.

Input/Parameter	Notation	Units
Saturated hydraulic conductivity	K_s	[m.s^{-1}]
Saturated soil moisture content	θ_s	[m^3.m^{-3}]
Residual soil moisture content	θ_r	[m^3.m^{-3}]
Pore size distribution index	λ	[-]
Bubbling pressure	ψ_c	[m]
Dry soil heat capacity	C_s	[J.K^{-1}.kg^{-1}]
Quartz content	q	[-]
Minimum soil resistance	$r_{s,min}$	[s.m^{-1}]
Soil bulk density	ρ_b	[kg.m^{-3}]

Topographic data

A topographic index map and a catchment map, which delineates the boundaries of the (sub)catchment(s), are required as inputs for the hydrological model. These maps are usually derived from a digital elevation model (DEM) of the region, prior

to model application. To each subcatchment within the study area, a number of properties are attributed (these are the parameters from the TOPMODEL) and are listed in a catchment look-up table (Table 6.3). If TOPLATS is run in statistical mode, no explicit spatial information about the topography is used, instead a probability distribution of the topographic index for each land cover type should be supplied by the model user.

Table 6.3: TOPMODEL parameters (time-invariant) from the catchment look-up table (CA-LUT) in the hydrological model.

Input/Parameter	Notation	Units
Baseflow at complete saturation	Q_0	$[m^3.s^{-1}]$
Hydraulic conductivity decay coefficient	f_K	$[m^{-1}]$
Initial average water table depth	$\overline{z}_0$	[m]

Vegetation data

A land cover map is used to attribute a land cover type to each grid cell within the catchment. In a land cover look-up table (LC-LUT), each land cover type is characterized by a set of model parameters. Some parameters are assumed to be constant in time (time-invariant parameters) while others can be updated following a user-defined time-interval (time-varying parameters). An overview of the vegetation parameters from the LC-LUT is given in Table 6.4. If TOPLATS is run in statistical mode, no explicit spatial information about the land cover is used, instead a fractional coverage of the different land cover types in the catchment needs to be specified.

Table 6.4: Vegetation parameters from the land cover look-up table (LC-LUT) in the hydrological model.

Input/Parameter	Notation	Units
Time-invariant parameters		
Emissivity	ϵ	[-]
Height of thermodynamic measurements	z_a	[m]
Height of wind speed measurements	z_w	[m]
Minimum stomatal resistance	$r_{st,min}$	$[s.m^{-1}]$
Maximum stomatal resistance	$r_{st,max}$	$[s.m^{-1}]$
Time-varying parameters		
Leaf area index	LAI	[-]
Albedo	α	[-]
Roughness length for momentum transfer	$z_{0,m}$	[m]
Roughness length for heat transfer	$z_{0,h}$	[m]
Zero plane displacement height	d_0	[m]

6.4. Runoff routing

TOPLATS calculates for each time step in the simulation the amount of baseflow leaving the catchment and the amount of surface runoff generated in the catchment. As the generated runoff at a particular location in the catchment flows over the land surface and travels in the river channels before it reaches the catchment outlet, there is a time lag between the moment of runoff generation and the moment when runoff leaves the catchment. The required time correction on the generated runoff is known as runoff routing and is a mathematical procedure to follow water movement from one place to another. Practically, the problem comes down to constructing the discharge hydrograph at a given point along the stream. Many different methods for runoff routing have been developed, of which most of them solve the Saint Venant equations to describe one-dimensional water transport on hillslopes and in open channels. In this thesis, it is chosen to apply the runoff routing model developed by Troch et al. (1994).

The unit hydrograph, which expresses the catchment response to an instantaneous unit input of runoff, is given by (Troch et al., 1994):

$$f_b(t) = \int_0^\infty f_h(t-\tau) f_c(\tau) d\tau \tag{6.25}$$

where t represents the time [s], and τ the time lag [s]. Above expression combines the response function for open channel flow (f_c) and the response function for overland flow (f_h). The unit impulse response function for open channel flow is expressed as:

$$f_c(t) = \int_0^\infty q_c(h,t) W(h) dh \tag{6.26}$$

where h is the distance along the channel [m], q_c is the discharge rate per unit width of channel caused by channel flow [$m^2.s^{-1}$] and $W(h)$ is the normalized width function. The latter represents the distribution of runoff entering the river network at a distance h and is determined as the ratio between the number of channel links at a given distance from the outlet and the total channel length. The unit impulse response function for overland flow is expressed as:

$$f_h(t) = \int_0^\infty q_h(h,t) H(h) dh \tag{6.27}$$

where q_h is the discharge rate per unit width of channel caused by overland flow [$m^2.s^{-1}$] and $H(h)$ is the normalized hillslope function, which can be interpreted as the probability density of runoff generated randomly in the catchment at a distance h from the river network. It is determined as the ratio between the number of computational units at a distance h from the river network and the total number of computational units that do not belong to the river network. The discharge

rates q_c and q_h are calculated following an analytical solution of the Saint Venant equations when neglecting the acceleration terms (Brutsaert, 1973):

$$q = \frac{h}{(2\pi)^{1/2} b t^{3/2}} e^{-(h-vt^2)/(2b^2 t)} \tag{6.28}$$

The parameters v [m.s^{-1}] and b [m^2.s^{-1}] are respectively referred to as the drift velocity and the diffusion coefficient, and are defined as:

$$v = (1 + a_0)V \tag{6.29}$$

$$b^2 = \frac{V^3}{g S_0 F^2} \left(1 - a_0^2 F^2\right) \tag{6.30}$$

where a_0 is an empirical constant [-], V is the vertically averaged velocity [m.s^{-1}], g is the acceleration of gravity [9.81 m.s^{-2}], S_0 is the channel bed slope [-] and F is the Froude's number [-]. The latter is calculated as $V/\sqrt{gH}$, where H is the water height above the bed in the river or hillslope [m]. Once the unit hydrograph $f_b(t)$ is known, the runoff at the outlet of the catchment Q_s [m^3.s^{-1}] can be calculated following a discrete convolution equation:

$$Q_s(t) = \sum_{i=1}^{t \leq N} A q_s(i) f_b(t - i + 1) \tag{6.31}$$

with A the area of the catchment [m^2], $q_s(i)$ the local surface runoff [m.s^{-1}] in the i-th time interval as simulated by TOPLATS (*cfr.* Eq. (6.24)) and N the total number of time steps. The total discharge at the outlet is then calculated as the sum of the baseflow $Q_b(t)$ (*cfr.* Eq. (6.19)) and the routed surface runoff $Q_s(t)$ (*cfr.* Eq. (6.31)).

7

Catchment and data description

The study site for application of the hydrological model is the catchment of the Bellebeek, which is a subcatchment of the Dender catchment in Belgium. The climate is temperate with a mean annual temperature of 11.5°C and a total annual rainfall of 750 mm (uniformly distributed throughout the year). The surface area of the catchment is 91.24 km^2 and is dominated by agricultural crops and pasture, intersected with urban area, forests and open water bodies. Figure 7.1 shows the location of the Bellebeek catchment. A list of the available data for the Bellebeek catchment is given below.

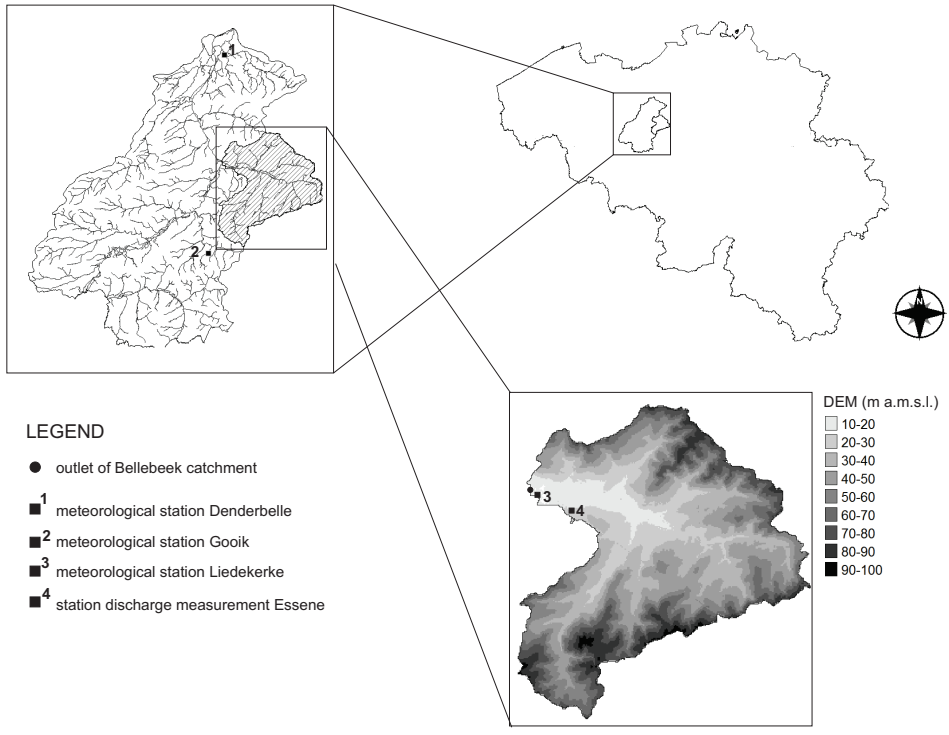


Figure 7.1: Situation of the Bellebeek catchment in Belgium with indication of the river network and the location of the measurement stations.

7.1. Geographical data

Geographical data for the Bellebeek catchment include a digital elevation model, a land cover map and a soil map. Most geographical data were available through different sources. Therefore, we selected two data sets for both the soil and land cover information, each differing in spatial resolution and/or in information content.

7.1.1. Digital elevation model

A digital elevation model (DEM) of the Bellebeek catchment (shown in Figure 7.1) was available through the Flanders Geographical Information Agency (FGIA). The map has a spatial resolution of 25 m and an accuracy of 0.07 m. Elevation in the catchment ranges between 10 and 100 m a.m.s.l. Based on the DEM and the location of the outlet (*cfr.* Figure 7.1), the catchment area was delineated by applying the algorithms of Jenson and Domingue (1988).

7.1.2. Soil maps

Soil texture information of the Bellebeek catchment was readily available through the Food and Agriculture Organization (FAO) and the Flemish government. The soil data sets differ in spatial resolution and information content and are briefly described below. An overview of the class labels in the soil maps is given in Table 7.1.

1. A soil map of the study area was extracted from the FAO digital soil map (1 km resolution). The soil types on the soil map correspond to the USDA soil texture classification (*cfr.* Figure 7.2). The particle size distribution used in the USDA classification scheme is as follows: the clay fraction has a diameter less than 2 μm, the sand fraction has a diameter larger than 2 mm and the silt fraction has a diameter between 2 μm and 2 mm. The FAO soil map of the catchment is further referred to as STM-FAO and is shown in Figure 7.3. The soil map indicates a dominant presence of loam and silty loam soils with local patches of clay soils.

2. Through the Flanders Geographical Information Agency, a soil map with a spatial resolution of 25 m was available for the Bellebeek catchment. This soil map indicates the soil type based on the Belgian soil classification scheme (*cfr.* Figure 7.2), which uses the same particle size distributions as in the USDA soil classification scheme. The Belgian soil map of the Bellebeek catchment is shown in Figure 7.3 and is further referred to as STM-BEL. Silt loam and sandy silt loam soils are dominantly present in the catchment.

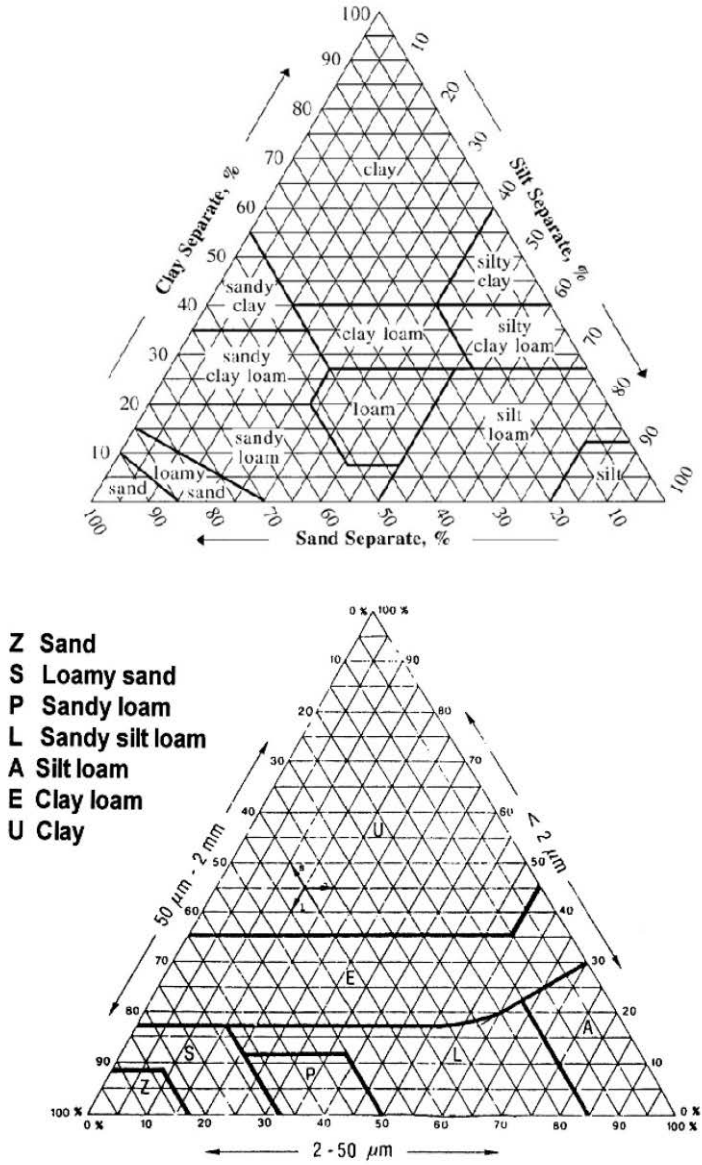


Figure 7.2: USDA soil classification triangle (top) and Belgian soil classification triangle (bottom).

Table 7.1: Overview of the different soil types in the soil type maps (STMs) for the Bellebeek catchment.

STM-FAO (1 km res.)			STM-BEL (25 m res.)		
Code	Soil Type	Area [%]	Code	Soil Type	Area [%]
L	loam	26.61	S	loamy sand	0.11
SiL	silt loam	73.39	P	sandy loam	4.30
			L	sandy silt loam	15.46
			A	silt loam	72.34
			E	clay loam	2.36
			G	stony	0.03
			V	peat	0.03
			Ur	urban	5.40

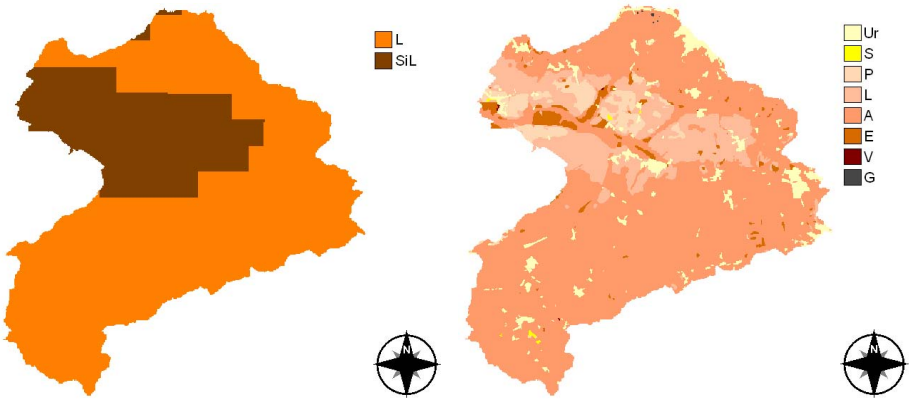


Figure 7.3: FAO soil map STM-FAO (left) and Belgian soil map STM-BEL (right) for the Bellebeek catchment. The class labels in the legends are explained in Table 7.1.

7.1.3. Land cover maps

Land cover information for the Bellebeek catchment is readily available through:

1. The Flanders Geographical Information Agency (FGIA): provides a land cover map at a resolution of 15 m (Figure 7.4). The classification scheme of the map includes 6 different land covers: urban area, water area, broad-leaved forest, needle-leaved forest, pasture and cropland. This general land cover map of the Bellebeek is further referred to as LCM-GEN.

2. The department Mestbank of the Flemish Land Agency (VLM): provides a databank with information on the crop type for a large number of agricultural fields. In the Bellebeek catchment, 7 specific land covers (crop types) were distinguished: grass (mixture of grass species and leguminosae species),

potatoes, beets, maize, wheat, barley and other crop species. However, for only 45% of the agricultural fields in the Bellebeek catchment, information on the crop type was available. The crop type information was therefore merged with the general land cover map to result in a more detailed land cover map of the catchment. This land cover map is further referred to as LCM-CROP (Figure 7.4).

An overview of the class labels in the land cover maps is given in Table 7.2.

Table 7.2: Overview of the different land cover types in the available land cover maps (LCMs) for the Bellebeek catchment.

LCM-GEN (15 m res.)			**LCM-CROP** (15 m res.)		
Code	Land Cover	Area [%]	Code	Land Cover	Area [%]
Pa	pasture	19.96	Pa	pasture	9.87
Cr	cropland	65.22	Cr	cropland	35.42
Bf	broad-leaved forest	4.12	Bf	broad-leaved forest	3.89
Nf	needle-leaved forest	0.01	Nf	needle-leaved forest	0.01
Wa	water bodies	0.11	Wa	water	0.11
Ur	urban area	10.58	Ur	urban area	10.79
			Gr	grassland	16.04
			Po	potatoes	2.43
			Be	beets	1.81
			Ma	maize	9.30
			Wh	wheat	6.60
			Ba	barley	1.29
			Oc	other crop type	2.45

7.2. Remote sensing data

Three cloud-free satellite images of the Bellebeek catchment were acquired on June 13 (DOY 164), June 30 (DOY 181) and July 5 (DOY 186) 2006, with a spatial resolution of 20 m. On June 13 and June 30, reflectance data in the green ($0.50 - 0.59\,\mu$m), red ($0.61 - 0.68\,\mu$m), near-infrared ($0.79 - 0.89\,\mu$m) and mid-infrared ($1.58 - 1.75\,\mu$m) wavelength regions were acquired by the SPOT 4 HRVIR. Mid-infrared radiances were not available for July 5 as the reflectance data were acquired by SPOT 2 HRV, on which a mid-infrared band is not present. The satellite images were orthorectified and geo-referenced. A radiometric correction was carried out to remove distortions due to differences in the sensitivity of the elementary detectors of the viewing instrument. Atmospheric correction was not carried out as the images were cloud-free such that the atmospheric influence

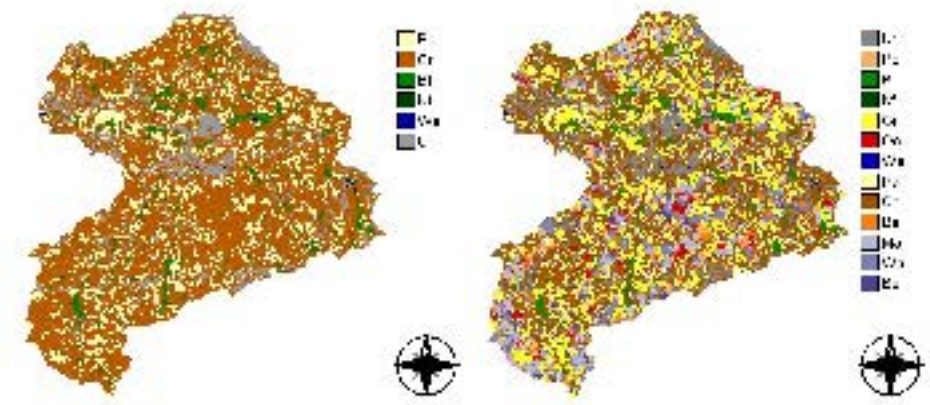


Figure 7.4: General land cover map LCM-GEN (left) and detailed land cover map LCM-CROP with indication of the crop type (right) for the Bellebeek catchment. The class labels from the legend are explained in Table 7.2.

was assumed constant over the entire image. Further, the optical images are predetermined for land cover classification such that comparison of the radiances among the different acquisition dates is not relevant.

7.3. Meteorological data

Wind speed, relative humidity, net radiation, rainfall, atmospheric pressure and temperature (dry bulb, wet bulb, dew point) were registered with a temporal resolution of 10 to 60 minutes at the meteorological station of Liedekerke, situated near the outlet of the catchment. Wind speed was measured at a height of 10 m while the other variables were measured at 2 m height. Missing data were complemented by measurements from nearby meteorological stations (at Gooik and Denderbelle). Measurements of incoming shortwave radiation were not readily available, but were calculated from the net radiation based on a regression (with a correlation of 0.96) between the shortwave and net radiation, measured at the meteorological station of Gooik (approximately 2 km south of the catchment). All meteorological forcings were collected for the period January 1, 2006 to December 31, 2007 and were aggregated to an hourly time step. The meteorological records point out that the weather conditions in the catchment of the Bellebeek apply to a temperate climate with an annual mean temperature of 11.5 °C and a total annual rainfall of 750 mm. The location of the meteorological stations is shown in Figure 7.1.

7.4. Soil moisture and discharge measurements

Discharge observations for the period January 1, 2006 to December 31, 2007 were available with an hourly time step at the outlet of the Bellebeek catchment. The discharge was derived from measurements of the water height through a so-called $Q - h$ relation. The water height was measured by means of a limnigraph with an accuracy of 1 cm. Measurements of the flow velocity at the limnigraph were used to construct the relationship between the discharge and the water height. The location of the discharge measurement station of Essene is shown in Figure 7.1.

At the station in Liedekerke, in situ soil moisture measurements were taken at 2.5 cm depth between May 13 and May 30, 2007 by means of a time domain reflectometer (TDR). This measurement point has coordinates 50.89 °N and 4.09 °E.

8

Calibration of the hydrological model

Hydrological models are defined by their states and parameters. Generally, two types of parameters exist: physical parameters and process parameters. The physical parameters can be measured directly in the field and are independent of the system response. The process parameters, on the contrary, cannot be measured directly in the field (at the relevant scale) and need to be inferred by indirect means (Gupta et al., 1998). Although these parameters are usually not expected to have a physical meaning, they are assumed to be related to system properties that have a physical relevance in hydrological processes. In order for a hydrological model to properly simulate the observed response of the catchment, the (process) model parameters need to be tuned or calibrated. Measurements of the forcing data and system response over a historical period of time are therefore required.

The hydrological model can be denoted as a non-linear function f which expresses the relationship between the model inputs (forcings) $\mathbf{x}$, the model parameters ϕ and the model outputs $\mathbf{y}$:

$$\mathbf{y} = f(\mathbf{x}, \phi) \tag{8.1}$$

Since it is impossible to find an exact solution of ϕ for a given set of observed outputs $\mathbf{y}$, the calibration problem reduces to an optimization problem in which an optimal set of parameters $\phi_{\mathbf{opt}}$ is determined such that the estimated output $\mathbf{y} = f(\mathbf{x}, \phi_{\mathbf{opt}})$ closely resembles the observed output $\mathbf{y}_{\mathbf{obs}}$. To solve the optimization problem, an appropriate calibration algorithm needs to be selected.

8.1. Calibration methods

A basic approach to the parameter optimization is the trial and error procedure, which involves a manual adjustment of the parameter values and is therefore labor-intensive. Furthermore, manual calibration uses a subjective process of visual inspection and evaluation of model performance measures (Boyle et al., 2000). Because of these disadvantages, automatic methods for calibration have become more popular (e.g. Sorooshian et al., 1993, Duan et al., 1992, Hogue et al., 2000, Boyle et al., 2000). Although such methods aim at finding a single best set of parameter values by minimizing an objective function o (a measure for

model performance), many studies have reported difficulties in finding this unique parameter estimates (e.g. Sorooshian et al., 1993, Duan et al., 1992) as many combinations of parameter sets may result in a similar objective function value. This problem is known as equifinality (Beven, 1993) (for more explanation the reader is referred to Section 3.3.1). Because a poor identifiability of the parameters may result in a considerable uncertainty in the model output, parameter uncertainty can be taken into account during model calibration by means of uncertainty analysis methods such as the Generalized Likelihood Uncertainty Estimation (GLUE, *cfr.* Section 3.3.1) (Beven and Binley, 1992), or Bayesian Markov chain Monte Carlo methods (MCMC) (Kuczera and Parent, 1998, Vrugt et al., 2003).

Quantifying the uncertainty in the process parameters is out of the scope of this dissertation, as the focus is on uncertainty in the land surface characterization, although we are aware that uncertainty in the calibrated (and hence assumed optimal) parameter values introduces additional uncertainty in the model output. Calibration of the hydrological model is therefore assumed to be well-posed such that the success of calibration does not depend on the model structure or on the observed data. Three different calibration algorithms were applied to calibrate the TOPLATS model: shuffled complex evolution (SCE), dynamically dimensioned search algorithm (DDS) and weight-adaptive recursive parameter estimation (WARPE). For an overview of global optimization algorithms for watershed model calibration, we refer to Duan (2003).

8.1.1. Shuffled complex evolution

SCE is a commonly used genetic algorithm for the calibration of hydrological models and was developed at the University of Arizona in 1992 (Duan et al., 1992). The goal of the method is to find the global optimum by minimizing an objective function. SCE is able to calibrate a large number of parameters, is robust to variations in parameter sensitivities and interactions, and small minima and maxima in the objective function do not hamper the optimization. The method is briefly summarized below, for a detailed description we refer to Duan et al. (1992, 1994).

The SCE algorithm roughly works as follows:

1. s points are randomly sampled from the feasible parameter space and the objective function o is computed for each sampled point. If no a priori information about the location of the global optimum is available, a uniform distribution is used to generate the samples.

2. The sampled points are ranked in order of increasing objective function value such that the first point receives the lowest value.

3. The ranked points are partitioned into l complexes, in which each complex

receives m points. The complexes are partitioned such that the first complex receives every $l(k-1)+1$ ranked point (with $k = 1, 2, ..., m$), the second complex receives every $l(k-1)+2$ ranked point (with $k = 1, 2, ..., m$), and so on. The distribution of the population in different complexes allows for a more free and more complete assessment of the parameter space.

4. Each complex is evolved according to the competitive complex evolution (CCE) algorithm, which is based on the Nelder and Mead (1965) Simplex search algorithm. In the CCE algorithm, subcomplexes are constructed by randomly selecting n points from the complex. Each subcomplex functions like a pair of parents, with the ability to participate in the process of reproducing offspring. Stronger parents survive better and produce healthier offspring than the weaker ones. For a detailed explanation of this algorithm, we refer to Duan et al. (1994).

5. The points in the evolved complexes are combined into a single population, are sorted in order of increasing objective function value and are re-partitioned into l complexes. This step is referred to as shuffling of the complexes and allows for an interchange of information.

6. The convergence of the algorithm is checked: if the convergence criterion (e.g. spread in the ranking) is satisfied, the algorithm is stopped. If this is not the case, the algorithm is repeated from step 3.

The SCE method contains many probabilistic and deterministic components that are controlled by some algorithmic parameters (m, l en n). These parameter values must be chosen carefully in order to reach an optimal performance of the method. Default values are given by Duan et al. (1994).

8.1.2. Dynamically dimensioned search

The DDS method is a simple single-solution based heuristic global search algorithm with as goal finding a good global solution within a specified maximum number of model evaluations (this is the only stopping criterion). The algorithm scales the search algorithm from global in the initial iterations to more local in the final iterations by dynamically and probabilistically reducing the dimension of the neigbourhood (*i.e.* the number of parameter values being changed to generate a new search neigbourhood) (Tolson and Shoemaker, 2007). The method is briefly summarized below, for a detailed description we refer to Tolson and Shoemaker (2007).

The DDS algorithm roughly works as follows:

1. The required inputs are defined: the maximum number of function evaluations N_{eval}, an initial solution $\phi_{\mathbf{0}}$, the lower bounds $\phi_{\mathbf{min}}$ and the upper bounds $\phi_{\mathbf{max}}$ for all parameters.

2. The objective function o is evaluated at the initial solution. At this point, the best solution $\phi_{\mathbf{opt}}$ equals $\phi_{\mathbf{0}}$.

3. From the parameter set, k parameters are randomly selected for inclusion in the search neigbourhood:

   (a) The probability that each parameter is selected, is calculated as a function of the current iteration number.

   (b) The parameters are added to the neigbourhood following the obtained probability.

   (c) If the neigbourhood is empty, a random parameter is selected and added to the neigbourhood.

4. The k parameters from the neigbourhood are perturbed using a standard normal random variable $N(0,1)$ and the perturbation size parameter r_p such that $\phi_{\mathbf{pert}} = \phi_{\mathbf{opt}} + r_p \left(\phi_{\mathbf{max}} - \phi_{\mathbf{min}}\right) N(0,1)$.

5. The objective function o is evaluated for the perturbed parameter set and the solution is updated if $o(\phi_{\mathbf{pert}}) \leq o(\phi_{\mathbf{opt}})$.

6. The stopping criterion is checked: if the maximum number of function evaluations N_{eval} is reached, the algorithm is stopped. If this is not the case, the algorithm is repeated from step 3.

The only parameter to set in the DDS method is the neigbourhood size perturbation parameter r_p, for which a default value of 0.2 is recommended (Tolson and Shoemaker, 2007).

8.1.3. Weight-adaptive recursive parameter estimation

WARPE is a method for model calibration that is based on the equations of the Extended Kalman Filter (Pauwels, 2008). The goal of the method is to minimize the uncertainty in the estimated model parameters. The consequence of this theoretical foundation is that if multiple observations are available, the mismatch between the observations and the simulations is not lumped together into a single objective function, but the parameter update procedure explicitly takes into account the mismatch for every single observation.

The WARPE algorithm roughly works as follows:

1. The parameter vector $\phi_{\mathbf{k}}$, the matrix $\mathbf{Q_k}$ containing the variance of the noise on the model parameters (assumed to be constant throughout the iteration), and the error covariance matrix $\mathbf{P_k}$, of which the diagonal values indicate the degree of uncertainty in the estimated parameter values, are initialized. The index k refers to the iteration number, initially equal to zero.

2. The observations of the model output are entered in the vector $\mathbf{y_{obs}}$ and the variance of the noise on the observations is entered in the matrix $\mathbf{R_k}$ (assumed to be constant throughout the iteration).

3. The parameter vector and the error covariance matrix are propagated from iteration $k-1$ to iteration k as follows:

$$\begin{cases} \phi_k^- = \phi_{k-1}^+ \\ \mathbf{P}_k^- = \mathbf{P}_{k-1}^+ + \mathbf{Q}_{k-1} \end{cases} \tag{8.2}$$

   where the superscript $.^+$ indicates an a posteriori estimate, *i.e.* after adjustment using the model output observations, and the superscript $.^-$ indicates an a priori estimate, *i.e.* before adjustment using the observations.

4. The simulated model output $\mathbf{y}_k^-$ is calculated by running the hydrological model with the parameter values from ϕ_k^- and the forcings $\mathbf{x}$.

5. The Jacobian matrix $\mathbf{H}_k$ is calculated as follows:

$$\mathbf{H_k}(i,j) = \frac{\partial f(\mathbf{x}, \phi_{\mathbf{k}}^-)(i)}{\partial \phi(j)} \tag{8.3}$$

6. The parameter vector and the error covariance matrix are updated using the observations:

$$\begin{cases} \mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} \left[\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \right]^{-1} \\ \phi_k^+ &= \phi_k^- + \mathbf{K}_k \left(\mathbf{y_{obs}} - \mathbf{y}_k^- \right) \\ \mathbf{P}_k^+ &= \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right] \mathbf{P}_k^- \end{cases} \tag{8.4}$$

   where $\mathbf{K}_k$ is the Kalman gain factor.

7. Steps 3 to 6 are repeated until convergence is achieved or when a predefined number of iterations N_{eval} has been reached.

8.2. Hydrological model setup

The hydrological model TOPLATS was used to simulate the discharge for the Bellebeek catchment. Therefore, we forced the model with the climatic data from the meteorological stations at Liedekerke and Gooik (*cfr.* Section 7.3). These point measurements were assumed to be representative for the entire study area and rainfall was assumed to be uniformly distributed over the catchment. Information on the soil type and the land cover was provided by respectively STM-FAO and LCM-GEN, information on the topography was provided by the DEM (*cfr.* Section 7.1). All geographical data of the study area were converted to a grid with a resolution of 50 m using a nearest neighbour resampling technique, as higher spatial

resolutions compromised the computational resources (CPU of 1 model run at 25 and 50 m resolution is respectively 32 and 8 hours). The resulting data consist of a grid with 280 rows and 264 columns, with a total of 73920 grid cells.

Values for the model parameters were taken from literature or were based on expert knowledge. The soil related parameters (e.g. hydraulic conductivity, heat capacity, soil resistance) were taken from literature (Vereecken et al., 1989, Rawls et al., 1982, Sleutel et al., 2006, Peters-Lidard et al., 1997) and were used to construct the soil type look-up table, which characterizes the different soil types from the soil map (*cfr.* Table 6.2). For the different land covers, both time-invariant and time-varying vegetation parameters (e.g. albedo, LAI) were specified and were used to construct the land cover look-up table (*cfr.* Table 6.4). In this study, we dealt with five time-varying vegetation parameters of which the value was updated on a monthly time basis: leaf area index (LAI), albedo (α), momentum roughness ($z_{0,m}$), heat roughness ($z_{0,h}$) and zero plane displacement height (d_0). Monthly values for LAI and α were taken from literature (Tian et al., 2004, Breuer et al., 2003, Tekalign and Hammes, 2005, Pauwels et al., 2007, Kotoda, 1986). Values for $z_{0,m}$ and d_0 were calculated from the canopy height following the relationship as reported in literature for forest, bare soil, cropland and grassland (respectively, Jarvis et al. (1976), Jones (1992), Campbell (1977) and Allen et al. (1988)). Values for the maximal crop height were taken from Breuer et al. (2003). The crop height was put equal to zero outside the growing season and was assumed to linearly increase within the growing period (Allen et al., 1988). The heat roughness $z_{0,h}$ was calculated from $z_{0,m}$ following the relationship $z_{0,h} = 0.1 \cdot z_{0,m}$ (Allen et al., 1988). Other vegetation parameters such as emissivity and stomatal resistance were assumed to be time-invariant for which values were taken from literature (Peters-Lidard et al., 1997, Chen et al., 1996, Jacquemin and Noilhan, 1990).

8.3. Results and discussion

In this part of the thesis, the TOPMODEL parameters and the parameters from the routing model were automatically calibrated in two sequential calibrations. Other parameters such as soil (*cfr.* Table 6.2) and vegetation parameters (*cfr.* Table 6.4) were not considered for calibration because the physical meaning of these values might be lost when treating them as tuning parameters. This is an important aspect w.r.t. the subsequent chapters in which these model parameters are the subject of uncertainty and sensitivity analysis and their values are varied within literature-based ranges that were derived from field measurements. In the first calibration step, the TOPMODEL parameters were tuned in order to obtain a good representation of the baseflow and the amount of generated runoff in the catchment. In the second calibration step, the routing parameters were tuned in order obtain a good representation of the shape of the peak discharges. Because of

the high computational cost of automatic calibration, the model was calibrated in aggregated statistical mode after which the calibrated parameter values were transferred to the distributed mode. As such, the computation time of one model application could be reduced from eight hours to two minutes. The discharge at the outlet of the Bellebeek was simulated for the period January 1, 2006 to December 31, 2006 as the sum of the baseflow and the routed surface runoff. Surface runoff was routed to the catchment outlet using the unit hydrograph approach as described by Troch et al. (1994) (*cfr.* Section 6.4). Finally, simulated and observed discharge (*cfr.* Section 7.4) were aggregated to an hourly time-basis and were compared by means of goodness-of-fit statistics.

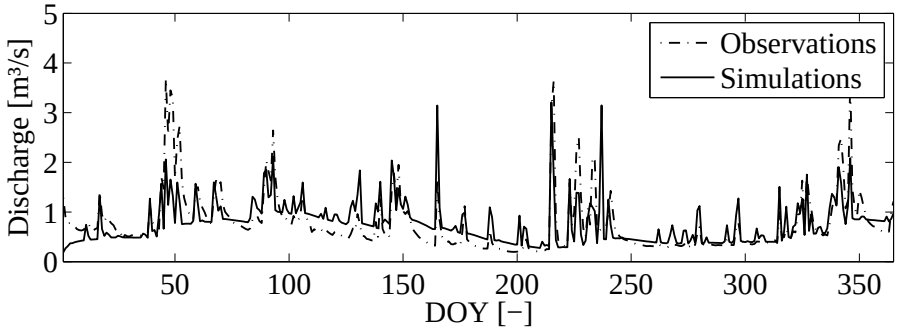
8.3.1. Calibration of the TOPMODEL parameters

The model parameters to be calibrated are the three TOPMODEL parameters: the baseflow at saturation Q_0 [$\mathrm{m}^3 \cdot \mathrm{s}^{-1}$], the hydraulic conductivity decay coefficient f_K [m^{-1}] and the initial average water table depth $\overline{z}_0$ [m]. The calibration methods used are SCE, DDS and WARPE. With the SCE and DDS method, an objective function needs to be minimized. Here, it was chosen to use the root mean squared error (RMSE), defined as:

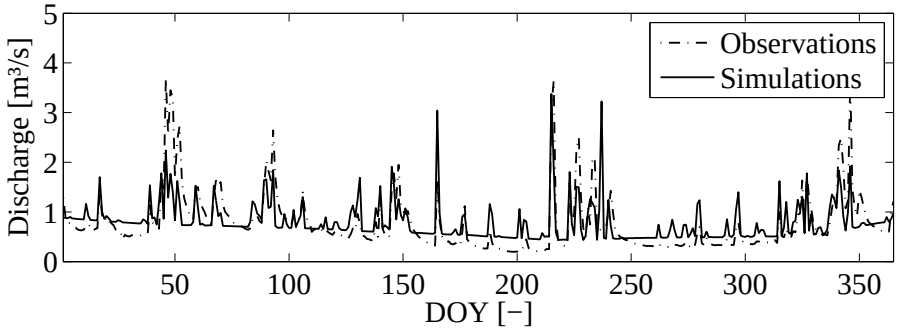
$$RMSE = \sqrt{\frac{\sum_{t=1}^{N} \left(y_{obs}(t) - y(t)\right)^2}{N}} \tag{8.5}$$

where N is the total number of time steps in the simulation (here $N = 365$), $y(t)$ is the observed model output at time step t and $y(t)$ is the simulated model output at time step t. For the algorithmic parameters in the SCE and DDS method, default values were taken from literature (Duan et al. (1994) and Tolson and Shoemaker (2007), respectively). For the calibration with WARPE, the error on the observations was defined in the matrix $\mathbf{R}_k$ as 1% of the observed value. The diagonal values in the matrix $\mathbf{Q}_k$ were set equal to the square of 10% of the corresponding parameter value in ϕ_k^-. Initially, the error covariance matrix $\mathbf{P}_k$ was set equal to $\mathbf{Q}_k$. All calibration algorithms were stopped after 1000 model evaluations ($N_{\mathrm{eval}} = 1000$). The initial values, the lower and upper bounds for the TOPMODEL parameters are listed in Table 8.1, together with the calibrated parameter values. The routing parameters were given a constant value of 0.2 (*cfr.* Table 8.3).

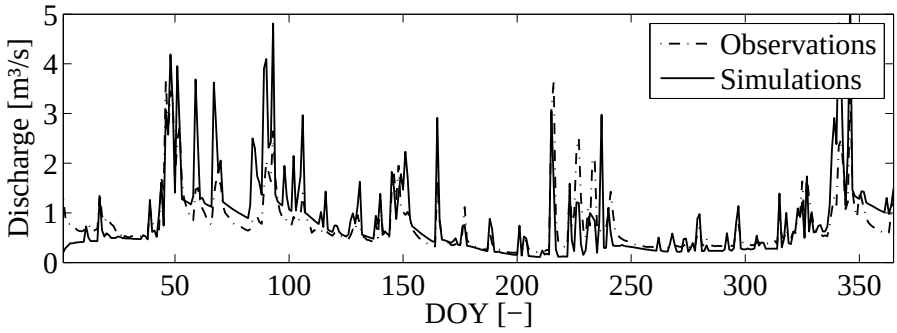
From Table 8.1, it can be seen that the obtained value for each of the parameters largely differs among the different calibration algorithms. This means that the optimization problem is ill-posed and that the resulting optimal parameter set depends on the choice of the method used. The calibrated values obtained with the SCE method lie near the specified upper bounds of the parameters which suggests a suboptimal performance of this calibration method. The goodness-of-fit for the



(a)

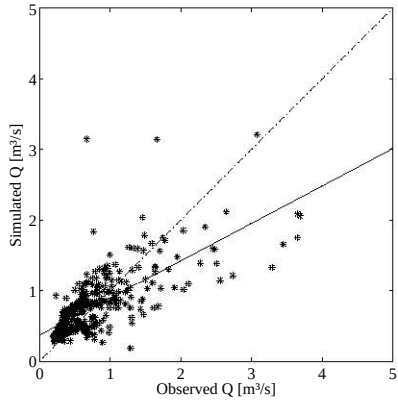


(b)

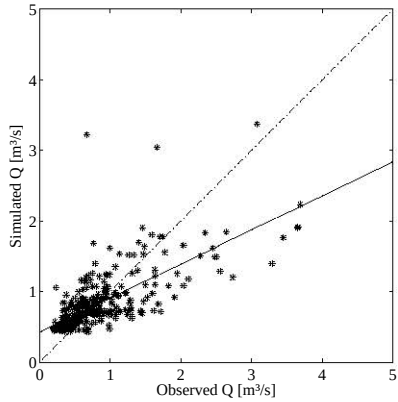


(c)

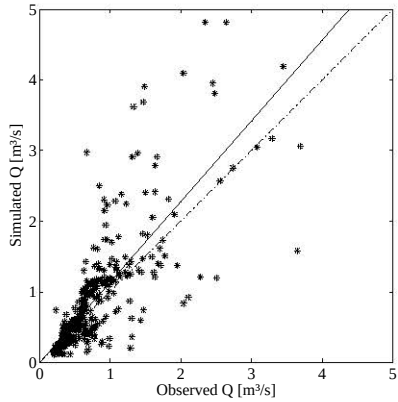
Figure 8.1: Timeseries of the observed discharge and the simulated discharge for the year 2006 obtained with the calibrated TOPMODEL parameters (TOPLATS run in statistical mode) as obtained with the different calibration methods: (a) SCE, (b) DDS, (c) WARPE.



(a)



(b)



(c)

Figure 8.2: Scatterplot between the observed and simulated discharge for the year 2006 as obtained with the calibrated TOPMODEL parameters from the SCE (a), DDS (b) and WARPE method (c) (TOPLATS run in statistical mode).

Table 8.1: Initial, minimum, maximum and calibrated value for the TOPMODEL parameters as obtained with the different calibration algorithms.

Parameter	Minimum	Maximum	Initial	Calibrated		
				SCE	*DDS*	*WARPE*
Q_0 [$m^3.s^{-1}$]	2.0	12.0	6.0	11.99	9.63	6.31
f_K [m^{-1}]	1.0	7.0	3.0	1.32	1.17	2.51
$\overline{z}$ [m]	0.5	4.0	2.5	3.39	1.93	1.51

different calibration results is evaluated in terms of the RMSE and the Pearson correlation coefficient $r_{y\hat{y}}$ (Table 8.2), and shows that the SCE method results in the lowest RMSE, but that the WARPE method results in the highest correlation. Based on these statistics, no clear statement about the best calibration result can be made. However, a visual inspection of the calibration result, *i.e.* comparing the time series of the calibrated discharge with the observed discharge, reveals that the difference between the calibration methods is clearly expressed w.r.t. the dynamics in the simulated discharge (Figure 8.1). Although the SCE method resulted in the lowest value of the objective function, the calibrated discharge was not able to capture the dynamics in the observed discharge. Especially the amount of generated surface runoff was ill-captured. The same problem was observed with the DDS method, where also the dynamics in the baseflow were ill-captured by the calibration results and were completely smoothed. On the contrary, the calibrated discharge as obtained with the WARPE method, closely resembles the observed discharge, as reflected by the high correlation coefficient. The same conclusion can be drawn from the scatterplot between the observed and calibrated discharge (Figure 8.2) which indicates an underestimation of the peak discharge and an overestimation of the baseflow for both the SCE and DDS method.

Table 8.2: Statistics of the calibrated discharge as obtained with the different calibration algorithms.

	SCE	DDS	WARPE
RMSE	0.4016	0.4095	0.5236
$r_{y\hat{y}}$	0.7147	0.7173	0.7967

8.3.2. Calibration of the routing parameters

The surface runoff was routed to the outlet of the Bellebeek catchment following the routing model of Troch et al. (1994) (described in Section 6.4). Based on the DEM of the catchment, the normalized width function $W(h)$ and the normalized hillslope function $H(h)$ were constructed. The resulting functions are shown in Figure 8.3. The unit hydrograph was obtained through application of Eq. (6.25) and is shown in Figure 8.3(c). In order to solve the Saint Venant equation (*cfr.*

Eq. (6.28)) in the routing model, four user-defined parameters need to be specified: the flow velocity in the channels V_w [m.s^{-1}], the flow velocity in the hillslopes V_h [m.s^{-1}], the water height above the bed in the rivers H_w [m], the water height above the bed in the hillslopes H_h [m].

The routing parameters define how one volume of water reaches the river in terms of travel time and flow rate. By tuning these parameters, the shape of the simulated peak discharges can be adapted such that it matches the observed peak. As the SCE and DDS method were found to be inadequate for calibrating the TOPMODEL parameters, the parameters from the routing model were only calibrated with the WARPE method. Matrices $\mathbf{Q_k}$, $\mathbf{P_k}$ and $\mathbf{R_k}$ were initialized in a similar manner as in the calibration of the TOPMODEL parameters. The initial values, lower and upper bounds of the routing parameters are listed together with the calibrated values in Table 8.3. For the TOPMODEL parameters, the calibrated values obtained with the WARPE method in the previous calibration step were used. After calibration of the routing parameters, the RMSE further decreased to 0.505.

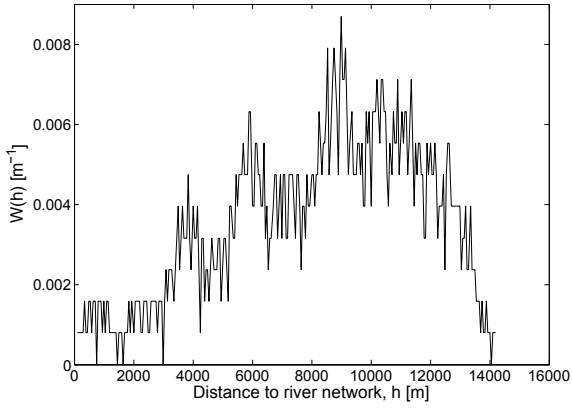
Table 8.3: Initial, minimum, maximum and calibrated value for the routing parameters as obtained with the WARPE method.

Parameter	Minimum	Maximum	Initial	Calibrated
V_w [m.s^{-1}]	0.01	2.00	0.2	0.122
V_h [m.s^{-1}]	0.01	2.00	0.2	0.950
H_w [m]	0.01	2.00	0.2	1.524
H_h [m]	0.01	2.00	0.2	0.894

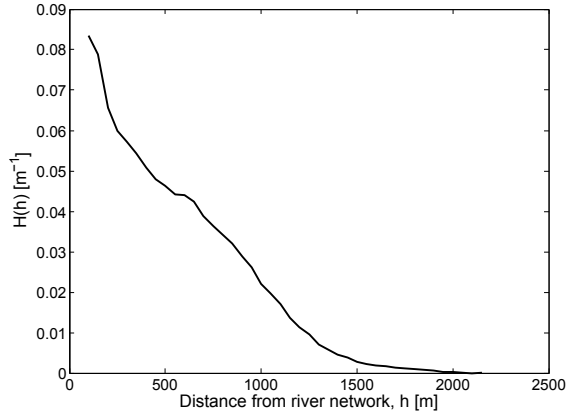
8.3.3. Validation of the calibrated model

For the period January 1, 2007 to December 31, 2007 the hydrological model was validated using the discharge observations at the catchment outlet (*cfr.* Section 7.4). Therefore, TOPLATS was run in distributed mode using the calibrated TOPMODEL and routing parameters (Tables 8.1 and 8.3).

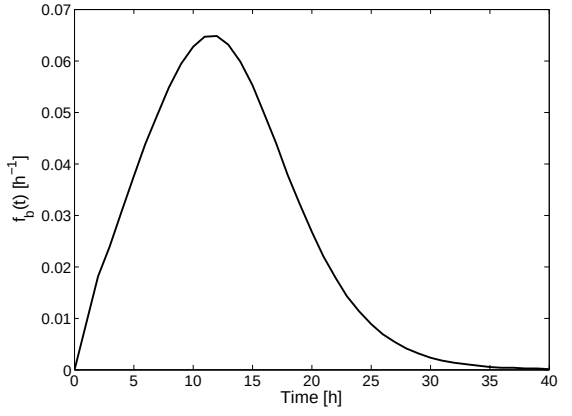
The time series of the validated model run is shown in Figure 8.4 together with the observed discharge in the Bellebeek catchment for the year 2007. Peak discharges are generally good approximated while the baseflow is generally overestimated. The same conclusions can be drawn from the scatterplot (*cfr.* Figure 8.5). The statistics of the validated model run indicate a slight improvement in the fit with an RMSE of 0.5958 and a correlation coefficient of 0.8870. These results show that transferring the calibrated model parameters from one period to another and from aggregated to distributed mode does not hamper the performance of the hydrological model.



(a)



(b)



(c)

Figure 8.3: The runoff routing functions for the Bellebeek catchment: the normalized width function (a), the normalized hillslope function (b) and the resulting unit hydrograph (c).

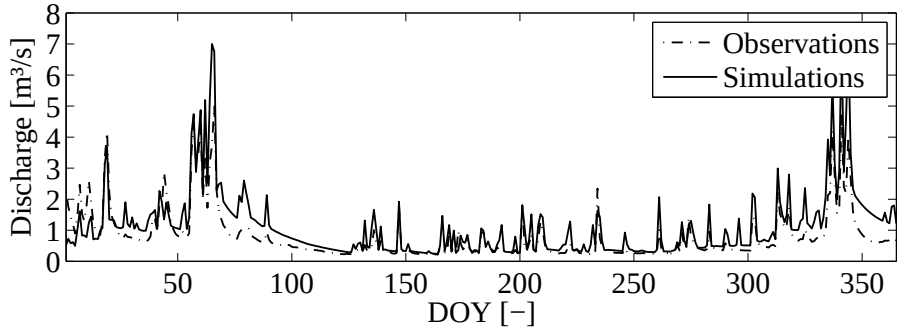


Figure 8.4: Time series of the observed discharge and the simulated discharge with the calibrated TOPMODEL and routing parameter values for the year 2007 (TOPLATS run in distributed mode).

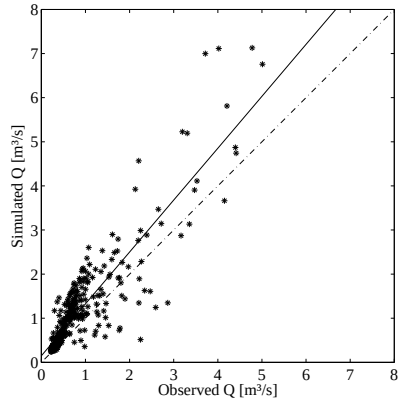


Figure 8.5: Scatterplot between the observed and simulated discharge for the validation period (year 2007) as obtained with the calibrated TOPMODEL and routing parameters from the WARPE method (TOPLATS run in distributed mode).

8.4. Conclusion

It can be concluded that automatic calibration of a distributed hydrological model is not evident and time-consuming. Therefore, we were forced to run the hydrological model in aggregated statistical mode, after which the calibrated model parameters were transferred to distributed mode. Although the calibrated parameter values showed a good model performance in distributed mode, care should be taken when transferring the calibrated parameter set to other model applications as a poor model calibration might result in a lack of sensitivity of the model predictions towards the land and soil surface inputs (Armstrong and Martz, 2008).

The DDS and SCE calibration algorithm both resulted in a minimization of the objective function (in this case the RMSE), however, these methods were not able to produce a sufficiently high correlation between the simulated and observed discharge such that the dynamics in the hydrological processes were not accurately captured. On the contrary, the WARPE method resulted in a simulated discharge that closely resembled the observed discharge. This was expressed by a high correlation coefficient, however, the value of the RMSE indicated that the WARPE method performed worse than the SCE and DDS method. The reason for this result can be found in the theoretical foundation of the calibration algorithms. While the SCE and DDS method minimize an objective function, which reflects the overall difference between simulated and observed discharge, the WARPE method minimizes the uncertainty in the estimated parameter values. Furthermore, with an objective function, differences between simulated and observed discharge are lumped in time which caused the peak discharge to be ill-captured in the calibration result. The WARPE method, on the contrary, takes into account the difference between the simulated and observed discharge for every single time step in the simulation period such that the calibration result clearly reflected the dynamics of the observed discharge. Results of this calibration study suggest that the choice of an automatic calibration algorithm is not arbitrary and that minimization of an objective function is not always effective to optimize discharge predictions. A better alternative would be to make use of multi-objective calibration as it is shown that minimization of the RMSE not necessarily implies a maximalisation of the correlation coefficient. Both statistics include complementary information concerning bias and dynamics, respectively. Further, evaluation of the goodness-of-fit based on statistics should be accompanied by a proper visual inspection of the calibration result.

Part III

Impact of uncertain land cover input on hydrological model predictions

9

Introduction

In this part of the dissertation, different aspects of uncertainty modelling w.r.t. land cover (vegetation) input are investigated. Usually, the land cover in spatially distributed hydrological models is described by a land cover map and a set of parameters that describe the biophysical characteristics for each of the land cover classes. As such, the effect of land cover on the hydrological model response is ultimately controlled by the biophysical parameters. The latter play a crucial role within the model because they determine the energy and moisture exchange between the land surface and the atmosphere, such that changes in the land cover parameters alter the energy budget in the coupled land-atmosphere-transfer scheme and directly affect evapotranspiration, runoff and infiltration. In addition, the land cover map determines the areal proportions of the different land covers in the watershed and the level of detail at which land cover information is present (as given by the land cover labels).

Although most of the vegetation parameters, such as LAI and crop height, can be measured in the field, hydrological model simulations are typically performed under conditions of data scarcity concerning the vegetation characteristics (including predictions in ungauged basins). Land cover information hence needs to be inferred by indirect means, e.g. from expert-knowledge, from values reported in literature or from remotely sensed data. This causes the land cover information to be (highly) uncertain and affects the reliability of the simulation result as the uncertainty propagates through the hydrological model. The uncertainty and sensitivity analyses within this part of the dissertation fit within the scope of water flux predictions in ungauged basins, *i.e.* no field-based measurements of the land cover characteristics are available. Depending on the study area under consideration, different types and levels of information concerning the land cover are available, such that also the uncertainty in the land cover input is highly variable. In this part, the land cover uncertainty is evaluated following the schematic in Figure 9.1. Two different sources of uncertainty are distinguished based on the status of the land cover information:

1. A land cover map of the basin is readily available from different data sources. Depending on the data source, differences in spatial resolution, class labeling, and information content may arise, ranging from detailed information about the crop type, e.g. beets or potatoes, to general information about the vegetation type, e.g. cropland or forest. In Chapter 10, the uncertainty introduced by two different land cover data sets in the prediction of water fluxes is quantified. It is investigated whether and when detailed land cover (crop) information is relevant for hydrological model simulations.

2. A land cover map of the basin is not readily available or does not have the required level of detail. In this case, a land cover map should be derived from remotely sensed data through classification analysis. Both optical and radar-based imagery can be used for this purpose. The latter has the advantage over the former of being less affected by cloud cover, but requires a higher dimensionality (e.g. frequency or temporal) to result in a good classification. In Chapter 11, the uncertainty in the land cover classification is quantified and is evaluated w.r.t. the dimensionality of the data. It is indicated how the acquisition of remote sensing imagery can be optimized to produce land cover maps with a low uncertainty. Chapter 12 discusses on the effect of classification uncertainty on the prediction of water fluxes.

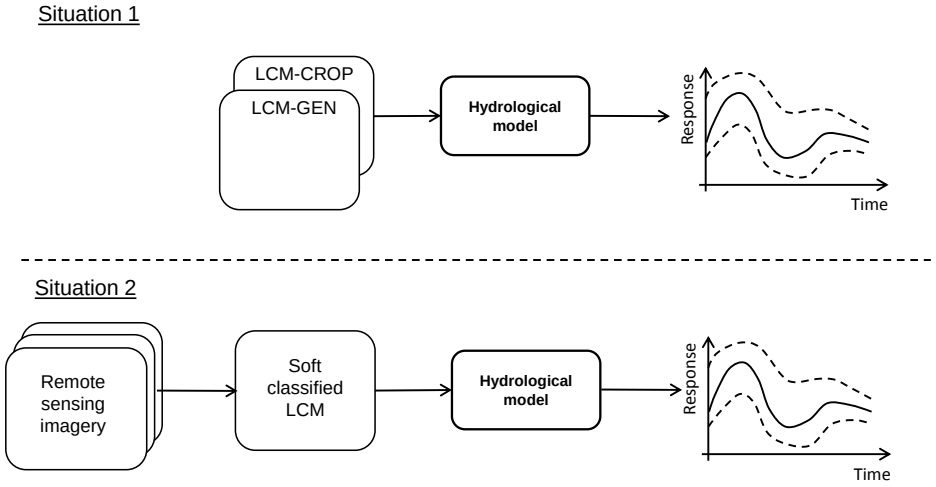


Figure 9.1: A schematic of the different uncertain situations w.r.t. land cover information to model the water fluxes for a catchment. In situation 1, a land cover map of the basin is readily available from different sources: a general land cover map LCM-GEN and a detailed land cover map LCM-CROP, in situation 2 the land cover map needs to be derived through classification of remotely sensed imagery.

10

Effect of land cover generalization on the prediction of water fluxes

10.1. Introduction

The land cover pattern is a critical factor when modelling the spatial and temporal dynamics of the water fluxes in a watershed. However, information about the spatial distribution and the areal proportions of the different land covers in the watershed is usually available from a variety of sources, covering a wide range of scales or levels of detail. Differences in the land cover data sets arise from differences in the method used for land cover survey, and from differences in the goal of the land cover survey. The choice of an appropriate level of detail to represent the land cover pattern, is an important issue in hydrological modelling as the land cover characteristics are directly linked to the generation of runoff. This has been demonstrated through several studies with regard to scaling issues, e.g. the impact of the grid resolution of the land cover map on the hydrological response (Armstrong and Martz, 2008), the transferability of land cover parameters to predict streamflow under other land use scenarios (Heuvelmans et al., 2004). Furthermore, previous studies have shown that historical changes in the land cover affect the hydrological response of the watershed and the evapotranspiration estimates (e.g. Ott and Uhlenbrook, 2004, Matheussen et al., 2000, Singh, 1998, Van Shaar et al., 2002, Wegehenkel, 2002). From these results, it follows that reducing the level of detail or generalizing the land cover information should also affect the water flux predictions, as it leads to a loss of information and a reduced spatial variability (and hence an increased homogeneity) in the land cover pattern. A better understanding of the relationship between the land cover and the hydrological processes would therefore enhance a successfull application of the hydrological model.

A study by Zhang and Schilling (2006) showed that land cover has a significant effect on the simulation of the soil moisture content and the water table depth. Other studies reported only a minor to moderate change in the hydrological response due to land use change (e.g. Armstrong and Martz, 2008, Ott and Uhlenbrook, 2004, Croke et al., 2004, Weber et al., 2001). Wegehenkel et al. (2006) have evaluated the output of a hydrological model (more specifically evapotranspiration and surface

runoff) for four different land cover data sets, each differing in spatial resolution and information content. They found that the impact of uncertainty in the land cover information on the simulated evapotranspiration is only minor. Also the spatial distribution of the simulated evapotranspiration was found to be similar for the different land cover data sets. However, simulated runoff rates were found to show a strong correlation with the areal proportions of the land covers: larger urban areas (which are considered impervious) corresponded with significantly higher amounts of annual runoff. Similar effects were reported by Ott and Uhlenbrook (2004). Further, Armstrong and Martz (2008) have evaluated the output of a hydrological model when several generalized land cover maps (obtained from a classified land cover map by means of aggregation methods) were used as input to a hydrological model, but have found relatively small differences in the predicted flow.

Several studies have drawn attention to the importance of a threshold below (above) which the impact of a land cover change is minor (major). The threshold effect is a function of the way in which transpiration and evaporation separately respond to land cover changes (Li et al., 2007). The exact response of the model to changes in the vegetation is a function of the model itself: how physical processes and parameterizations are represented and how the structure of the model is built (Li et al., 2007). As such, it was investigated by Eckhardt et al. (2003) whether different land cover scenarios could be distinguished in their effect on the model output (surface runoff, streamflow, groundwater recharge). Different land cover maps, with a single land cover type that uniformly covers the watershed, were found to result in significant differences in hydrological response. Nevertheless, when subjecting only a part of the watershed to land cover change, a threshold needs to be threspassed in order to have a significant impact on the model output (the lowest threshold was set at 15%).

From the brief literature survey, it is clear that the effect of variations in the detail of the land cover information, as given by the land cover map, is not straightforward and that the results of several studies are contradicting. It follows that the impact of reducing the land cover detail depends on the type of hydrological model used, on the model output considered and on the extent of the information reduction. The objective of this chapter is to investigate whether detailed information about the crop type is relevant for the prediction of evapotranspiration, runoff and baseflow by the spatially-distributed and physically-based TOPLATS. The model set-up to simulate the water fluxes for the catchment of the Bellebeek is described in Section 10.3. Two land cover data sets of the watershed are readily available and are compared in Section 10.4 w.r.t. (i) their information content and (ii) their corresponding water flux predictions using the techniques for map comparison as described in Section 10.2.

10.2. Techniques for thematic map comparison

Differences between categorical maps can be measured and characterized in a variety of ways through visual or quantitative analysis. Quantitative map comparison is often performed at the pixel level and is based upon measures of classification agreement. In this section, three techniques for thematic map comparison are briefly discussed: (i) cross-walking, which is based on the kappa coefficient of agreement, (ii) average mutual information, which is based on the information theory and (iii) fuzzy agreement, which is based on expert knowledge and fuzzy logic.

10.2.1. Cross-walking

The cross-walking approach is based on a one-to-many relationship between the legends of two maps. In case the maps have a different set of class labels, a conversion of the class labels is required such that there is a correspondence between both sets. The latter can be achieved by collapsing the set of classes directly to obtain comparable classification schemes, while assuring that class labels are consistent in their meaning. The cross-tabulation between the land cover maps is treated as a standard confusion matrix such that the correspondence between the maps can be assessed using standard measures of accuracy (Foody, 2006).

A commonly used measure for map comparison is the kappa coefficient of agreement. This measure is a rescaled version of the proportion of cases for which the labels in the maps agree and is defined as (Cohen, 1960):

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{10.1}$$

with p_e the proportional agreement expected by chance and p_o the proportional observed agreement, respectively calculated as:

$$p_o = \sum_{i=1}^{n_{\text{class}}} \frac{n_{ii}}{n_{\text{tot}}} \tag{10.2}$$

$$p_e = \sum_{i=1}^{n_{\text{class}}} \frac{(n_{.i} n_{i.})}{n_{\text{tot}}} \tag{10.3}$$

where n_{class} is the number of classes, n_{tot} the total number of elements, n_{ii} the number of elements in the diagonal cell ii, $n_{.i}$ and $n_{i.}$ are the totals of column i and row i in the confusion matrix, respectively. If there is a perfect agreement between the observations and the predictions ($p_o = 1$), κ is 1. No agreement ($p_o = p_e$) results in a κ value of 0. In case there is no agreement with an inverse relation between predictions and observations ($p_o < p_e$), κ becomes negative (Landis and

Koch, 1977).

10.2.2. Average mutual information

Foody (2006) has developed a method to compare maps, even though they have different classification schemes, that is based on the degree of mutual or shared information. The amount of information shared by maps X_1 and X_2 is represented by the average mutual information (AMI), which is calculated as (Finn, 1993):

$$AMI = K^A \sum_i \sum_j p(x_2(i), x_1(j)) \cdot \log\left(\frac{p(x_2(i)|x_1(j))}{p(x_2(i))}\right) \tag{10.4}$$

with K^A a constant (usually equal to one), $p(x_2(i))$ the proportion of pixels in class i on map X_2, $p(x_2(i), x_1(j))$ the joint probability of a pixel being class i on map X_2 and class j on map X_1 and $p(x_2(i)|x_1(j)) = p(x_2(i), x_1(j))/p(x_1(j))$ the conditional probability of a pixel belonging to class i on map X_2, given that it belongs to class j on map X_1. The AMI can be expressed as a percentage of the entropy in a reference map:

$$\%AMI = \frac{AMI}{H(X_1)} \cdot 100 \tag{10.5}$$

with $H(X_1)$ the Shannon entropy of the reference map, here X_1. The entropy is calculated as (Shannon, 1948):

$$H(X_1) = -K^S \sum_i p(x_1(i)) \cdot \log(p(x_1(i))) \tag{10.6}$$

where $p(x_1(i))$ is the proportion of class i on map X_1 and K^S is a constant, usually set equal to one. The entropy may be used as an indicator of the confidence of class allocation (Foody, 1995) by expressing the heterogeneity of the membership probabilities. The entropy is maximal when the probability of class membership is evenly partitioned between all classes and is minimal when the probability of class membership is entirely associated with one class. The main disadvantages of the information theory-based approach for map comparison are: (i) there is no spatial component and (ii) it is specifically interesting from a map producer point of view (Foody, 2006). Nevertheless, this approach allows for comparing two maps with the focus on specific classes. To evaluate the amount of information about class i that is shared by two maps X_1 and X_2, the *a posteriori* entropy for map X_1, given the class label information from map X_2, is calculated:

$$H(X_1|x_2(i)) = -K^S \sum_j p(x_1(j)|x_2(i)) \cdot \log(p(x_1(j)|x_2(i))) \tag{10.7}$$

A high *a posteriori* entropy means that the knowledge that a pixel belongs to class

i in map X_2 provides little information on its class membership in map X_1. The change in entropy about the class of a pixel in map X_1 through knowledge of the class in map X_2 is then expressed as a percentage of the total entropy in map X_1:

$$\%AE(X_1|x_2(i)) = \frac{H(X_1) - H(X_1|x_2(i))}{H(X_1)} \cdot 100\% \tag{10.8}$$

where $H(X_1)$ is the entropy in map X_1 calculated following Eq. (10.6) and $H(X_1|x_2(i))$ the *a posteriori* entropy calculated following Eq. (10.7).

10.2.3. Fuzzy agreement

Fritz and See (2005) developed an approach for map comparison that allows for a one-to-many mapping but without cross-walking the legends. This approach therefore preserves the maximum information content in both maps. The methodology consists of two phases: in a first phase, all possible overlaps between the legends are identified, and in a second phase, the uncertainty and the importance of disagreement between the legend categories are mapped based on fuzzy evaluation measures (in order to capture expert knowledge). The two maps are finally compared on a pixel-by-pixel basis to evaluate the spatial agreement. The methodology is applied through following steps (Fritz and See, 2008):

1. Construct a *legend look-up table* to indicate whether there is agreement or similarity between the categories of the legend in map X_1 and map X_2.

2. Create an *overlap matrix* to indicate the amount of overlap (calculated or based on expert knowledge) between the categories for which agreement was identified. The overlap ranges from 0 to 1 and specifies the degree of similarity.

3. Assess the *importance of disagreement*: rate all disagreements on a scale from 1 (least important) to 5 (most important). A value of 0 indicates complete agreement.

4. Calculate the *degree of fuzzy agreement*: convert the ratings from the matrix indicating the importance of disagreement into fuzzy membership values using a simple linearly increasing fuzzy set (Fritz and See, 2005).

5. Create a *thematic (un)certainty layer*: build a cross-classification map between map X_1 and map X_2 and attribute the value from the overlap matrix (constructed in step 2) to each of the categories in the cross-classification map to obtain a certainty layer. The uncertainty layer is then obtained as the complement of the certainty layer.

6. Create an *agreement map*: link the degree of fuzzy agreement (as given by the matrix constructed in step 4) to the categories in the cross-classification

map. The resulting agreement map allows for identifying areas of high spatial disagreement for which further validation and accuracy assessment are advisable.

10.3. Hydrological model set-up

The hydrological model TOPLATS was applied to the catchment of the Bellebeek. Therefore, we forced the model with the climatic data for the period January 1, 2006 to December 31, 2006 (*cfr.* Section 7.3). Information on the soil type and on the topography was provided by respectively the STM-FAO and DEM (*cfr.* Section 7.1). Land cover information was available from two different sources with the class labels as given in Table 7.2:

- LCM-GEN: a general land cover map that provides information on the vegetation type, e.g. cropland and forest.
- LCM-CROP: a detailed land cover map that provides information on the crop type in the agricultural areas, e.g. wheat and barley.

Both land cover maps provide different levels of land cover information as a consequence of the underlying classification scheme. The classification scheme can range from level 1 to level 4, as illustrated in Figure 10.1 (Foody and Atkinson, 2002). A higher level in the classification implies an increase in the information content. LCM-GEN has a classification scheme at level 2 and provides information about the general vegetation type, whereas LCM-CROP has a level 4 classification scheme and provides information about the crop species grown in the agricultural areas.

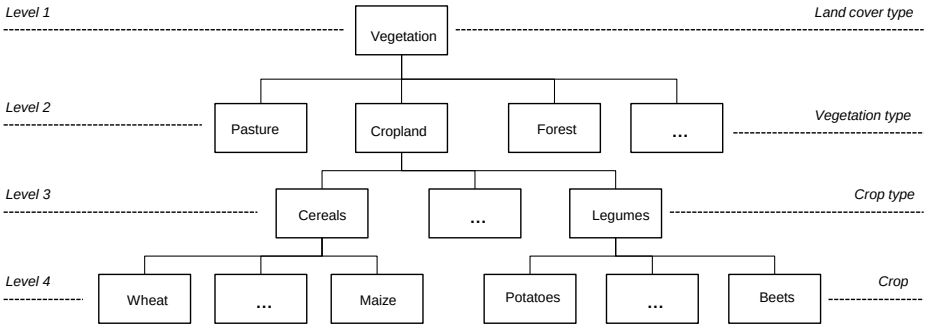


Figure 10.1: Different levels of generalization in a land cover classification scheme (modified from Foody and Atkinson (2002)).

All geographical data of the study area were converted to a grid with a resolution of 50 m. The corresponding soil and vegetation parameters were attributed literature-

based values as described in Section 8.2. Calibrated values for the TOPMODEL and routing parameters were used (*cfr.* Tables 8.1 and 8.3). For a detailed description of the hydrological model, we refer to Chapter 6.

10.4. Results and discussion

10.4.1. Comparison of land cover information

Cross-walking
Land cover maps LCM-GEN (map X_1) and LCM-CROP (map X_2) were cross-referenced, for which the cross-tabulation is given in Table 10.1. It shows that the crop type classes (Gr Oc) mainly correspond to cropland and pasture in LCM-GEN. For a limited number of pixels, LCM-CROP also attributes a crop type class to pixels identified as non-agricultural (e.g. Bf, Ur) in LCM-GEN. In order to quantify this type of disagreement between the different data sources, the class labels Gr, Po, Be, Ma, Wh, Ba and Oc were aggregated into a single class label Cr, such that the labelling in LCM-CROP corresponds to the labelling in LCM-GEN. After cross-walking the legends, the kappa coefficient of agreement (calculated following Eq. (10.1) was found to be 0.93. This value indicates a high agreement between the two land cover maps with respect to their information content about the general vegetation type (level 2). The main source of disagreement at this information level concerns the identification of cropland pixels on LCM-CROP as pasture on LCM-GEN. From the high value of κ, it follows that the difference between both land cover maps is mainly situated at the level of the information content concerning the crop type (level 4).

Average mutual information
Based on the pixel frequencies from the cross-tabulation, the probabilities $p(x_1(j))$ (with $j \in \{1, 2, ..., 6\}$) and $p(x_2(i))$ (with $i \in \{1, 2, ..., 13\}$) were determined. The probabilistic system was then used to calculate the entropies $H(X_1)$ and $H(X_2)$ following Eq. (10.6) in which K^S was put equal to 1. This resulted in an entropy value of 0.42 and 0.86 for LCM-GEN and LCM-CROP, respectively. This means that the heterogeneity in LCM-CROP is much higher than in LCM-GEN, such that the amount of shared information (AMI) by the two land cover maps is only 0.34 (*cfr.* Eq. (10.4)), indicating a poor agreement between both maps. When considering LCM-GEN as reference map, it follows that LCM-CROP shares 39.29% (*cfr.* Eq. (10.5)) of the information contained in LCM-GEN. Or stated otherwise, 60.71% of the land cover information in LCM-CROP is more detailed than the corresponding information in LCM-GEN, which can be inferred from the fact that the poor agreement is mainly attributed to differences in the level of information detail.

Table 10.1: Cross-tabulation between land cover maps LCM-GEN and LCM-CROP. Explanation of the class labels can be found in Table 7.2.

	LCM-GEN						
LCM-CROP	Pa	Cr	Bf	Nf	Wa	Ur	Total
Pa	23962	0	0	0	0	0	23962
Cr	0	85966	0	0	0	0	85966
Bf	0	0	9447	0	0	0	9447
Nf	0	0	0	32	0	0	32
Wa	0	0	0	0	264	0	264
Ur	73	1028	0	0	0	25077	26178
Gr	21305	16823	502	0	3	294	38927
Po	363	5507	2	0	0	15	5887
Be	245	4129	0	0	0	10	4384
Ma	997	21466	38	0	0	61	22562
Wh	448	15528	2	0	0	38	16016
Ba	172	2948	1	0	0	11	3132
Oc	869	4886	12	0	0	169	5936
Total	48434	158281	10004	32	267	25675	242693

The land cover maps were also compared w.r.t. the specific classes by means of the $\%AE$ (*cfr.* Eq. (10.8)). Although the information content in both maps differed in level of detail, the entropy in LCM-CROP consistently decreased through knowledge about the general vegetation type in LCM-GEN (Table 10.2). The entropy in LCM-GEN decreased with 24.31% through integrating information about the cropland area in LCM-GEN. However, a large *a posteriori* entropy was associated with the cropland area, which indicates that knowledge that a pixel belongs to class Cr in LCM-GEN provides little information about the corresponding class label in LCM-CROP. For the other classes, the high value of $\%AE$ indicates that one was correctly informed that the labeling in LCM-CROP corresponds to the labeling in LCM-GEN.

Table 10.2: Change in entropy of LCM-CROP (map X_2) given the class label information in LCM-GEN (map X_1).

$x_1(i)$	$\%AE(X_2\|x_1(i))$
Pa	49.46
Cr	24.31
Bf	87.96
Nf	100.0
Wa	96.88
Ur	92.81

Fuzzy agreement

Following the procedure of Fritz and See (2005) (*cfr.* Section 10.2.3), a legend

Table 10.3: Legend lookup table to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2.

	LCM-GEN					
LCM-CROP	Pa	Cr	Bf	Nf	Wa	Ur
Pa	X					
Cr		X				
Bf			X			
Nf				X		
Wa					X	
Ur						X
Gr	X					
Po		X				
Be		X				
Ma		X				
Wh		X				
Ba		X				
Oc		X				

Table 10.4: Overlap matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2.

	LCM-GEN					
LCM-CROP	Pa	Cr	Bf	Nf	Wa	Ur
Pa	0.5	0.142	0	0	0	0
Cr	0	1	0	0	0	0
Bf	0	0	1	0	0	0
Nf	0	0	0	1	0	0
Wa	0	0	0	0	1	0
Ur	0	0	0	0	0	1
Gr	0.5	0	0	0	0	0
Po	0	0.142	0	0	0	0
Be	0	0.142	0	0	0	0
Ma	0	0.142	0	0	0	0
Wh	0	0.142	0	0	0	0
Ba	0	0.142	0	0	0	0
Oc	0	0.142	0	0	0	0

Table 10.5: Importance of disagreement matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2.

	LCM-GEN					
LCM-CROP	Pa	Cr	Bf	Nf	Wa	Ur
Pa	0	1	4	4	2	3
Cr	3	0	3	2	2	4
Bf	4	5	0	2	3	5
Nf	3	4	2	0	3	5
Wa	2	2	3	3	0	1
Ur	4	3	5	5	1	0
Gr	0	3	4	4	2	4
Po	3	0	4	4	2	3
Be	3	0	4	4	2	4
Ma	3	0	4	4	2	4
Wh	3	0	4	4	2	4
Ba	3	0	4	4	2	4
Oc	3	0	3	3	2	3

Table 10.6: Fuzzy agreement matrix to evaluate the fuzzy agreement between LCM-GEN and LCM-CROP. The class labels are explained in Table 7.2.

	LCM-GEN					
LCM-CROP	Pa	Cr	Bf	Nf	Wa	Ur
Pa	1	1	0.2	0.2	0.6	0.2
Cr	0.4	1	0	0	0.6	0.
Bf	0.2	0	1	0.6	0.4	0
Nf	0.4	0.2	0.6	1	0.4	0
Wa	0.6	0.6	0.4	0.4	1	0.8
Ur	0.2	0.4	0	0	0.8	1
Gr	1	0.4	0.2	0.6	0.6	0.2
Po	0.4	1	0.2	0.2	0.6	0.4
Be	0.4	1	0.4	0.6	0.6	0.4
Ma	0.4	1	0.2	0.2	0.6	0.4
Wh	0.4	1	0.2	0.2	0.6	0.4
Ba	0.4	1	0.2	0.2	0.6	0.4
Oc	0.4	1	0.4	0.4	0.6	0.4

look-up table (Table 10.3, similarity is indicated with a cross) and an overlap matrix (Table 10.4) were constructed to indicate the similarity between LCM-CROP and LCM-GEN. The degree of similarity was calculated as follows:

1. Identical class labels were attributed a similarity degree of one.
2. Class labels that partially overlap, *i.e.* a level 4 class label that is connected with a level 2 class label (*cfr.* Figure 10.1), were given a degree of similarity equal to $1/n_{L4}$ in which n_{L4} is the number of specific classes (level 2 class) belonging to the corresponding general class (level 4 class).
3. If no agreement between the class labels was identified, the degree of similarity was set equal to 0.

For the cases with agreement (situation 1 or 2 from the above list), the importance of disagreement was attributed a value of zero since the level 2 class label and its connected level 4 class labels are non-exclusive. The importance of disagreement between non-overlapping class labels was then rated between 1 and 5 based on expert knowledge (note that this choice is subjective and could have been defined in another way). The resulting importance of disagreement matrix (Table 10.5) was converted into a matrix of fuzzy agreement (Table 10.6) such that the values {1,0.8,0.6,0.4,0.2,0} correspond to a degree of disagreement of {0,1,2,3,4,5} respectively. The values from Tables 10.4 and 10.6 were attributed to the categories in the cross-classification map, which resulted in respectively a map of the thematic uncertainty and a map of the fuzzy agreement between land cover maps LCM-CROP and LCM-GEN, as shown in Figure 10.2. The class labels were obtained after reclassifying the degree of uncertainty and the degree of agreement as follows: 'low' for values between 0 and 0.3, 'medium' for values between 0.3 and 0.5, 'high' for values between 0.5 and 0.7 and 'very high' for values exceeding 0.7. The uncertainty map (*cfr.* Figure 7.4) clearly shows that the uncertainty about the land cover for the pixels in the agricultural areas is very high, which is due to the lack of knowledge about the cultivated crop type. For the non-agricultural areas, the uncertainty about the land cover is low, wich was already suggested by the cross-walking approach. The fuzzy agreement map shows a slightly different pattern as compared to the thematic uncertainty map. The level of agreement between the land cover maps was generally very high with only a limited number of agricultural fields for which a low agreement was detected. For the latter fields, there is a mismatch between the class label from LCM-GEN and the corresponding class label from LCM-CROP, and hence an inconsistency in the level 2 information given by both maps. Tables 10.7 and 10.8 summarize the number of pixels falling into each category of respectively map uncertainty and map agreement. For comparison, these fuzzy measures were also converted to a Boolean measure as follows:

1. Pixels with an uncertainty equal to 0 were assigned to a class 'no uncertainty' while all other pixels (uncertainty>0) were assigned to a class 'full

uncertainty'.

2. Pixels with a fuzzy agreement between 0 and 0.5 were assigned to a class 'no agreement', while pixels with a fuzzy agreement larger than 0.5 were assigned to a class 'full agreement'.

Results show that the fuzzy approach leads to a lower map uncertainty and a higher map agreement than when using a Boolean approach.

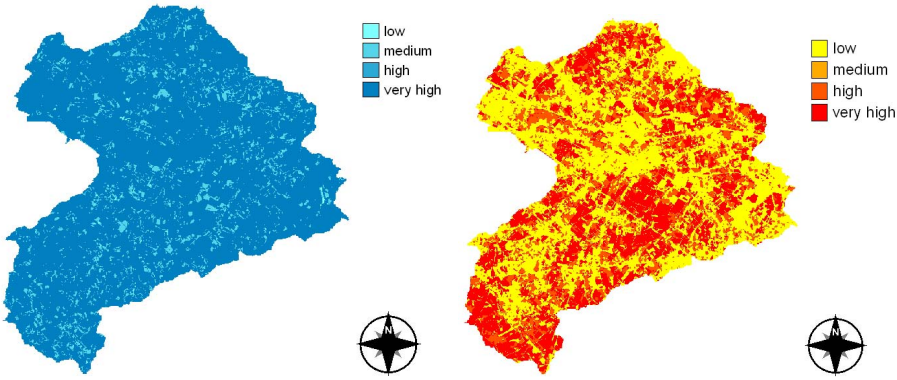


Figure 10.2: Fuzzy agreement map (left) and thematic uncertainty map (right).

Table 10.7: Number of pixels showing no, partial and complete uncertainty together with the average uncertainty in function of the map type.

Map type	No	Low	Medium	High	Very high	Full	Average
Boolean	35 978	-	-	-	-	206 715	0.852
Fuzzy	-	120 525	0	45 267	76 901	-	0.632

Table 10.8: Number of pixels showing no, partial and full agreement together with the average agreement in function of the map type.

Map type	No	Low	Medium	High	Very high	Full	Average
Boolean	22 173	-	-	-	-	220 520	0.907
Fuzzy	-	1032	21 141	3	220 517	-	0.943

10.4.2. Comparison of predicted water fluxes

Following the methodology as described in Section 10.3, the hydrological response of the Bellebeek catchment was simulated for two different scenarios: (i) when the land cover information is given by LCM-GEN, (ii) and when the land cover information is given by LCM-CROP, while keeping all other inputs unchanged. The surface flow Q_s, the baseflow Q_b and the evapotranspiration rate ET were simulated for the year 2006 using an hourly time step and were aggregated on a

daily time basis. A time series of the simulated water fluxes is shown in Figure 10.3. Results show that the impact of changing the detail of land cover information is rather small. Evapotranspiration was generally higher, whereas surface flow and baseflow were consistently lower when map LCM-CROP was replaced by map LCM-GEN. Differences in the simulated water fluxes varied in time and were the highest in spring and summer, which is straightforward as differences in the crop characteristics are most prominent at that time of the year.

The results can be summarized by means of an uncertainty ratio (UR) to reflect the average deviation in the water flux predictions when changing the land cover information, relative to a chosen reference prediction:

$$UR_{\text{map}} = \frac{\left|\sum_{t=1}^{N} y(t)^{\text{LCM-GEN}} - y(t)^{\text{LCM-CROP}}\right|}{\sum_{t=1}^{N} y(t)^{\text{LCM-GEN}}} \tag{10.9}$$

where N is the number of time steps in the simulation (here $N = 365$) and $y(t)^{\text{LCM-GEN}}$ and $y(t)^{\text{LCM-CROP}}$ is the value of model output Y on time step t using respectively LCM-GEN and LCM-CROP as input. The resulting value of UR_{map} is given in Table 10.9 for the different water fluxes and shows that the uncertainty about the crop type introduces the largest uncertainty in the surface runoff predictions, with an average uncertainty ratio of 2.4%. Predictions of evapotranspiration and surface flow were slightly less affected with UR_{map} equal to 2.0%. A spatial analysis of the uncertainty ratio w.r.t. the predicted evapotranspiration rates (Figure 10.4) indicates that the highest uncertainties were detected for the agricultural fields and that the local impact of varying the detail of land cover information can be very high with uncertainty ratios exceeding 50%. This is consistent with the thematic uncertainty map (Figure 10.2), which implies that the uncertainty pattern in the land cover information is reflected in the hydrologic prediction uncertainty.

Table 10.9: Comparison between LCM-GEN and LCM-CROP in terms of water flux predictions and uncertainty ratio.

	LCM-GEN	LCM-CROP	UR_{map}
ET [mm]	392.4	384.5	2.0%
Q_s [mm]	111.4	113.7	2.4%
Q_b [mm]	224.2	229.5	2.0%

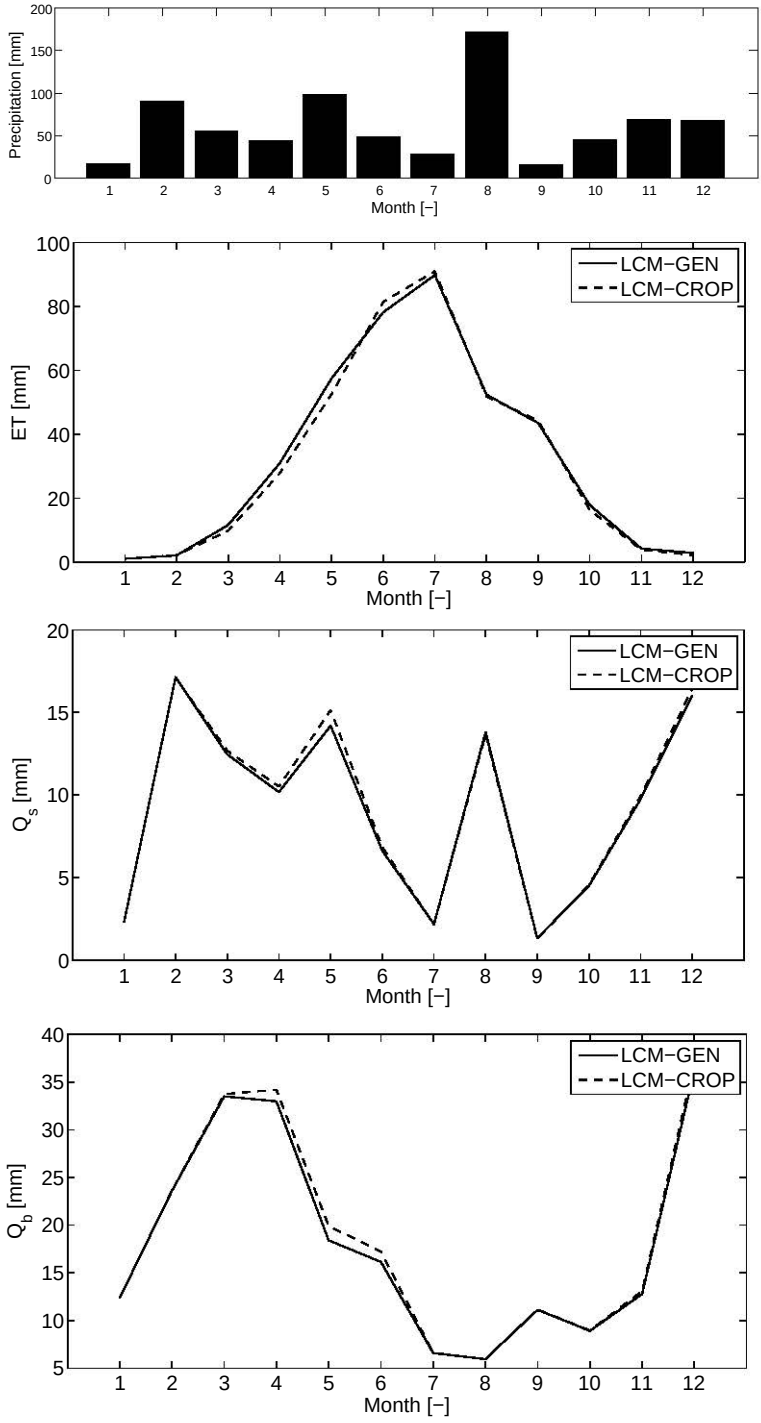


Figure 10.3: Simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b for the Bellebeek catchment during 2006, using two different land cover maps LCM-GEN and LCM-CROP.

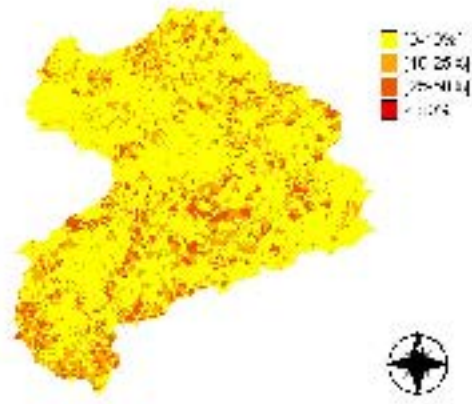


Figure 10.4: Uncertainty map of the simulated evapotranspiration as a result of generalization in the land cover information by replacing land cover map LCM-CROP with LCM-GEN.

10.5. Conclusion

It was investigated whether and how the level of information detail in the land cover map, and more specifically the knowledge about the crop type, affects the prediction of water fluxes by a spatially-distributed hydrological model. Therefore, evapotranspiration, surface flow and baseflow were simulated for the Bellebeek catchment using two different land cover maps: (i) a general land cover map LCM-GEN, providing information about the vegetation type (level 2), and (ii) a detailed land cover map LCM-CROP, providing information about the crop type (level 4).

In a first part of the study, the land cover maps were compared w.r.t. their information content by means of several techniques for map comparison in order to locate and identify the source of land cover uncertainty. The cross-walking approach learned that disagreement between the land cover maps is almost completely attributed to differences in their level of information detail. Almost 61% of the land cover information in LCM-CROP is more detailed than the corresponding information in LCM-GEN. Following the information theory approach, knowledge about which pixels are cropland in LCM-GEN decreased the entropy in LCM-CROP with 24%. The thematic uncertainty map pointed out that when information about the cultivated crop is lacking, a high uncertainty in the land cover map is present. From the results of the map comparison, we can argue that uncertainty in the land cover is best represented and analysed through a fuzzy method. As such, we could identify the level of information detail as the main source of uncertainty and the cropland area as the main location of uncertainty.

In a second phase, the land cover maps were used as input to TOPLATS in order to

simulate the area-averaged hydrological fluxes for the Bellebeek catchment. It was found that varying the level of the land cover information introduced an uncertainty in the water flux predictions that ranged between 2.0 and 2.4%. The impact of land cover information on the water fluxes showed a temporal variation with a peak around spring. Although the land cover maps LCM-GEN and LCM-CROP had little information in common, they did not result in large differences in the area-averaged water flux predictions. The limited effect of adding crop level information can be explained by the fact that the effect of land cover is ultimately controlled by the vegetation parameters from the LC-LUT (*cfr.* Table 6.4). Here, the values for the general cropland class were able to reflect the average parameter values over the different crop classes. As such, adding information about the crop type to the land cover map was not of crucial importance to estimate the area-averaged water fluxes.

A spatial analysis of the simulated evapotranspiration revealed that, in contrast with the area-averaged predictions, local predictions of the annual ET were highly affected by adding information about the crop type. However, these results reflect only a part of the potential effect of uncertain land cover information reported by Wegehenkel et al. (2006), who indicated changes in the runoff prediction up to 70% when using different sources of land cover input. Here, only uncertainty in the land cover class due to differences in information detail is considered, while making abstraction of all other sources of uncertainty. Results show that an increased heterogeneity in the land cover resuls in an increased heterogeneity in the hydrological response and locally reduces the uncertainty in the model prediction up to 50%. This implies that when a spatial distribution of the water flux predictions is intended, it is meaningful to increase the level of land cover information by adding knowledge about the cultivated crop types in order to increase the spatial variability. As such, the best suitable land cover data set can be selected among the available data sources in function of the purpose of the hydrologic simulation. Considering the high correspondence between uncertainty in the model predictions and uncertainty in the land cover map, the thematic uncertainty map was found to be useful to locate the regions for which the water flux predictions will be affected when generalizing the land cover information.

11

Evaluation of uncertainty in land cover from image classification

11.1. Introduction

Thematic mapping, and more specifically land cover classification, is one of the most common applications of remote sensing. Land cover maps obtained from the classification of remote sensing data are frequently used as input to spatially explicit hydrological models. Besides the map itself, users are also interested in the associated map quality. Generally, map quality is assessed at the global level, expressing the thematic accuracy of the hard classification. Many standard methods to estimate the classification accuracy exist (Foody and Atkinson, 2002), but mostly rely on the confusion matrix (Congalton, 1991). This matrix contains information on the pattern of misclassification represented by two types of error: omission (false exclusion) and commission (false inclusion), but provides no information regarding the spatial variability in classification quality within the produced land cover map. However, this type of information is especially interesting if error propagation through the spatially distributed hydrological model is intended as it is widely acknowledged that model inputs can cause considerable errors in the model output (e.g. Pauwels and Wood, 2000, Romanowicz et al., 2005, Miller et al., 2007, Oleson et al., 1997, De Fries and Los, 1999). Furthermore, knowledge on local map quality has shown to be valuable in the process of filtering undesirable pixels from a training set (Goncalves et al., 2009). Local estimates of the classification

This chapter is based on:

1. Loosvelt L., Peters J., Skriver H., Lievens H., Van Coillie F.M.B., De Baets B., Verhoest N.E.C., Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification, *International Journal of Applied Earth Observation and Geoinformation*, 19, 173-184, 2012.

2. Loosvelt L., Peters J., Skriver H., De Baets B., Verhoest N.E.C., Impact of reducing polarimetric SAR input on the uncertainty of crop classifications based on the Random Forests algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 50(10), 4185-4200, 2012.

quality are assessed by means of the classification uncertainty, which is defined as a quantitative measure of doubt when a classification decision is made in a hard way (Unwin, 1995). Consequently, the uncertainty is related to the posterior probability that a pixel belongs to a particular land cover class (Goodchild et al., 1992). Standard measures for uncertainty at the pixel-level are $U = 1 - p_{\max}$, with $p_{\max}$ the maximum probability from the probability vector, and the entropy H of the probability vector (Peters et al., 2011). Although uncertainty measures have shown to provide valuable information about the land cover classification of SAR imagery (Li and Zhang, 2011), aspects of classification uncertainty have been rarely addressed in the remote sensing literature (e.g. McIver and Friedl, 2001, Goncalves et al., 2009, Ibrahim et al., 2005, Brown et al., 2009, Silvan-Cardenas and Wang, 2008).

The posterior probabilities can be assessed by a probabilistic classifier. Hence, in order to account for a spatial representation of the classification uncertainty, an appropriate classification algorithm is required. Conventional statistical methods such as maximum likelihood classifiers (Skriver et al., 2011, Waske and Braun, 2009, Ferro-Famil et al., 2001), the complex-Wishart classifier (Lee et al., 1999, Skriver et al., 2011, Skriver, 2012) and the Hoekman and Vissers classifier (Hoekman and Vissers, 2003, Skriver et al., 2011, Skriver, 2012), have been extensively used for crop classification based on SAR image data. However, these parametric methods rely on the assumption that the probabilities of class memberships can be modelled by a specific probability distribution function, which hampers the performance of the algorithm. Machine learning algorithms, such as neural networks (Frate et al., 2003), support vector machines (Pal, 2005), classification trees (Waske and Braun, 2009) and Random Forests (Breiman, 2001), overcome this shortcoming since they are non-parametric and relax the need for assumptions concerning the statistical frequency distributions. These methods have shown to be able to produce higher accuracies compared to the conventional parametric methods (Alberga et al., 2008, Waske and Braun, 2009, McNairn et al., 2009, Zou et al., 2010). The classification algorithm applied in this study is Random Forests (Breiman, 2001), which is an ensemble of tree-based classifiers and which offers, apart from a probabilistic outcome, several advantages compared to other non-parametric methods: (i) the potential to deal with high-dimensional data sets, (ii) few user-defined model parameters and (iii) the availability of integrated modules to calculate additional measures such as variable interaction, variable importance and proximities.

Although optical remote sensing images are generally used as data sources to perform the land cover classification, synthetic aperture radar (SAR) image data has received considerable research attention (Lee et al., 1999, Stankiewicz, 2006, Skriver et al., 2011), partly by being unaffected by cloud cover, in contrast to optical images (Shi et al., 1994). The distinction between crops in SAR image data is based on their different dielectric properties which affect the received backscatter signal and depends on the plant structure (e.g. size, orientation, shape), the

cropping system (e.g. row density, cover fraction), and the soil characteristics (e.g. soil moisture, soil roughness) (Ulaby et al., 1986). Consequently, each crop can be characterized by a set of polarimetric signatures as was described thoroughly by Skriver et al. (1999), Saich and Borgeaud (2000) and Ferrazzoli (2002). Unfortunately, a single-configuration SAR image is often insufficient for crop discrimination, which imposes the need for multi-configuration data sets. The latter implies the availability of information from multiple dates, multiple frequencies, multiple polarizations, multiple angles, etc. Former classification studies, dealing with the difference between single-frequency and dual-frequency classifications (Skriver et al., 2011, McNairn et al., 2009, Ferro-Famil et al., 2001, Lee et al., 2001), between single-date and multi-date classifications (Stankiewicz, 2006, Frate et al., 2003, McNairn et al., 2009, Tso and Mather, 1999, Blaes et al., 2005, Skriver, 2012) or between single-polarization and full-polarization classifications (Lee et al., 2001, Skriver, 2012), have demonstrated that multi-configuration in SAR images improves the accuracy of the land cover classification. However, the effect of multi-configuration on the local classification uncertainty has not been investigated at present. Due to the high dimensionality of the multi-configuration data sets, some limitations to the efficiency of the classification model may be experienced. It would therefore be computationally beneficial to reduce the number of model features, within a multi-configuration scenario, by only withholding the polarimetric features with the highest impact on the classification result. Furthermore, a reduction of the feature space may enhance the classification performance (Zou et al., 2010) and the assessment of feature importances may assist the interpretability of the classification results, as it provides quantitative information on which features are most suitable to capture the crop characteristics. Methodologies for feature reduction, such as principal component analysis and linear discriminant analysis, have been extensively evaluated for land cover mapping based on multi-spectral and hyper-spectral data (Del Frate et al., 2009, Xu and Gong, 2007). However, studies concerning their application for SAR-based classification are rather limited (Mittal and Singh, 2008, Park and Chi, 2008, 2006, Camps-Valls et al., 2010, Waske et al., 2006).

The objective of this chapter is twofold:

1. The introduction of Random Forests for the probabilistic mapping of vegetation from high-dimensional remote sensing data and presentation of a comprehensive methodology to assess and analyze classification uncertainty based on the local probabilities of class membership. The method is applied to polarimetric SAR image data in order to evaluate the impact of multi-configuration in the data set on the local estimates of classification uncertainty. Data from the Danish airborne SAR (EMISAR) (Madsen et al., 1991, Christensen et al., 1998, Christensen and Dall, 2002), acquired on four different dates in 1998 and in two different frequencies (L- and C-band), are used and are described in Section 11.2. A series of land cover classifications,

in which the focus is on the use of multi-frequency and multi-date SAR data sets, is carried out through Random Forests (discussed in Section 11.3). Based on the soft classification result, the conditions for which multi-configuration in SAR images realizes a decrease in the uncertainty of the land cover map are indicated, variations in uncertainty within the growing season and between crop types are assessed, and a spatial analysis of the classification uncertainty is carried out in Section 11.4. The uncertainty analysis is coupled to an assessment of the classification accuracy at the overall, class- and pixel-level. The pattern of misclassification associated with different types of multi-configuration is identified and is compared to the single-configuration alternative. An accuracy analysis allows for indicating under which conditions the use of a classification model with less data requirements is sufficient.

2. A reduction of the number of polarimetric SAR features, and hence of the dimensionality of the model input for land cover classification. Two different strategies for feature reduction are presented in Section 11.3: (1) an accuracy-oriented reduction, in which the preservation of the model accuracy is the main goal, (2) an efficiency-oriented reduction, in which a small decrease in accuracy is tolerated in favour of less features. Multi-date L-band EMISAR data are used to perform a land cover classification with Random Forests. The methodology for feature reduction is based on the variable importance measure within the Random Forests algorithm. For each input variable, an importance score is generated by permuting the variable of interest and recording the change in classification accuracy. According to the importance ranking and the objectives of the selected reduction strategy, a number of features is omitted from the single-date data sets in order to construct a multi-date input with a strongly reduced dimensionality. On this reduced data set, a multi-date classification model is built using the Random Forests algorithm. The impact of the feature reduction on both model predictions and model uncertainties is investigated. Therefore, a class- and pixel-based comparison between the land cover map obtained with the reduced input and the corresponding land cover map obtained with the non-reduced input, is carried out. The results are discussed in Section 11.5.

Because of the lack of SAR data and because of the limited availability of SPOT data for the Bellebeek catchment (*cfr.* Section 7.2), above analyses could not be performed on the study site for which the hydrological model was applied. Therefore, it is assumed (i) that the same principles for uncertainty assessment and feature reduction as demonstrated for SAR data sets in this chapter can be applied to optical imagery and (ii) that the conclusions drawn from this chapter can be transferred to other SAR and optical data sets.

11.2. Study site and data

The study was conducted at the Foulum test site, which is located near the Research Centre Foulum (RCF) of the Danish Institute of Agricultural Sciences in Jutland, Denmark. The study site is mainly on agricultural area, but some lakes and forested areas are also present. The average elevation of the study site is 54 m.a.s.l., with a variation of 3 m and a maximal slope of 5 °. An aerial photograph of the Foulum test site is shown in Figure 11.1.



Figure 11.1: Aerial photograph of the Foulum test site in Denmark.

11.2.1. SAR imagery

In the period 1993 until 1999, a large number of acquisitions was carried out over the Foulum test site by the Danish Airborne SAR System (EMISAR) (Christensen et al., 1998, Christensen and Dall, 2002), which was flown on a Royal Danish Air Force Gulfstream G-3 Craft at an altitude of approximately 12500 m. The fully polarimetric EMISAR system operates at two different frequencies, L-band (1.25 GHz) and C-band (5.3 GHz) (Christensen et al., 1998). The nominal one-look spatial resolution is 2 m by 2 m. In 1998, simultaneous L- and C-band images were acquired on April 17 (DOY 107), May 20 (DOY 140), June 16 (DOY 167) and July 15 (DOY 196) with four different incidence angles ranging from 20° to 65°. The effect of the different incidence angles on the single-date classification results and the effect of introducing multi-geometry in the multi-date configuration models are out of the scope of this study.

The images were processed and fully calibrated using an advanced internal calibration system (Christensen and Dall, 2002). Corrections of the local incidence angle

due to terrain slope were not carried out since the study site has a very limited relief. The images were then co-registered and orthorectified by identifying ground control points and by using an interferometric DEM (Dierking and Skriver, 2002). Speckle was reduced using a cosine-squared weighted 9 × 9 filter, resulting in a new spatial resolution of 5 m by 5 m with an estimated equivalent number of looks between 9 and 11. The images used in the present study cover an area of about 5 by 5 km.

11.2.2. Ground measurements

In situ measurements at the Foulum test site were carried out in 1998 by the DTU (Technical University of Denmark) and RCF (Research Centre Foulum) on the dates of EMISAR imagery acquisition (DOY 107, 140, 167 and 196). More than 350 fields were visited, and for each of them the crop type, the plant height and the crop development stage were registered. In the present study, a subimage with 36 fields is used. The crop types present in the area include following winter crops: winter barley (Wb), winter wheat (Ww), rye (Ry) and grass (Gr), and following spring crops: beet (Be), pea (Pe), spring barley (Sb) and oats (Oa). The number of reference fields in the used subimage varies between 1 and 8 for each crop (Table 11.1). The observed crop height from April to July is indicated in Table 11.1. The phenological development stage of the crops was described using the BBCH scale (Hack et al., 1992) (Meier, 2001, and references therein), which is a decimal coding system, based on the cereal coding system of Zadoks et al. (1974), where the first and the second digit indicate the principal and secondary growth stage, respectively. Figure 11.2 shows the evolution of the development stage according to the BBCH scale as a function of time. In early spring (DOY 107) a clear distinction is observed between winter and spring crops. Winter crops are in the stage of stem elongation and shoot development (BBCH values between 20 and 30), whereas the spring crops are within the first development stage of germination, sprouting and bud development at that time. The variability in development between the winter and spring crops themselves, however, is very limited in early spring. In late spring (DOY 140), the range of BBCH values covered by the different crops is highest (from 10 to 60), and within crop variations are highest at this time as well. Around mid-summer (DOY 196) nearly all winter and spring crops approach the final stages of their growing season and BBCH values range from about 70 (development of fruit) to 90 (senescence, beginning of dormancy).

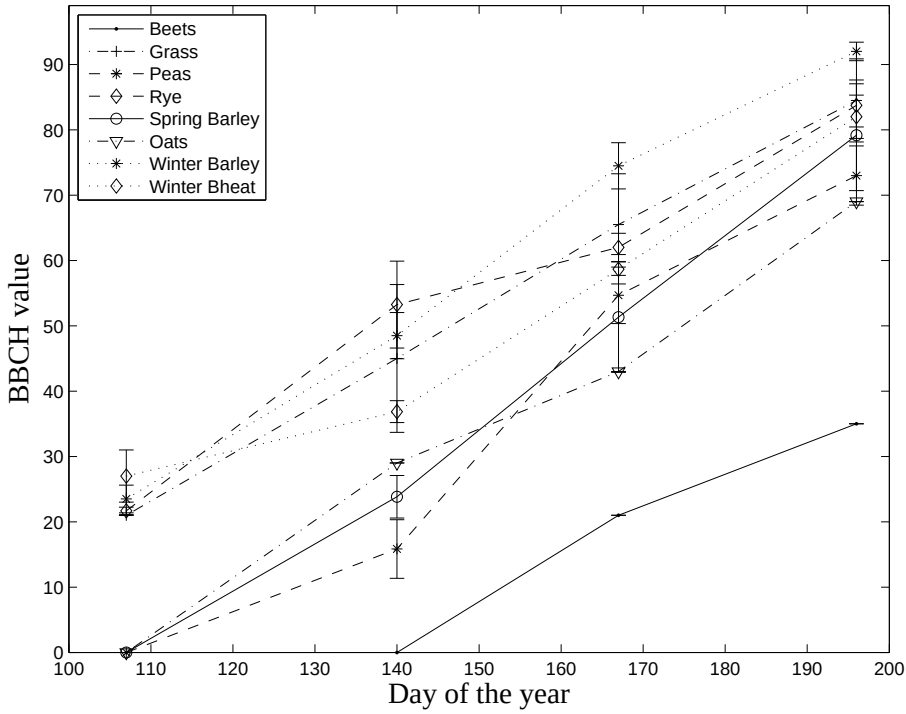


Figure 11.2: Crop development stage according to the BBCH scale for the individual crop types as a function of time. The first digit of the BBCH scale refers to the principal growth stage, the second digit to the secondary growth stage. BBCH values were linearly interpolated between the dates of field survey. The principal growth stages may be summarized as: (0) germination, sprouting, bud development, (1) leaf development, (2) formation of side shoots/tillering, (3) stem elongation or rosette growth, shoot development, (4) development of harvestable vegetative plant parts or vegetatively propagated organs, booting, (5) inflorescence emergence, heading, (6) flowering, (7) development of fruit, (8) ripening or maturity of fruit and seed, and (9) senescence, beginning of dormancy. For detailed information on the scaling system, the reader is referred to Hack et al. (1992) and Meier (2001).

Table 11.1: Overview of the number of fields (n) surveyed per crop type and the average observed crop height (h) per crop type together with the standard deviation.

Crop	n [-]	h [cm]			
		April	May	June	July
Be	1	0	0	0	26
Gr	1	12	26	54	0
Pe	6	0 (±0)	12 (±1)	50 (±13)	61 (±9)
Ry	4	5 (±11)	58 (±6)	100 (±7)	88 (±11)
Sb	7	0 (±0)	12 (±2)	56 (±12)	63 (±3)
Oa	1	0	22	66	99
Wb	2	0 (±0)	58 (±24)	84 (±10)	69 (±3)
Ww	6	5 (±7)	42 (±4)	72 (±5)	77 (±8)

11.3. Methodology

11.3.1. Random Forests

The Random Forests (RF) classifier is an ensemble learning technique that builds multiple decision trees based on random bootstrapped samples of the training data (sampled with replacement) (Breiman, 2001). Consequently, each tree is constructed using a different bootstrap subset X_l from the original training data X, containing about two third of the cases (*cfr.* Figure 11.3). At each node d in the decision tree, m variables are selected at random out of the n_{var} predictive variables and the best split among these m variables is used to split the node. By changing the set of predictive variables and the bootstrap subset over the different trees, Random Forests introduce diversity among the classification trees. Each possible split at node d is evaluated by means of the Gini index, which is a measures for the impurity (heterogeneity) at the node:

$$Gini(d) = \sum_{i=1}^{n_{\text{class}}} p_d(i)(1 - p_d(i)) \tag{11.1}$$

in which $p_d(i)$ is the proportion of class i at node d and n_{class} is the number of classes in the classification. The Gini index is calculated before (at the parent node) and after (at the descendent nodes) the split. The decrease in impurity is called the Gini gain and quantifies how well the classes are being separated based on the splitting rule:

$$Gain(d) = Gini(d) - Gini(d-1) \tag{11.2}$$

The split that produces the maximum Gini gain (*i.e.* lowest impurity) is the most appropriate one such that every time a split of a node is made, the Gini index

for the two descendent nodes is less than for the parent node. Splitting the nodes in each classification tree continues until the largest extent possible (no pruning of the decision tree). By means of a majority vote of the classifier ensemble, the model output is determined. The general procedure to grow a Random Forest is given in Algorithm 1. The cases X_l^c left out of the construction of the l-th tree (usually containing one third of the cases), called out-of-bag elements, are put down the tree to get a test classification (*cfr.* Figure 11.3). The test error is calculated as the proportion of times that the predicted class is not equal to the true class averaged over all cases and is further referred to as the out-of-bag error (oob error) (Algorithm 2). As such, there is no need for cross-validation or for a separate test data set to get an unbiased estimate of the test set error. The classification error depends on: (i) the correlation between any two trees in the forest, and (ii) the strenght of each individual tree in the forest. Decreasing the correlation and increasing the strength lead to a lower error rate. The error rate can be minimized by optimizing the two user-defined parameters in the algorithm, *i.e.* the number of trees (n_{tree}) and the number of variables used to split the nodes (m). The number of trees equals the number of bootstrap subsets used to construct the Random Forest and should be taken large enough in order to allow for convergence of the oob error. Random Forests do not overfit, so there is no upper limit for the number of trees. Increasing (decreasing) the value of m increases (decreases) both the correlation and the strength of the trees, such that the optimal value of m is situated somewhere in between 1 and n_{var}. A Fortran code of the Random Forests algorithm is freely available on http://www.stat.berkeley.edu/∼breiman/RandomForests/, but is also available as a package (`randomForest` (Liaw and Wiener, 2002)) for `R`.

Former studies have demonstrated a good performance of Random Forests as a tool for e.g. land cover mapping (Waske and Braun, 2009, Gislason et al., 2006, Pal, 2005) or ecohydrological distribution modelling (Peters et al., 2007, 2011, 2009). Furthermore, Random Forests is a soft classifier and therefore allows for a quantification of the prediction probability at the pixel level. Each pixel is attributed a probability distribution, with the probability $p(i)$ of being classified into land cover class i, defined as:

$$p(i) = k(i)/n_{\text{tree}} \tag{11.3}$$

where n_{tree} is the total number of trees involved in the classification process (here $n_{\text{tree}} = 200$) and $k(i)$ is the number of trees classifying the land cover as cover type i. Based on the resulting probability distribution, many realizations of the land cover map can be generated and can be further propagated through the spatially distributed hydrological model to assess the impact of classification uncertainty on the model result.

Feature importance

The Random Forests classifier allows for indicating which variables are important

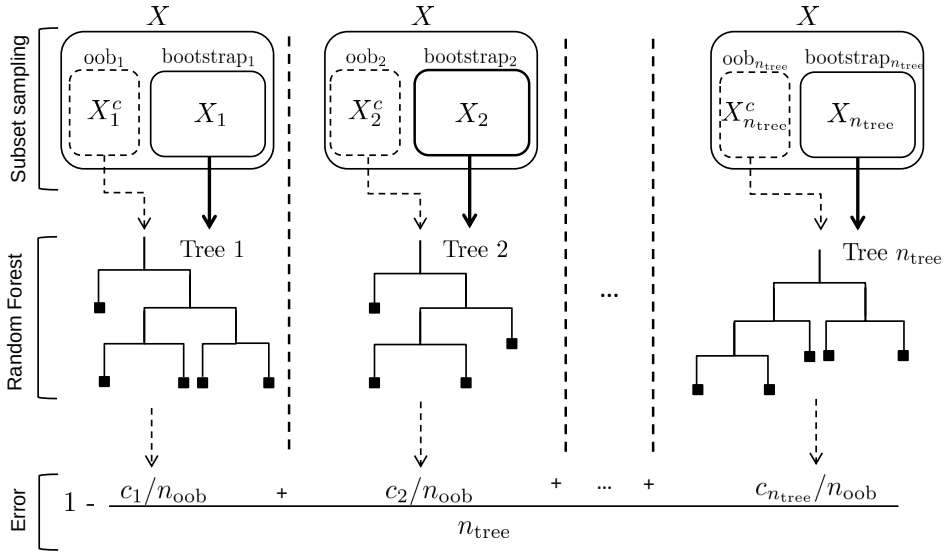


Figure 11.3: Schematic overview of the construction of a Random Forest (full arrows) and of the calculation of the model error (dotted arrows) based on respectively the bootstrap subsets and the out-of-bag subsets sampled from data set X. c_l indicates the number of correctly classified pixels by the l-th tree.

in the classification. Several alternative measures for variable importance are available: (i) the selection frequency: counts the number of times each variable is selected in the tree-building process by all individual trees in the forest, (ii) the Gini importance: describes the mean improvement in the Gini gain splitting criterion produced by each variable over all trees (Friedman, 2001), and (iii) the permutation importance: describes the difference in prediction accuracy before and after permuting the variable of interest, averaged over all trees (Breiman, 2001). The latter is preferred since it takes into account the impact of each predictor variable individually as well as its interaction with the other input variables. The permutation importance of variable j is determined by permuting variable j while keeping all other variables in the out-of-bag (oob) samples unchanged and running all trees on the oob samples. The number of votes for the correct class in the permuted data set is subtracted from the number of votes for the correct class in the untouched data set. The average of this number over all trees in the forest is called the 'mean permutation accuracy importance measure' of variable j, further referred to as C_j. By dividing this raw (unscaled) importance value by its standard error, a scaled version of the permutation importance is obtained and is called the z-score (Breiman, 2001). The procedure to calculate the variable importance is given in Algorithm 3. The permutation importance is implemented in `R` as the function `importance` from the `randomForest` package. Only limited application of the variable importance tool has been reported (e.g. Peters et al., 2008, Gislason

Algorithm 1: Construct a Random Forest on data set X

Data: data set X, containing n_{var} predictor variables and n_{tot} elements
Result: Random Forest classifier

```
begin
    define n_tree, the number of trees;
    define m, the number of variables to split the nodes;
    for each tree l ∈ {1, 2, .., n_tree} do
        draw a bootstrap subset X_l (usually containing 2/3 of n_tot cases);
        for each node d in tree l do
            randomly select m out of the n_var predictor variables;
            for each variable j ∈ {1, 2, ..., m} do
                find the best split u_j;
                calculate the Gini gain for u_j (Eqs. 11.1-11.2);
            end
            select the predictor variable j with the highest Gini gain at u_j to
            define the splitting rule at node d (i.e. the Gini gain splitting
            criterion);
            send cases with x_j < u_j or x_j = u_j to the left descendant, and
            cases with x_j ≥ u_j or x_j ≠ u_j to the right descendant if variable j
            is continuous or categorical, respectively;
        end
    end
end
```

et al., 2006).

Although the Random Forests algorithm is becoming increasingly popular, it has been argued that the permutation importance is less reliable than alternative methods if predictor variables are correlated (Nicodemus and Shugart, 2007), since Random Forests prefer correlated variables in the tree-building process (Strobl et al., 2008). However, this observation has been contradicted by Nicodemus and Malley (2009) and Nicodemus et al. (2010) who observed a slight preference for uncorrelated predictor variables over all splits in the tree-building process. Alternatively, conditional variable importance measures that preserve the correlation structure between variable j and the other variables may be used (Strobl et al., 2008). The conditional measure differs from the unconditional one since the former only permutes variable j within groups of observations (Strobl et al., 2008). Unfortunately, the conditional variable importance calculation is only applicable to small data sets ($n_{\text{tot}} < 500$) due to computational limitations (at present) (Nicodemus et al., 2010) and could therefore not be used in this study.

Moreover, the variable importance measure has shown to be biased in situations where the predictor variables vary in scale of measurement or number of categories (Strobl et al., 2007). Suboptimal predictor variables may therefore be artificially

Algorithm 2: Calculate the out-of-bag (oob) error of a Random Forest

Data: - data set X, containing n_{var} predictor variables and n_{tot} elements;
- Random Forest classifier with n_{tree} trees

Result: oob error

begin
  for *each tree* $l \in \{1, 2, .., n_{\text{tree}}\}$ **do**
    define n_{oob} out-of-bag elements X_l^c (elements not used to construct the l-th tree, usually containing 1/3 of n_{tot} cases);
    classify the oob elements with the l-th tree;
    count the number of correct classifications c_l;
  end
  calculate the oob error as $1 - \frac{1}{n_{\text{tree}}} \sum_{l=1}^{n_{\text{tree}}} \frac{c_l}{n_{\text{oob}}}$;
end

preferred, which induces a biased variable selection when building an individual tree from the forest, and a biased bootstrap sampling with replacement. In order to avoid this biased variable selection, an alternative unbiased Random Forests function has been implemented in `R` as `cforest`. In contrast with `randomForest`, `cforest` does not build the classification trees based on the Gini gain splitting criterion (Breiman et al., 1984), but on a conditional inference framework resulting in unbiased classification trees (Hothorn et al., 2006). In this study, the (unscaled) permutation importance was calculated with the `varimp` and `cforest` functions from the `party` package in `R`. Bootstrap sampling was performed without replacement, since the importance measure was found to be affected by bootstrap sampling with replacement (Strobl et al., 2007).

11.3.2. Model evaluation

The performance of a classification model can be evaluated and compared to alternative models with respect to the model predictions and the underlying probabilities by means of performance measures. Several measures are available to evaluate the overall model performance, its discriminating ability between different land covers and the uncertainty related to model predictions. Some commonly used accuracy and uncertainty measures are described below.

Accuracy measures

The confusion matrix of the hard classification result is used to assess the overall accuracy, OA [%], the class-specific accuracies, CAs [%], and the kappa index of agreement, κ (Cohen, 1960). The overall accuracy is defined as the total proportion of correctly classified pixels, whereas the class-specific accuracy is defined as the proportion of correctly classified pixels per total number of reference pixels within a land cover class. Cohen's kappa coefficient is a statistical measure for categorical

Algorithm 3: Calculate the variable importance with Random Forests

Data: - data set X, containing n_{var} predictor variables and n_{tot} elements;
- Random Forest classifier with n_{tree} trees

Result: importance of each predictor variable

begin
  for *each tree* $l \in \{1, 2, .., n_{\text{tree}}\}$ **do**
    classify the oob elements X_l^c with the l-th tree;
    count the number of correct classifications c_l;
    for *each variable* $j \in \{1, 2, ..., n_{\text{var}}\}$ **do**
      randomly permute the value of variable j in X_l^c;
      classify the permuted oob elements with the l-th tree;
      count the number of correct classifications $c_{l,\tilde{j}}$;
      calculate the raw importance score $\Delta c_{l,j} = (c_l - c_{l,\tilde{j}})/n_{\text{oob}}$;
      replace the permuted oob elements by the untouched oob elements;
    end
  end
  for *each variable* $j \in \{1, 2, ..., n_{\text{var}}\}$ **do**
    calculate the mean importance score over all trees as $C_j = \sum_{l=1}^{n_{\text{tree}}} \Delta c_{l,j}/n_{\text{tree}} \cdot 100$;
    calculate the scaled importance score as z-score$= C_j/\sigma_{\Delta c_{l,j}}$;
  end
end

agreement and is more robust than the simple overall accuracy since it takes into account agreement occurring by chance. The κ coefficient is defined in Eq. (10.1). Perfect agreement results in a κ value of 1, while no agreement is given by a κ value of 0 (Landis and Koch, 1977).

In addition, a set of confusion matrices can be constructed corresponding to different probability thresholds above which the land cover type is modelled to be present. In such a confusion matrix, predictions and observations are compared for each class separately (Table 11.2). By plotting the true positive rate on the ordinate and the false positive rate on the abscis for each probability threshold, receiver operating characteristic (ROC) curves are constructed. The FPR is defined as the fraction of pixels where the land cover class is incorrectly predicted as being present, hence FPR=FP/(FP+TN). The TPR is defined as the fraction of pixels where the land cover class is correctly predicted as being present, hence TPR=TP/(TP+FN). The area under the ROC curve (ranging from 0 to 1), the AUC, describes the likelihood that the observed land cover class for a pixel has a higher modelled probability of occurrence in comparison with pixels of other land cover classes. The AUC is hence a measure for the ability of the model to discriminate between pixels where the land cover type is absent versus present (Fawcett, 2006). Random guessing

results in a straight line between (0,0) and (1,1) in ROC space, leading to an AUC of 0.5. By averaging the AUC over the different classes, an overall measure for the quality of the model predictions is obtained:

$$\mathrm{AUC}_{\mathrm{avg}} = \sum_{i=1}^{n_{\mathrm{class}}} \mathrm{AUC}(i) \cdot w(i) \tag{11.4}$$

with n_{class} the number of classes (here $n_{\mathrm{class}} = 10$), $\mathrm{AUC}(i)$ the area under the ROC curve for land cover class i, and $w(i)$ the weighing factor associated with class i. This weighing factor is determined with respect to the contribution of each class in the test data set. Since each of the 10 classes was equally present in the test data set, all classes were attributed a weight of 0.1.

Table 11.2: Confusion matrix with indication of the true positive (TP), false positive (FP), false negative (FN) and true negative (TN) elements.

	observed	
predicted	present	absent
present	TP	FP
absent	FN	TN

Statistical model comparison

Predictions made by two different models (model 1 and model 2) are explicitly compared by means of a contingency table. The latter is constructed as follows:

number of pixels misclassified by both model 1 and model 2 n_{00}	number of pixels misclassified by model 1 but not by model 2 n_{01}
number of pixels misclassified by model 2 but not by model 1 n_{10}	number of pixels misclassified neither by model 1 nor model 2 n_{11}

where $n_{\mathrm{tot}} = n_{00} + n_{01} + n_{10} + n_{11}$ is the total number of elements in the test data set (here $n_{\mathrm{tot}} = 9920$). Based on this contingency table, the McNemar test (McNemar, 1947) is performed. The McNemar test statistic is a χ^2 goodness-of-fit test (with 1 degree of freedom) that compares the distribution of counts under the null hypothesis ($n_{01} = n_{10}$) to the observed counts (Everitt, 1992), and is calculated as:

$$M = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \tag{11.5}$$

If the null hypothesis of no significant difference in error rates between the two models is correct, then the probability that this quantity is larger than $\chi^2_{1,0.95} = 3.84$ is less than 0.05.

Uncertainty measures

A probabilistic classifier generates a vector of probabilities for each prediction made and is therefore able to characterize the uncertainty in the classification result. The probability vector $\mathbf{p} = (p(1), p(2), ..., p(n_{\text{class}}))$ contains the probability $p(i)$ that a pixel is classified into class i. From this probability vector, uncertainty measures $U = 1 - p_{\max}$ and the Shannon entropy H are derived. For an overview on uncertainty measures for classification, we refer to Unwin (1995).

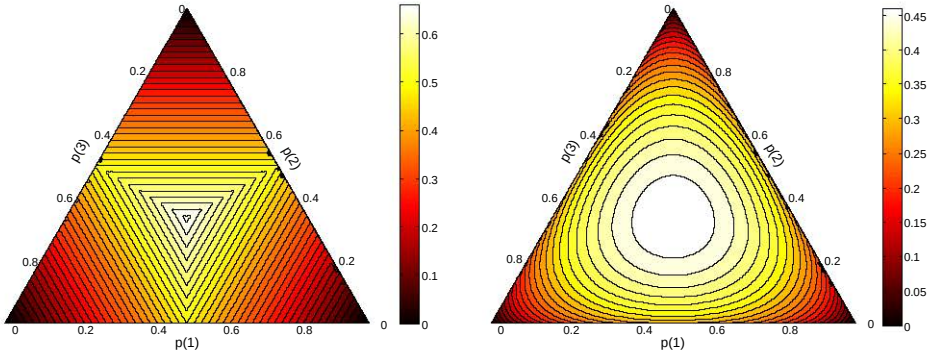
The maximum probability in the probability vector, $p_{\max}$, determines the land cover class that is most likely to occur for a chosen pixel, according to the classification model. This class is mostly used to convert the probabilistic model outcomes into a hard prediction result. The value of $U = 1 - p_{\max}$ is related to the degree of confusion with other classes and is therefore a measure for the strength of the class assignment. If the value of U is high, the classification result is dubious due to e.g. mixed pixels, heterogeneous classes or vague boundaries between classes (Goodchild et al., 1992). A low value of U, on the contrary, indicates that the model has limited doubts about the predicted class. Although the uncertainty measure U allows for a fast and simple uncertainty assessment, it fails to capture the entire distribution of the probabilities within the probability vector, as it only takes into account the maximum probability. By omitting the other probabilities from the evaluation, a lot of valuable information about the classification uncertainty is lost. For example, if $p_{\max} = 0.51$ and $n_{\text{class}} = 10$, the second ranked probability can have a value between 0.054 and 0.49. In case its value is 0.49, there is an apparent confusion between the classes ranked first and second. On the contrary, if the class ranked second has a probability of 0.054, the confidence in the classification result is higher. Weighted uncertainty measures that are able to summarize the information contained in the probability vector can be applied. The Shannon entropy H, originating from information theory (Shannon, 1948), is a commonly used measure:

$$H = -\sum_{i=1}^{n_{\text{class}}} \left(p(i) \cdot \log\left(p(i)\right)\right) \tag{11.6}$$

The entropy of a pixel equals zero if the maximum probability in its probability vector is 1. In this case, the classifier provides strong evidence for a certain land cover class. If the probability is uniformly distributed over the different classes of the probability vector ($p(i) = 1/n_{\text{class}}$), the entropy reaches its maximum value. In this case, none of the classes is preferred and the uncertainty about the pixel's true class is maximal. The classifier is not able to distinguish between the different land cover classes based on the given predictor variables.

The uncertainty measures U and H are both calculated on the modelled probability vector $\mathbf{p}_u$ of each pixel u in the study site. However, both uncertainty measures behave differently because U only uses the maximum probability whereas H exploits the entire probability vector. This difference is illustrated for a three-class example in Figure 11.4, which shows for a given pixel u the value of U and H as a function

of $\mathbf{p}_u = (p(1), p(2), p(3))$. The triangle plot reveals a linear behaviour of U, whereas the entropy shows a nonlinear behaviour.



(a) (b)

Figure 11.4: Triangle plot of uncertainty showing (a) $U = 1 - p_{\max}$ and (b) entropy H of pixel u as a function of probabilities $p(1)$, $p(2)$ and $p(3)$ (three-class example) from the probability vector $\mathbf{p}_u$. The linear behaviour of U contrasts with the nonlinear behaviour of H.

11.3.3. Definition of polarimetric features

A multitude of features may be synthesized from polarimetric SAR data by which information is provided on the soil (e.g. soil moisture), the vegetation biomass, permittivity and structure (e.g. leaf area index, shape of the leaves). For this study, we selected 20 different features based on their usefulness for land cover classification as demonstrated in literature (Stankiewicz, 2006, Alberga, 2007, Alberga et al., 2008, McNairn et al., 2009, Zhang et al., 2009). In the following, the feature calculation is briefly summarized. The basis for the calculation of the features is the covariance matrix. Assuming reciprocity, the covariance matrix is defined as:

$$\mathbf{C} = \begin{pmatrix} \langle S_{\mathrm{hh}} S_{\mathrm{hh}}^* \rangle & \langle S_{\mathrm{hh}} S_{\mathrm{hv}}^* \rangle & \langle S_{\mathrm{hh}} S_{\mathrm{vv}}^* \rangle \\ \langle S_{\mathrm{hv}} S_{\mathrm{hh}}^* \rangle & \langle S_{\mathrm{hv}} S_{\mathrm{hv}}^* \rangle & \langle S_{\mathrm{hv}} S_{\mathrm{vv}}^* \rangle \\ \langle S_{\mathrm{vv}} S_{\mathrm{hh}}^* \rangle & \langle S_{\mathrm{vv}} S_{\mathrm{hv}}^* \rangle & \langle S_{\mathrm{vv}} S_{\mathrm{vv}}^* \rangle \end{pmatrix}$$

where $\langle \cdot \rangle$ represents an ensemble averaging, $.*$ the complex conjugation and S_{tr} the complex scattering amplitude when the transmitted and received signals have a polarization t (horizontal, h, or vertical, v) and r (h or v), respectively. In this study, 20 simple polarimetric features were derived from this covariance matrix. They are divided into five groups.

Horizontal and vertical polarization features
The transmitted and received signals either have a horizontal (h) or a vertical (v) polarization. For the co-polarized (hh and vv) and cross-polarized (hv) responses, the corresponding angle-corrected gamma naught backscatter coefficient (γ^0_{tr}) is computed as:

$$\gamma^0_{tr} = \frac{\sigma^0_{tr}}{\cos\beta} \tag{11.7}$$

where β is the local incidence angle. The gamma naught backscatter coefficient γ^0_{tr} is an alternative to the sigma naught backscatter coefficient σ^0_{tr}, and is slightly less dependent on the incidence angle. The former is therefore expected to be more suitable for classification than the latter if local variations in the incidence angle due to topography are taken into account.

Circular polarization features
Using polarization synthesis, the backscatter coefficient can be calculated for other polarization modes. In case of a circular polarization, the transmitted and received signals can either be right-handed (r) or left-handed (l), resulting in a different response (rr, ll or rl).

45° polarization features
The backscatter coefficient can also be determined for a +45° linear (also indicated as + or +45) or -45° linear (also indicated as - or -45) polarization, corresponding to the angle of the transmitted and received beams. This type of polarization is included for example in Hoekman and Vissers (2003).

Cloude-Pottier decomposition features
A polarimetric decomposition theorem based on the eigenvalue analysis of the coherency matrix was introduced by Cloude and Pottier (1997). From this decomposition, three roll-invariant features can be derived: the entropy H_e, the anisotropy A and the angle $\overline{\alpha}$, which are respectively defined as:

$$H_\mathrm{e} = -\sum_{j=1}^{3} p_j \log_3(p_j) \tag{11.8}$$

$$A = \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \tag{11.9}$$

$$\overline{\alpha} = \sum_{j=1}^{3} p_j \alpha_j \tag{11.10}$$

with $p_j = \frac{\epsilon_j}{\sum_{l=1}^{3} \epsilon_l}$, also called the pseudo-probability of the eigenvalue ϵ_j, which indicates the relative importance of this eigenvalue. The entropy H_e reflects the randomness of a scattering process and ranges from 0 (isotropic scattering) to 1 (totally random scattering). Additional to the entropy, the anisotropy A provides a measure for the relative importance of the different scattering mechanisms. Information on the type of scattering mechanism is contained in $\overline{\alpha}$.

Other polarimetric features

Other commonly used polarimetric SAR observables for classifying SAR images include the correlation coefficient (ρ_{hhvv}) and the phase difference (ϕ_{hhvv}) between the co-polarized linear responses, which are defined as:

$$\rho_{\mathrm{hhvv}} = \left(\frac{\langle S_{\mathrm{hh}} S_{\mathrm{vv}}^{*} \rangle}{\sqrt{\langle S_{\mathrm{hh}} S_{\mathrm{hh}}^{*} \rangle \langle S_{\mathrm{vv}} S_{\mathrm{vv}}^{*} \rangle}} \right) \tag{11.11}$$

$$\phi_{\mathrm{hhvv}} = \psi_{\mathrm{hh}} - \psi_{\mathrm{vv}} \tag{11.12}$$

where ψ_{hh} and ψ_{vv} represent the phase (or argument) of the complex numbers S_{hh} and S_{vv}, respectively. Both features include information on the type of scattering mechanism. A low value of ρ_{hhvv} indicates the dominance of volume scattering, whereas a high value could be caused by surface scattering. For ϕ_{hhvv}, values close to zero indicate single-bounce scattering, whereas values close to $\pm\pi$ could arise from double-bounce scattering. An overview of the polarimetric features and their symbolic notation is given in Table 11.3.

In order to gain more insight in the correlation structure between the different polarimetric features, the Spearman correlation coefficient r (Spearman, 1904) was computed for the different acquisitions. The correlation pattern was found to be similar among the different acquisition dates and frequencies and is illustrated in Figure 11.5 for the features of the L-band acquisition on May 20. The results indicate that the major part of the feature pairs had a rather low correlation, *i.e.* r values lower than 0.5. For example, a low correlation was detected for the feature pairs consisting of γ_{rl} or γ_{vv} on the one hand, and of $\gamma_{\mathrm{hh}}/\gamma_{\mathrm{vv}}$, $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$ and $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{vv}}$ on the other hand. Nevertheless, some feature pairs showed high positive correlations ($r > 0.8$). Examples of such pairs are $\{\gamma_{\mathrm{hh}}, \gamma_{++45}\}$ and $\{\gamma_{\mathrm{ll}}, \gamma_{+-45}\}$. Only a small number of feature pairs showed strong negative correlations ($r < -0.8$). This applies for example to $\{\overline{\alpha}, \gamma_{++45}/\gamma_{+-45}\}$ and $\{\phi_{\mathrm{hhvv}}, H_{\mathrm{e}}\}$. In general, small variations of r between the different acquisition dates were observed, in contrast with large variations of r between the feature pairs. However, the general correlation pattern as shown in Figure 11.5 is only an indication of the overall behaviour between two specific features, disregarding potential class-specific differences. When analyzing the highly correlated features from Table 11.4 at the class-level, it appeared that the high correlation between these features only applied to a limited number of land cover classes. This is illustrated for the feature pair $\{\gamma_{hv}, \gamma_{ll}\}$ in Table 11.4, for which a large variation of r between the crop types and the acquisition dates was observed. For forest pixels, the correlation between γ_{hv} and γ_{ll} was always high ($r > 0.8$) whereas for water pixels it was always low ($r < 0.5$). Other land covers showed moderate to low values of r depending on the time of SAR retrieval.

These findings show that the correlation pattern between the SAR features is very complex with large variations of r between feature pairs, land covers and acquisition dates. Although the general correlation pattern attributed a high

Table 11.3: Overview of the 20 polarimetric SAR features used for land cover classification.

Feature	Symbol	Description
1	$\overline{\alpha}$	mean alpha angle in the Cloude-Pottier decomposition
2	A	anisotropy in the Cloude-Pottier decomposition
3	ρ_{hhvv}	correlation coefficient between the signal in the hh and vv basis
4	H_{e}	entropy in the Cloude-Pottier decomposition
5	γ_{hh}	backscatter coefficient in the hh basis
6	γ_{hv}	backscatter coefficient in the hv basis
7	γ_{ll}	backscatter coefficient in the ll basis
8	ϕ_{hhvv}	phase difference between the signal in hh and vv basis
9	γ_{+-45}	backscatter coefficient in the $+45\,^\circ/-45\,^\circ$ basis
10	γ_{++45}	backscatter coefficient in the $+45\,^\circ/+45\,^\circ$ basis
11	$\gamma_{\mathrm{hh}}/\gamma_{\mathrm{vv}}$	ratio between backscatter coefficient in the hh basis and backscatter coefficient in the vv basis
12	$\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$	ratio between backscatter coefficient in the hv basis and backscatter coefficient in the hh basis
13	$\gamma_{\mathrm{hv}}/\gamma_{\mathrm{vv}}$	ratio between backscatter coefficient in the hv basis and backscatter coefficient in the vv basis
14	$\gamma_{\mathrm{ll}}/\gamma_{\mathrm{rl}}$	ratio between backscatter coefficient in the ll basis and backscatter coefficient in the rl basis
15	$\gamma_{++45}/\gamma_{+-45}$	ratio between backscatter coefficient in the $+45\,^\circ/+45\,^\circ$ basis and backscatter coefficient in the $+45\,^\circ/-45\,^\circ$ basis
16	$\gamma_{\mathrm{rr}}/\gamma_{\mathrm{ll}}$	ratio between backscatter coefficient in the rr basis and backscatter coefficient in the ll basis
17	$\gamma_{\mathrm{rr}}/\gamma_{\mathrm{rl}}$	ratio between backscatter coefficient in the rr basis and backscatter coefficient in the rl basis
18	γ_{rl}	backscatter coefficient in the rl basis
19	γ_{rr}	backscatter coefficient in the rr basis
20	γ_{vv}	backscatter coefficient in the vv basis

correlation coefficient to several feature pairs, the class-level correlation pattern did not reveal a consistently high correlation. Moreover, considering the contradictory conclusions w.r.t. the effect of correlation on the importance score (*cfr.* Section 11.3.1) and the complex correlation pattern among the polarimetric features, it can be argued that if the importance score is affected by the presence of strong correlations, this only applies to a limited number of predictor variables, e.g. $r > 0.8$, in a limited number of cases, e.g. when classifying crop type x with SAR data retrieved on date y.

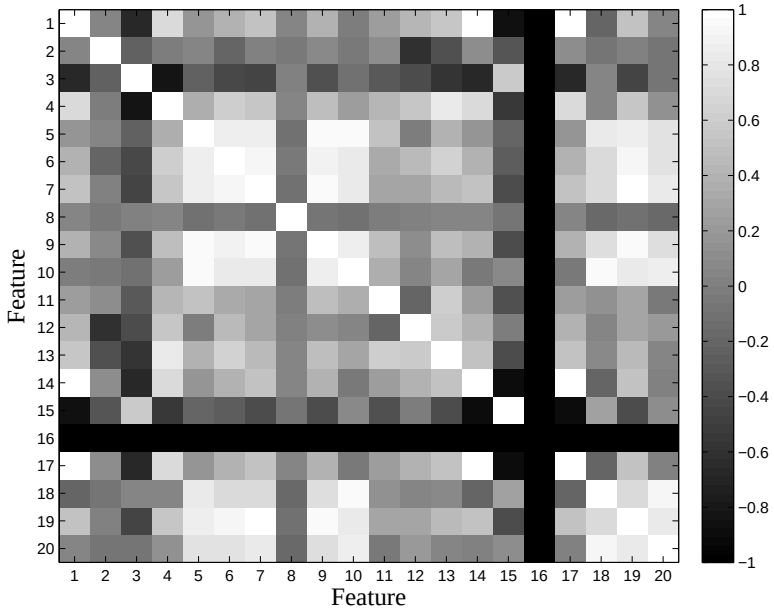


Figure 11.5: Grid representation of the Spearman correlation coefficient between the 20 polarimetric SAR features of the L-band acquisition on May 20. The feature numbers corresponds to the features as listed in Table 11.3

11.3.4. Feature reduction

Depending on the purpose of the land cover map, maintaining the maximal level of accuracy may be the main concern. In such case, the objective of the feature reduction is to reduce the dimensionality of the multi-date data set in order to obtain a classification model that is not significantly worse than the non-reduced model. This reduction strategy is referred to as *accuracy-oriented feature reduction* and aims at identifying a temporal set of features, reflecting the type of features that is most informative according to the date of SAR retrieval. The procedure consists of an evaluation of all features j with respect to their importance score C_j, on the different acquisition dates. According to the importance ranking, the features

Table 11.4: Spearman correlation coefficients r between γ_{hv} and γ_{ll}. Values of r lower than 0.5 are in italic and values of r higher than 0.8 are in boldface.

	April 17	May 20	June 15	July 16
Wa	*0.12*	*0.23*	*0.24*	*0.45*
Fo	**0.86**	**0.85**	**0.85**	**0.86**
Be	*0.47*	0.77	**0.85**	0.59
Pe	**0.86**	0.59	0.64	0.65
Sb	0.73	0.72	*0.46*	**0.81**
Oa	0.57	*0.22*	0.62	0.52
Ry	0.57	*0.45*	*0.27*	*0.49*
Wb	0.62	*0.32*	*0.46*	0.50
Ww	0.74	0.66	0.64	0.62
Gr	0.71	*0.46*	0.77	**0.85**

were cumulatively removed from the corresponding single-date input after which the reduced single-date inputs were concatenated into a multi-date input. The accuracy-oriented feature reduction implies that the selected set of features depends on the date of acquisition, since a feature may not have the same importance rank on all dates. The minimal number of features per date to be included in the multi-date classification was determined based on the following criterion: removing one additional feature from the single-date data sets does not substantially lower the classification accuracy of the derived multi-date model (*i.e.* $> 0.5\%$ with respect to the non-reduced model).

If the accuracy of the land cover map is subordinate to the feature requirements for classification, the objective of the feature reduction can be reformulated as follows: reduce the dimensionality of the multi-date data set at the expense of the classification accuracy in order to obtain a classification model with a performance comparable to the non-reduced model by using a minimal number of features. This reduction strategy is referred to as *efficiency-oriented feature reduction* and aims at identifying an overall set of features that allows for a good land cover classification on multi-date SAR. The advantage of this method over the former is that it provides the most informative type of features irrespective of the date of SAR retrieval. In this reduction strategy, the features with a strong predictive capacity during the entire period (April 17 - July 16) are identified and the same subset of features is withhold on all dates. Therefore, the permutation importance C_j of each feature j is averaged over the different acquisition dates to result in $C_{j,\text{avg}}$. Based on this value, the features are ranked from less important (low value) to more important (high value) and are cumulatively removed from the single-date inputs according to the importance ranking. In the efficiency-oriented feature reduction, only one importance ranking that applies to all acquisition dates, is determined. The number of features to be included in the multi-date model is determined based on a trade-off

between the classification performance and the feature requirements: the benefit of removing an additional feature from the single-date data sets does not justify the loss in classification accuracy of the derived multi-date model (*i.e.* $> 5\%$ with respect to the non-reduced model).

11.4. Effect of multi-configuration on the classification quality

For pixels within the Foulum test site where a field-based observation of the crop type was available, the corresponding values of the 20 polarimetric SAR features (Table 11.3) were extracted, and as such a labeled reference data set was compiled. Both the number and the size of the reference fields were highly variable among the different crop types, causing the number of samples per class to vary between 1984 and 52996. Since large disproportions in the number of samples diminish the classification accuracy of the underrepresented classes, the data set was subjected to a stratified random sampling in order to obtain the same number of pixels per land cover class. Therefore, the data set was divided into $n_{\text{class}} = 10$ strata based on the observed crop type and 1984 samples were drawn idependently in each of the different strata.

Different configurations in the SAR imagery (*cfr.* Section 11.2) were applied: (i) single date, (ii) single frequency, (iii) multi-frequency, (iv) multi-date and (v) multi-date-multi-frequency. For each configuration, the corresponding set of polarimetric features was extracted (*cfr.* Section 11.3.3) and was used to construct a land cover classification model through Random Forests with $n_{\text{tree}} = 200$ and $m = 5$ (default is the square root of n_{var}) (*cfr.* Section 11.3.1). The training of the classifier was performed on 9920 elements from the reference data set. These training pixels were drawn following a stratified random sampling, yielding 992 elements per land cover class. The remaining 9920 elements were used as test data.

11.4.1. Map quality analysis

The modelled land cover maps were evaluated in terms of the quality measures as listed in Section 11.3.2. The quality of the different land cover maps was compared with respect to the classifier input dimensionality. Single-configuration alternatives, by which a classification was performed on single-frequency imagery acquired at a single date, were benchmarked against multi-configuration alternatives. Multi-configuration alternatives included (i) multi-frequency classifications, for which the classifier input consisted of C- and L-band imagery acquired on one date, (ii) multi-date classifications, for which the classifier input consisted of single-frequency imagery acquired on the four distinct dates, and (iii) a multi-date-frequency

classification, for which the classifier input consisted of C- and L-band imagery acquired on the four distinct dates. As such, the dimensionality of the feature space increased from 20 features for the single-configurations, 40 features for the multi-frequency configurations, 80 features for the multi-date configurations and 160 features for the multi-date-frequency configuration.

Accuracy of the classification

Table 11.5 gives an overview of the accuracy for the different configuration classifications. From the single-configuration classifications it was observed that L-band imagery outperformed C-band imagery on all dates for land cover classification. However, a high variability in accuracy was observed since the date of image acquisition did also determine the classifier performance. The highest accuracies were observed when the classification was based on L-band imagery from May or C-band imagery from June, whereas a decreased classification accuracy was observed in July. For the multi-frequency alternatives, the lowest accuracy was obtained from April imagery, accuracies from May, June and July imagery being quite similar. On all dates, the classifier performance is substantially improved by concatenating imagery from both C- and L-band. Multi-date classifications did not differ, and were consistently better than the single-configuration and multi-frequency configuration classifications. Especially the minimum of the user's and producer's accuracies increased considerably for the multi-date classification compared to the multi-frequency classification. Obviously, the multi-date-frequency configuration resulted in the highest values for all evaluation measures with an almost perfect agreement between classified and reference land cover maps.

Although the overall accuracy was often high, the discrepancy between the minimum and the maximum of the user's and producer's accuracies indicated that most single- and multi-frequency configuration classifications failed to attain similar accuracy levels for all land cover classes. For the single-configuration classification based on L-band imagery, especially the erroneous identification of winter wheat and winter barley compromised the classifier performance (*cfr.* Table 11.6). Low class-specific accuracies with C-band imagery were detected for a broader range of crops, including both spring crops (peas, spring barley) and winter crops (rye, winter barley, winter wheat). Additionally, C-band imagery was also found to be inappropriate for classifying forest on June or July data. For most crops, the class-specific accuracies showed high variations among the four distinct dates. Grass, beets and oats were the only crops that were accurately classified irrespective of the frequency or time of SAR retrieval. The multi-frequency configuration based on April imagery showed a lower accuracy boundary of 0.57 and 0.66 for user's and producer's accuracy, respectively. This model substantially increased the accuracy for peas, winter barley and winter wheat in comparison to the single-configuration model on data from April. On the other dates, all class-specific accuracies of the multi-frequency configuration models exceeded 0.75 (not shown in Table 11.6). Based on the accuracy results reported in Tables 11.5 - 11.6, specific

Table 11.5: Overall classification accuracy (OA), model user's accuracy (MA), producer's accuracy (PA) and Cohen's κ for single configuration and multi-configuration classifications. The MA and PA are summarized by their range over the different land cover classes. The last column shows the results from a pairwise classification comparison by the McNemar test at the 0.05 significance level. Different numbers indicate significantly different performances, and models are ranked from the best performing classification model (1) to the worst performing model (11).

Config.	Date	Band	OA	MA	PA	κ	Rank
Single	April 17	C	0.61	[0.34, 0.99]	[0.37, 0.99]	0.56	11
Single	April 17	L	0.65	[0.30, 0.99]	[0.38, 0.99]	0.61	9
Single	May 20	C	0.61	[0.26, 0.99]	[0.37, 1]	0.57	11
Single	May 20	L	0.88	[0.71, 0.98]	[0.75, 0.99]	0.87	4
Single	June 16	C	0.71	[0.42, 0.99]	[0.52, 1]	0.68	8
Single	June 16	L	0.77	[0.43, 0.99]	[0.55, 1]	0.75	6
Single	July 15	C	0.64	[0.39, 0.99]	[0.45, 1]	0.60	10
Single	July 15	L	0.73	[0.55, 0.99]	[0.53, 1]	0.70	7
Multi-freq	April 17	C and L	0.79	[0.57, 1]	[0.66, 0.99]	0.77	5
Multi-freq	May 20	C and L	0.92	[0.80, 0.99]	[0.81, 1]	0.91	3
Multi-freq	June 16	C and L	0.92	[0.80, 0.99]	[0.81, 1]	0.91	3
Multi-freq	July 15	C and L	0.87	[0.77, 1]	[0.75, 1]	0.86	4
Multi-date	all dates	C	0.97	[0.94, 1]	[0.92, 1]	0.96	2
Multi-date	all dates	L	0.97	[0.93, 1]	[0.94, 1]	0.97	2
Multi-date & multi-freq	all dates	C and L	0.99	[0.98, 1]	[0.97, 1]	0.99	1

data requirements that meet the objectives of the classification study can be defined. For the classification of certain crop types, it might be sufficient to use a single-date data set instead of a multi-configuration data set.

Table 11.6: Class-specific user's accuracies (CA) resulted from the single-configuration classifications. The CA is indicated with -, + or ++ if its value exceeded 0.25, 0.50 or 0.75, respectively. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Be), peas (Pe), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww).

Band	Date	Wa	Fo	Gr	Be	Pe	Sb	Oa	Ry	Wb	Wb
C	April 17	++	++	++	+	-	+	+	-	-	-
C	May 20	++	++	+	++	+	-	+	-	+	-
C	June 16	++	-	+	+	+	+	++	++	++	-
C	July 15	++	-	++	++	+	+	++	-	-	+
L	April 17	++	++	++	+	-	+	++	+	-	-
L	May 20	++	++	++	++	++	++	++	++	+	+
L	June 16	++	++	++	+	++	++	++	+	+	-
L	July 15	++	++	++	++	+	+	++	+	+	+

Uncertainty of the classification

All classification models computed a probability vector for each test pixel within the study site. The class corresponding to the maximum probability is the class assigned to the test pixel. In this section, we quantified the uncertainty of the classifier at pixel-level by calculating $U = 1 - p_{\max}$ and the entropy H (*cfr.* Eq. (11.6)) based on the probability vector of each test pixel. For both the correctly and incorrectly classified pixels, an empirical frequency distribution of U and H was built (Figure 11.6). The shape of both distributions is indicative for the prediction strength of the classification model. Frequency distributions of U show that a high proportion of correctly classified pixels obtained low values of U (< 0.2), whereas the majority of incorrectly classified pixels obtained a value of U in the range 0.3 to 0.6. Furthermore, different models showed different cumulative distribution functions for U, with decreasing values of U starting from the single-configuration models, followed by the multi-frequency models, the multi-date models and finally the multi-date-frequency models. This trend was much less pronounced for the incorrectly classified pixels, as all models obtained similar values of U for similar proportions of pixels.

For entropy, both the correctly and incorrectly classified pixels showed a frequency distribution similar to that of U. The majority of pixels that were incorrectly classified obtained an entropy value between 0.4 and 0.7, with little variability between the models. For the correctly classified pixels, the frequency distribution of H showed a high proportion of pixels with low entropy (< 0.2). The entropy of the correctly classified pixels consistently decreased from single-configuration over multi-frequency, multi-date and multi-date-frequency models. Within the different configuration schemes, classification models using C-band imagery generally

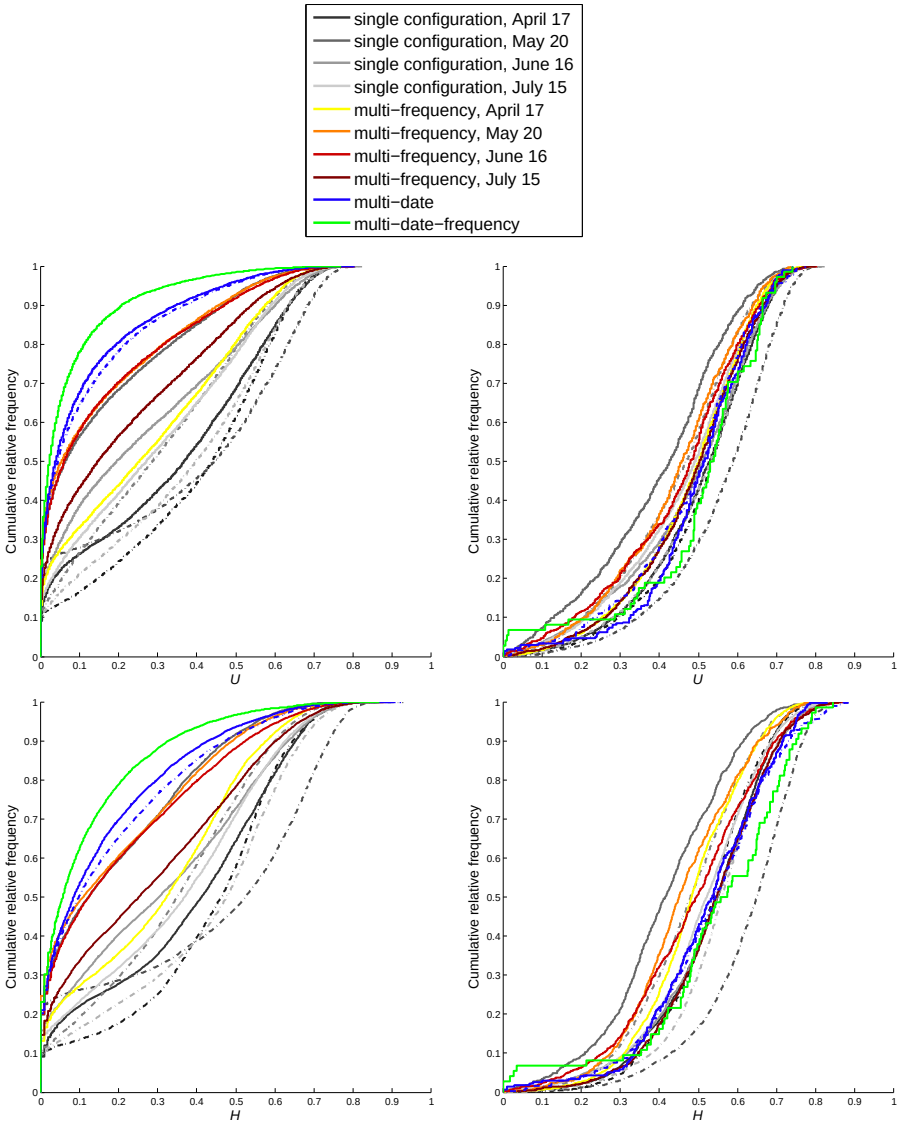


Figure 11.6: Cumulative relative frequency distribution of $U = 1 - p_{\max}$ and entropy H for correctly classified and incorrectly classified pixels by the different classification models. Results for classification models applying C-band imagery are given in dashed lines, results for models applying L-band imagery are given in solid lines.

performed worse compared to the L-band models.

A class-specific uncertainty assessment by means of U and H showed marked differences between the land cover classes (Tables 11.7 - 11.8). All the models showed an extremely low classification uncertainty for water pixels, with values of U and H close or equal to 0. A similarly low classification uncertainty was observed for forest pixels. The C-band single-configuration models, however, did perform worse for this land cover class and showed an increasing uncertainty when the growing season proceeds. The classification uncertainty for grass was also very low for all multi-configuration models and most of the single-configuration models. Results on the class-specific classification uncertainty for grass showed an increased variability among the different single-configuration models in C-band.

The multi-date or multi-date-frequency models generally obtained very low classification uncertainty ($U < 0.14$; $H < 0.24$) for all agricultural crops. On the contrary, the uncertainty of the single-configuration classification models showed a lot of variability among the four dates and seven crop types. When comparing the classification uncertainty between dates for each crop individually, both C- and L-band models obtained lower values in May or June in contrast to April and July, although crop-specific uncertainty from C-band classifications was most frequently higher than from L-band classifications. The highest uncertainty was generally recorded for rye, winter wheat or winter barley in case L-band models were applied. C-band models additionally produced high uncertainty values for peas and spring barley. The multi-frequency models improved the classification uncertainty, especially of winter crops. The lowest uncertainty was recorded for beets or peas, depending on the time of SAR retrieval. Classification uncertainty remained the highest for the winter crops (rye in the July model, winter wheat in the May and June model), but also for peas in the April model. These uncertainty trends were observed from both the analysis of $U = 1 - p_{\max}$ and H. However, for some models, uncertainty measures U and H identified a different crop type as the class with the least uncertain classification result (Tables 11.7 - 11.8).

In summary, it can be said that in addition to higher classification accuracy levels, multi-configuration also decreased the uncertainty of the classification as compared to the single-configuration alternatives (Figure 11.6). However, improved uncertainty levels were only observed for correctly classified pixels. As the dimensionality of the configuration increased, a larger portion of correct predictions were made with lower uncertainty, *i.e.* lower values of U and H. This means that if the model voted for the correct class, there was little confusion with other land classes. On the contrary, incorrectly classified pixels mostly had moderate uncertainties, independent of the classifier input dimensionality. This means that if the model voted for an incorrect class, the prediction was uncertain due to confusion between - most likely - similar land cover classes with similar probabilities of occurrence. Although it would have been expected that multi-configuration also decreased the

Table 11.7: Median values of $U = 1 - p_{\max}$ for the single-configuration and the multi-configuration classifications for each of the land cover classes separately. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Be), peas (Pe), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww). For each classification model the minimum value of U among the agricultural crops is in boldface, the maximum value of U is underlined.

			Land cover class									
Configuration	Date	Band	Wa	Fo	Gr	Be	Pe	Sb	Oa	Ry	Wb	Ww
single	April 17	C	0	0.08	0.22	**0.30**	0.47	0.42	0.36	0.55	0.56	0.53
single	May 20	C	0	0.03	0.27	**0**	0.52	0.56	0.53	0.57	0.36	0.45
single	June 16	C	0	0.39	0.33	0.24	0.29	0.23	0.10	0.15	**0.09**	0.42
single	July 15	C	0	0.52	0.19	**0.09**	0.47	0.40	0.19	0.47	0.50	0.29
single	April 17	L	0.04	0	0.03	0.36	0.52	0.40	**0.21**	0.40	0.55	0.44
single	May 20	L	0	0	0.02	**0.01**	0.04	0.23	0.02	0.10	0.24	0.23
single	June 16	L	0	0	0.09	0.13	0.17	**0.12**	0.20	0.43	0.29	0.41
single	July 15	L	0	0	0.21	**0.15**	0.24	0.45	0.22	0.39	0.26	0.22
multi-freq	April 17	C and L	0	0	0.03	**0.20**	0.42	0.30	0.22	0.36	0.39	0.37
multi-freq	May 20	C and L	0	0	0.04	**0**	0.04	0.24	0.04	0.11	0.22	0.27
multi-freq	June 16	C and L	0	0	0.08	0.11	**0.02**	0.08	0.05	0.15	0.03	0.23
multi-freq	July 15	C and L	0	0.01	0.09	**0.02**	0.18	0.25	0.12	0.31	0.26	0.18
multi-date	all dates	C	0	0.02	0.12	**0**	0.04	0.07	0.02	0.14	0.05	0.13
multi-date	all dates	L	0	0	0.02	**0.02**	0.05	0.10	**0.02**	0.10	0.12	0.11
multi-date-freq	all dates	C and L	0	0	0.04	**0**	0.02	0.06	0.02	0.08	0.04	0.08

Table 11.8: Median entropy values for the single-configuration and the multi-configuration classifications for each of the land cover classes separately. The land cover classes include water (Wa), forest (Fo), grass (Gr) and the agricultural crops beets (Br), peas (Pr), spring barley (Sb), oats (Oa), Rye (Ry), winter barley (Wb) and winter wheat (Ww). For each classification model the minimum entropy value among the agricultural crops is in boldface, the maximum entropy value is underlined (excluding classes Wa, Fo and Gr).

			Land cover class									
Configuration	Date	Band	Wa	Fo	Gr	Be	Pe	Sb	Oa	Ry	Wb	Ww
single	April 17	C	0	0.15	0.35	0.41	0.50	0.40	**0.37**	0.53	0.55	0.55
single	May 20	C	0	0.01	0.40	**0**	0.61	0.64	0.62	0.65	0.38	0.42
single	June 16	C	0	0.36	0.42	0.36	0.33	0.30	0.18	0.24	**0.16**	0.45
single	July 15	C	0	0.56	0.29	**0.15**	0.53	0.47	0.30	0.55	0.56	0.39
single	April 17	L	0.10	0	0.07	0.44	0.58	0.39	**0.25**	0.49	0.57	0.50
single	May 20	L	0	0	0.05	**0.02**	0.09	0.32	0.06	0.18	0.31	0.29
single	June 16	L	0	0	0.16	0.23	0.28	**0.21**	0.32	0.47	0.40	0.44
single	July 15	L	0	0.01	0.34	**0.24**	0.33	0.51	0.33	0.47	0.37	0.31
multi-freq	April 17	C and L	0	0.01	0.07	0.30	0.46	0.34	**0.26**	0.42	0.44	0.45
multi-freq	May 20	C and L	0	0	0.08	**0**	0.08	0.34	0.08	0.19	0.30	0.34
multi-freq	June 16	C and L	0	0.01	0.16	0.22	**0.06**	0.16	0.11	0.23	**0.06**	0.34
multi-freq	July 15	C and L	0	0.03	0.17	**0.05**	0.29	0.37	0.22	0.44	0.39	0.27
multi-date	all dates	C	0	0.04	0.24	**0.01**	0.09	0.15	0.04	0.24	0.10	0.22
multi-date	all dates	L	0	0	0.05	**0.05**	0.10	0.18	**0.05**	0.18	0.22	0.20
multi-date-freq	all dates	C and L	0	0	0.09	**0.01**	0.05	0.12	0.04	0.17	0.09	0.16

uncertainty in the incorrect predictions, *i.e.* predict them with higher values of U and H, this was not the case.

11.4.2. Effect of image frequency

As the crop-specific uncertainty analysis (*cfr.* Tables 11.7-11.8) showed a substantial agreement with the results from the accuracy analysis, the classification quality is used as an overall measure to discuss the effect of SAR frequency: a high (low) quality corresponds to a high (low) classification accuracy in combination with a low (high) classification uncertainty.

An evaluation of the single-configuration models indicated a higher classification quality resulting from L-band backscatter information as compared to C-band (*cfr.* Table 11.5). This offers positive perspectives towards the upcoming missions of ALOS-II, equipped with a quad-polarized L-band sensor. The largest deficiency of C-band models, as compared to L-band models, is the low quality by which forest is classified in mid-summer (*cfr.* Table 11.6), which is due to the high biomass of other crops and the restricted penetration depth of C-band. Former studies also reported on more accurate land cover classification based on L-band imagery. Soares et al. (1997) found a superior L-band contribution for crop discrimination based on C- and L-band SAR image data, and Simard et al. (2002) observed a higher accuracy of an L-band classifier (66%) in comparison to a C-band classifier (61%) for mapping tropical coastal vegetation.

Some land cover classes such as water, forest, grass and beets were characterised by a high classification quality, irrespective of the frequency of the input data. The quality for the remaining classes was low in case of the single-frequency classifications. By combining both frequency bands in multi-frequency models, a broader characterization of the land cover types could be attained. As a consequence, the classification quality substantially increased, but its magnitude depended on the time of SAR retrieval. When imagery from late spring or later were used, class-level quality was high and similar between all land cover classes. Multi-frequency models produced the lowest classification quality for rye, winter wheat and winter barley, which is attributed to a combination of similarities in plant physiology and similarities in crop development stage. On average, the accuracy levels were within the range reported by Chen et al. (1996).

11.4.3. Effect of image acquisition date

The determinative effect of the time of data retrieval within the growing period on classification performance is widely acknowledged (Lin et al., 2009, Wang et al., 2010). Likewise, the results of this study indicate the importance of acquisition date on the classification quality (*cfr.* Tables 11.5-11.6). The potential of the

single-configuration models to classify spring crops increased from April until June, and decreased towards July. In case early-crop inventory (in April) is required, it is recommended to use L-band imagery, rather than C-band imagery. The effect of time of data retrieval on the classification quality can be explained referring to the crop development stages (*cfr.* Figure 11.2). In April, all the spring crops (oats, spring barley and peas) were in a similar stage of development (early germination), and SAR polarimetric signatures of these spring crops were very alike. The winter crops (rye, winter barley and winter wheat) were in a tillering, shoot development or stem elongation stage at that time, and were readily distinguished by the model. As the growing season proceeded, interspecific differences in growth, development stage, crop morphology, and cropping systems became more apparent, resulting in more distinct backscatter signals for the different crops in late spring. Imagery obtained during late spring included backscatter information of crops within six or seven development stages. The models applied to imagery obtained during this period therefore showed the highest classification quality (Table 11.5). The models developed on imagery from July resulted in a lower quality, since only four different development stages were recorded at that time: rosette growth (exclusively for beets), development of fruit, ripening or maturity of fruit and seed, and senescence.

Through the use of polarimetric features computed from SAR images of different dates, multi-date configuration models significantly improved the classification quality compared to the single-date configuration alternatives (*cfr.* Table 11.5). Moreover, the difference in performance between C- and L-band models was no longer present in the multi-date configuration. For all classes, high classification qualities were recorded. According to the acquisition timing in the growing season, polarimetric SAR features from the early spring image have resulted in subdividing winter from spring crops. The separation of the winter and spring crops themselves is based on the distinct growth stages and the resulting distinct backscatter signatures of the crops. These results emphasize the importance of acquisition timing in the growing season, and support the conclusions of several authors (Wegmüller and Werner, 1997, Stankiewicz, 2006, Wang et al., 2010, Skriver, 2012) for adjustment of imagery acquisition to the crop development. As compared to the multi-frequency models, multi-date models resulted in a higher and more similar classification quality for all land covers. This is an important conclusion in prospect of future missions such as SENTINEL-1 (not polarimetric, C-band) and the RADARSAT constellation (polarimetric, C-band), aiming at a frequent revisit time.

11.4.4. Spatial analysis of uncertainty

To spatially locate the strenghts and weaknesses of the classification model, a map of the associated classification uncertainty (represented by the entropy) was constructed. Figure 11.7 shows the result for the multi-date L-band model. Dif-

ferent patterns in the spatial distribution of the classification uncertainty were detected. The *boundary field pattern* attributes high uncertainty values to boundary pixels, indicated by bright borders (*cfr.* Figure 11.7(a), e.g. area 1). Comparison of classification uncertainty of boundary pixels with inner-field pixels from the reference fields revealed, on average, a slightly higher uncertainty for the former ones (*cfr.* Figure 11.7(b)). However, the boundary effect in the reference fields is partly lost due to digitization of the field borders, which caused a large number of boundary pixels to be omitted from the field. This resulted in an underestimation of the average boundary pixel uncertainty. Secondly, *regular inner-field patterns* were observed as parallel lines of high uncertainty (*cfr.* Figure 11.7(a), e.g. area 2). This pattern in classification uncertainty was assessed by the correlation of the grey-level co-occurrence matrices (GLCM) (Haralick et al., 1973) for a range of offset distances. Some reference fields did show a particular correlation pattern (Figure 11.7(c)) with repeating high correlations over a five pixel interval. These patterns were observed mainly for winter wheat fields. Finally, *irregular inner-field patterns* and *random patterns* occurred as respectively clustered (*cfr.* Figure 11.7(a), e.g. area 3), or randomly scattered pixels with high uncertainty. The observed pattern in uncertainty is also valid for the other types of model configuration. However, depending on the configuration type, some patterns may be less or more pronounced. The multi-date C-band model showed an expansion of the boundary field pattern, whereas the single-date models were dominated by random patterns of high entropy.

With exception of the random type, the uncertainty patterns are explained by means of observable spatial phenomena. For the pixels in the boundary of a field, the received signal contained mixed characteristics from two or more adjacent land cover classes and caused confusion in the classification process. Because of this confusion, the model attributed similar probabilities to these classes, which resulted in a high entropy. Moreover, selecting the most probable class as the predicted class was often incorrect for these boundary pixels. Entropy values at the inner-field were generally lower. This provided strong evidence to predict one particular land cover class with a low associated uncertainty at the pixel level. The regular inner-field patterns reflect agricultural practices such as ploughing and sowing in parallel rows. These practices induced repeated gradients in vegetation cover and probably also in soil characteristics, e.g. compaction. The bare soil zone between two vegetated rows was subjected to the mixed-pixel effect, similar to the boundary zone, and therefore suffered from high uncertainty. Irregular inner-field patterns were very likely to be caused by local differences in soil conditions, e.g. soil roughness, soil texture and soil topography, which caused differences in soil moisture (e.g. ponding water). The soil moisture gradient in its turn induced differences in vegetation density and crop development within the same crop type and caused the model to erroneously classify these pixels. Therefore, heterogeneity in soil conditions tends to lower the strength of the classification model, which is clearly reflected in locally

increased entropy values.

11.5. Effect of feature reduction on the classification quality

On each single-date L-band data set (April 17, May 20, June 16 or July 15), the model input features were determined (*cfr.* Section 11.3.3) and were used to carry out a land cover classification with Random Forests ($n_{\text{tree}} = 200$, $m = 5$, *cfr.* Section 11.3.1). Initially, all 20 polarimetric features (Table 11.3) were included in the model development. In a next step, the number of input features was reduced following two different strategies as described in Section 11.3.4: the accuracy-oriented and efficiency-oriented feature reduction. On the reduced inputs, a series of Random Forests classifications was carried out. In a second step, the impact of reducing the number of input features, and hence the model dimensionality, on the predicted land cover and its associated quality was analyzed based on the accuracy and uncertainty measures as described in Section 11.3.2.

11.5.1. Feature importance analysis

On all single-date L-band data sets, a Random Forests classification was carried out and a variable importance analysis was performed with 10 repetitions. Figure 11.8 gives the resulting importance score for each of the 20 input features. The importance pattern was found to depend on the time of acquisition. This is not unexpected since the polarimetric characteristics of the vegetated area change along with the crop development stage. Consequently, the informative contribution of the different polarimetric features changed when other scattering mechanisms started to dominate (*cfr.* Section 11.3.3). From Figure 11.8, some general trends in the importance pattern were detected. Irrespective of the acquisition date, γ_{hv} clearly occurred as the most important predictor for land cover. The ranking of all other features was found to vary from early spring (April 17) to mid-summer (July 16). The most apparent variation in importance was observed for $\overline{\alpha}$, γ_{++45}, $\gamma_{\text{hv}}/\gamma_{\text{hh}}$ and $\gamma_{++45}/\gamma_{+-45}$. On April 17 and May 20, $\overline{\alpha}$ was identified as a very important feature, whereas on June 15 and July 16 its importance was considerably lower. A similar phenomenon was observed for γ_{++45}, which was found to have a clearly higher importance on May 20 and July 16 than on April 17 and June 15. Furthermore, $\gamma_{++45}/\gamma_{+-45}$ and $\gamma_{\text{hv}}/\gamma_{\text{hh}}$ were important predictors on respectively May 20 and June 15, whereas on the other dates the opposite was true. The analysis of the predictive capacity (indicated by its importance value C_j) of a feature as a function of time, revealed some general information on its potential to detect differences in crop phenology. Concerning $\overline{\alpha}$ and $\gamma_{++45}/\gamma_{+-45}$, these

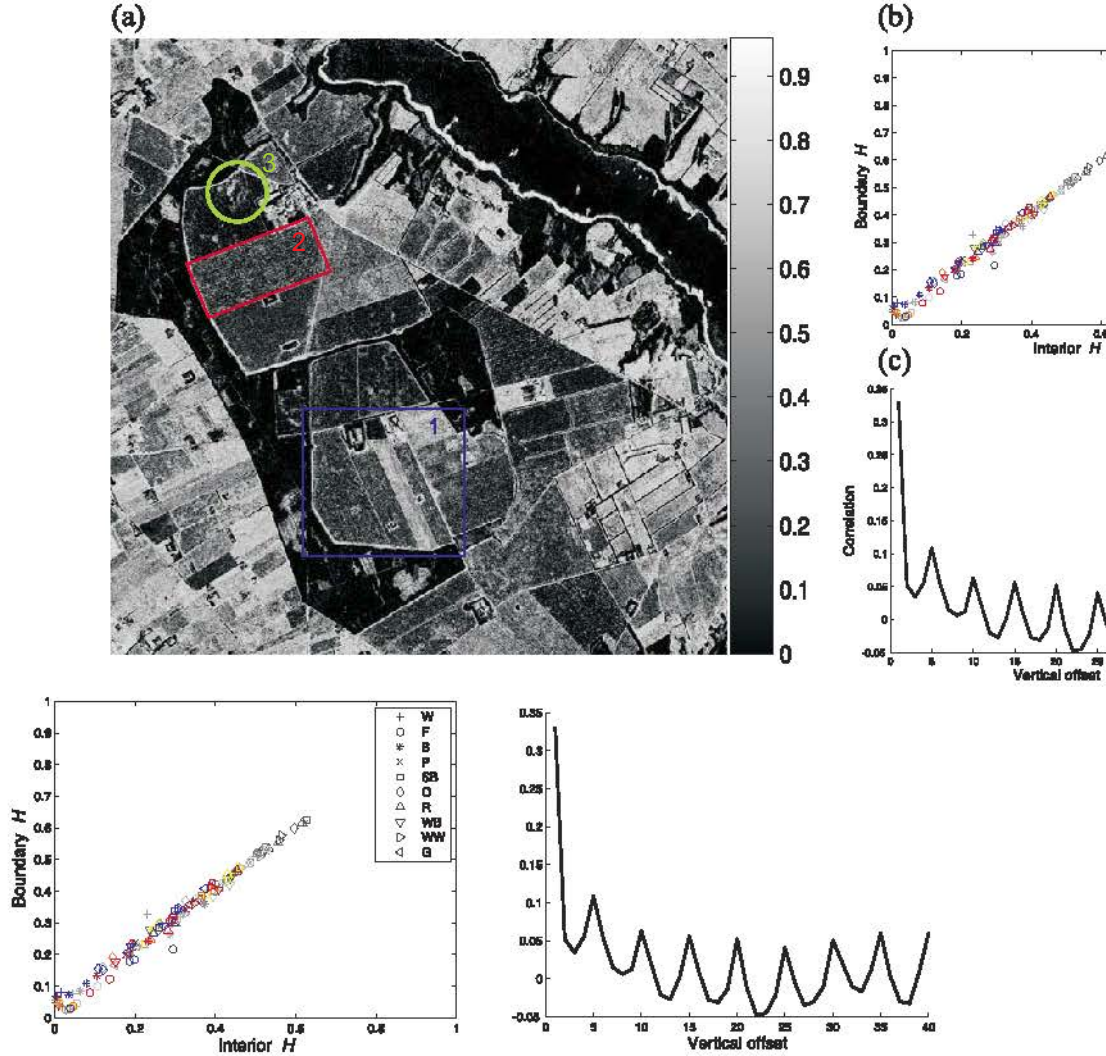


Figure 11.7: (a) Uncertainty map of the multi-date classification model in L-band using entropy (H). (b) Scatterplot of entropy values for inner-field pixels and boundary pixels for the 15 different classification models averaged over the different reference fields for each crop separately. Different crops have different scatter symbols, whereas different models have different colours (for the colour legend, *cfr.* Figure 11.6). (c) The within-field variability of uncertainty values reflects the agricultural practices for some crops, e.g. winter wheat.

features seem to be useful for crop identification in spring, when winter and spring crops can be clearly distinguished based on their crop development stage (Figure 11.2). At that time, there is a high variation in crop phenology: spring crops have just started leaf and shoot development, whereas the winter crops have already passed this stage and begin to grow their vegetative parts. On the other hand, $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$ and γ_{++45} appeared as important features in summer, when nearly all crops reach the final stages of crop development. Phenological differences are then situated at the level of the flowering stage. For cereal crops, phenological differences are determined by the position of the heads (hanging or upright).

11.5.2. Feature reduction

Accuracy-oriented reduction

For each single-date data set, the features were first ranked with respect to their importance score and the feature with the lowest importance was removed from each single-date data set independently. The removed features were therefore not necessarily the same on all dates (Figure 11.8). Secondly, a multi-date data set was built by concatenating the four reduced single-date data sets, and a multi-date classification model was constructed. This procedure was repeated until only 1 feature per acquisition date, *i.e.* the most important one, was left to construct the multi-date land cover model. For each classification, the overall classification accuracy OA [-], Cohen's κ [-] and the $\mathrm{AUC}_{\mathrm{avg}}$ [-] were calculated (*cfr.* Section 11.3.2). The accuracy of the classification model, in terms of OA and κ, was subjected to marginal changes if less than 13 features per date were removed from the model input (Figure 11.9). A gradual decrease in OA and κ was observed when more than 12 features per date were omitted. The $\mathrm{AUC}_{\mathrm{avg}}$ remained high and unaffected as long as no more than 16 features were removed from the input.

The optimal number of features was selected on the basis of the most restrictive accuracy measure (here κ), following the criterion as mentioned in Section 11.3.4. It was found that an equally performing, but less complex multi-date model could be constructed if only the eight most important features per date were used in the classification process. An overview of the selected features per date (Table 11.9) shows that the multi-date model as resulted from the accuracy-oriented feature reduction required the calculation of 14 different features, as not all 8 selected features were the same on each date. From this subset of features, γ_{hv}, ϕ_{hhvv}, γ_{ll} and γ_{rr} were selected as important predictors on all or on the majority of dates. These features were designated as 'general' features for crop classification because their potential to detect differences in crop characteristics was hardly affected by the time of acquisition within the growing season. The high discriminating potential of the cross-polarization mode can be attributed to its specific sensitivity towards random orientation of the canopy scattering elements. Former research already

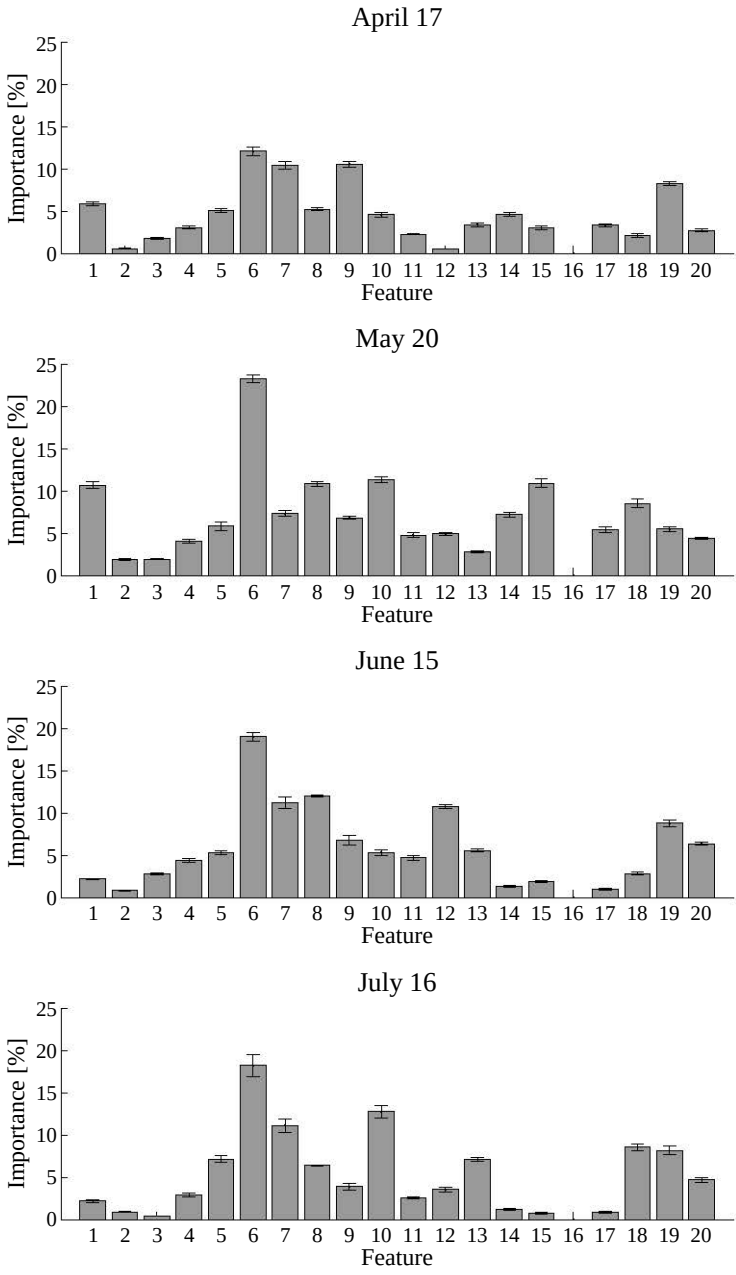


Figure 11.8: Importance score (calculated following Algorithm 3) of the 20 features derived from L-band single-date SAR imagery (acquisitions on April 17, May 20, June 15 and July 16 in 1998) used to classify the land cover at the Foulum test site (Denmark), averaged over 10 model runs (with standard deviation). The numbers on the abscis correspond to specific features from Table 11.3.

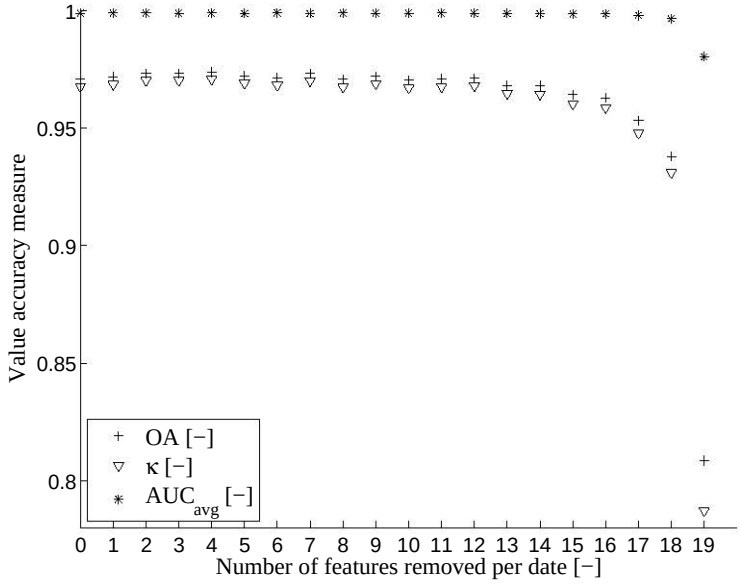


Figure 11.9: Overall accuracy, Cohen's kappa and $\mathrm{AUC_{avg}}$ as a function of the number of features removed from the single-date data sets. Feature removal is based on the feature importance ranking per date, in which the removed feature is not necessarily the same on all dates.

pinpointed that cross-polarization is a crucial polarization for crop discrimination in both C-band and L-band SAR (McNairn et al., 2009, Lee et al., 2001). The other features from the subset were only selected once (e.g. γ_{vv} on June 15) or twice (e.g. $\overline{\alpha}$ on April 17 and May 20) and were therefore designated as 'specific' features for crop classification. For the latter type of features, the importance score and hence the predictive capacity strongly depends on the time of acquisition. The selected set of most important features showed a mixture of highly and lowly correlated features. However, strong correlations were not the motive behind the feature selection. For example, γ_{hv}, the most important feature, showed a strong general correlation ($r > 0.8$) with γ_{++45} and γ_{+-45} (Figure 11.5), but there was no evidence that on all dates these latter features were automatically selected as being important too. This supports the earlier assumption that the presence of high correlations does not necessarily bias the feature selection.

To gain insight into the relative importance of the different dates, a variable importance analysis was performed on the selected multi-date model. The resulting importance for each input feature is shown in Figure 11.10 with indication of the corresponding date from which the feature was derived. The γ_{hv} from the May 20 acquisition was almost twice as important as the second important feature. Based on the average importance $\overline{C}$ of the eight selected features per date (*cfr.* Table 11.9), the different acquisition dates were ranked relative to their contribution to

Table 11.9: Overview of the eight SAR features with the highest permutation importance C_j (ranked from 1 to 8 with decreasing C_j) on each of the four acquisition dates.

Rank	April 17	May 20	June 15	July 16
1	γ_{hv}	γ_{hv}	γ_{hv}	γ_{hv}
2	γ_{+-45}	γ_{++45}	ϕ_{hhvv}	γ_{++45}
3	γ_{ll}	$\gamma_{++45}/\gamma_{+-45}$	γ_{ll}	γ_{ll}
4	γ_{rr}	ϕ_{hhvv}	$\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$	γ_{rl}
5	α	α	γ_{rr}	γ_{rr}
6	ϕ_{hhvv}	γ_{rl}	γ_{+-45}	$\gamma_{\mathrm{hv}}/\gamma_{\mathrm{vv}}$
7	γ_{hh}	γ_{ll}	γ_{vv}	γ_{hh}
8	$\gamma_{\mathrm{ll}}/\gamma_{\mathrm{rl}}$	$\gamma_{\mathrm{ll}}/\gamma_{\mathrm{rl}}$	$\gamma_{\mathrm{hv}}/\gamma_{\mathrm{vv}}$	ϕ_{hhvv}

the multi-date classification. The highest ranked date was May 20 ($\overline{C} = 7.20$) while the lowest in rank was June 15 ($\overline{C} = 3.39$). April 17 ($\overline{C} = 3.54$) and July 16 ($\overline{C} = 4.42$) had a slightly higher contribution than June 15. The higher importance of the features from the May 20 acquisition reflects the high observed discrepancy in crop development at that time (*cfr.* Table 11.2). This observation is in contrast with former research that indicated July as the month with the highest potential for accurate classification based on the overall accuracy, both for C-band (Frate et al., 2003) and L-band SAR (McNairn et al., 2009). When moving from a single-date (Figure 11.8) to a multi-date model (Figure 11.10), a modification in the ranking of the features occurred. For example, $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{vv}}$ was found to be the lowest ranked feature from June 15 in the single-date model, whereas this feature appeared to be the fourth ranked feature from June 15 in the multi-date model. Nevertheless, γ_{hv} remained the most influential predictor in the multi-date classification for each of the contributing dates.

Efficiency-oriented reduction

The importance score of each feature j was averaged over the different acquisition dates to result in a mean importance value, $C_{j,\mathrm{avg}}$ (Figure 11.11). Large standard deviations on $C_{j,\mathrm{avg}}$ were observed for the 'specific' features, e.g. $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$ and $\gamma_{++45}/\gamma_{+-45}$. Based on the value of $C_{j,\mathrm{avg}}$, the features were ranked from less (low value) to more important (high value). In iterative model applications, the feature with the lowest $C_{j,\mathrm{avg}}$ was removed from the single-date data sets and the resulting multi-date data set was used for model calibration and evaluation. If less than 10 features were removed, the accuracy of the derived multi-date model remained very high and almost unaffected (*cfr.* Figure 11.12). Removing more than 10 features caused the model performance to gradually deteriorate, followed by a dramatic decrease in accuracy when more than 17 features were removed. In view of the objective of efficiency-oriented reduction (*cfr.* Section 11.3.4), it was chosen to limit the input to three different features per aquisition date, *i.e.* γ_{hv}, ϕ_{hhvv} and γ_{ll}. The subset of features determined by the efficiency-oriented

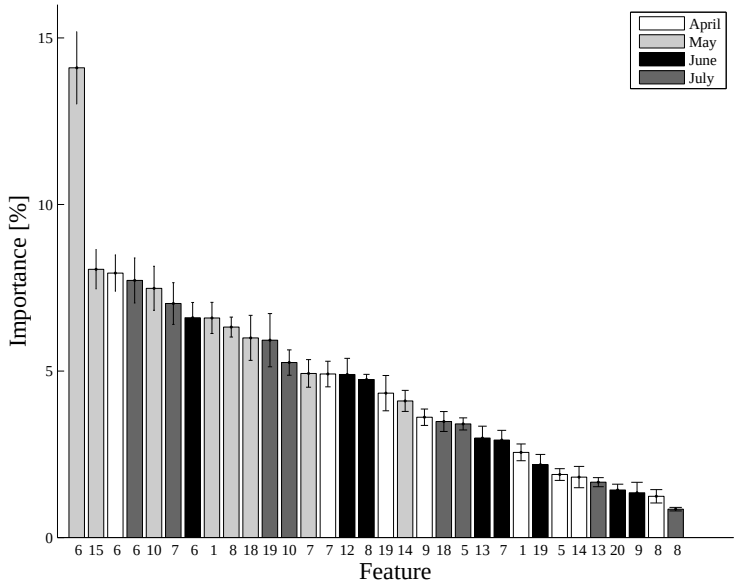


Figure 11.10: Importance score (averaged over 10 model runs with indication of standard deviation, and calculated following Algorithm 3) for the classification on the 8 most important features per date (April 17, May 20, June 15 and July 16). The numbers on the abscis correspond to specific features from Table 11.3.

reduction is consistent with the results from the accuracy-oriented reduction since the features with the highest $C_{j,\text{avg}}$ in the former strategy, were also the features that were identified as 'general' features in the latter strategy. Although features γ_{hv} and γ_{ll} were indicated as highly correlated in Figure 11.5, they were found to provide complementary information for crop classification. Especially for the identification of rye, winter wheat and winter barley, a combination of γ_{hv} and γ_{ll} was nonredundant and the correlation between both features at the class-level was low (*cfr.* Table 11.4). On the contrary, these features showed a very low correlation with ϕ_{hhvv}.

11.5.3. Map quality analysis

In this section, the effect of feature reduction on the quality of the resulting land cover map is evaluated. Evaluation is restricted to four multi-date classification models, representing four different degrees in dimensionality:

1. *Complex model*: constructed with all 20 polarimetric features per acquisition date (Table 11.3).

2. *Optimal model*: constructed with the subset of eight most important features per acquisition date (Table 11.9) as resulted from the accuracy-oriented

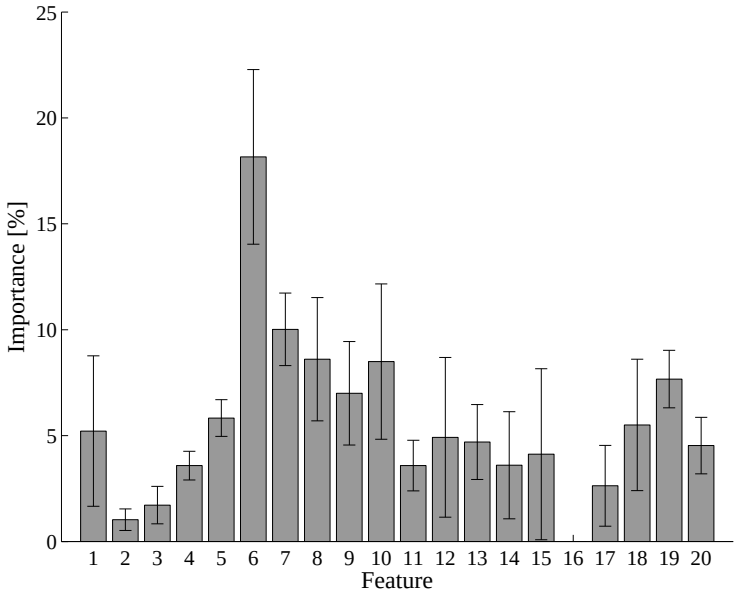


Figure 11.11: Average importance score $C_{j,\mathrm{avg}}$ (error bars indicate the standard deviation) of the input variables over all acquisition dates (April 17, May 20, June 15, July 16). The numbers on the abscis correspond to specific features from Table 11.3.

feature reduction.

3. *Practical model*: constructed with the subset of three most important features over the entire period (April 17-July 16), *i.e.* γ_{hv}, ϕ_{hhvv} and γ_{ll}, as resulted from the efficiency-oriented feature reduction.

4. *Simple model*: constructed with only the most important feature, *i.e.* γ_{hv} on all dates.

An overview of the obtained OA, κ and $\mathrm{AUC}_{\mathrm{avg}}$ for all four models is given in Table 11.10. The complex model and the optimal model showed negligble differences in their accuracy. The practical model resulted in a slighty lower classification accuracy, whereas the accuracy was substantially lower in case of the simple model. The decrease in OA of the practical model was mainly attributed to its lower performance for classifying winter wheat and winter barley (Table 11.11), although decreases in CA were limited to 4%. Despite the dramatic decrease in OA of the simple model, its value remained high (80.79%). Table 11.11 shows that the simple model failed to identify peas and winter wheat with an acceptable accuracy level (64.31% and 57.96%, respectively) due to confusion between peas and beets and between winter wheat, winter barley and rye (Table 11.12). Deteriorations in performance at the class-level were not observed in case the optimal model was used. A pairwise comparison at the pixel-level between the complex model and its reduced

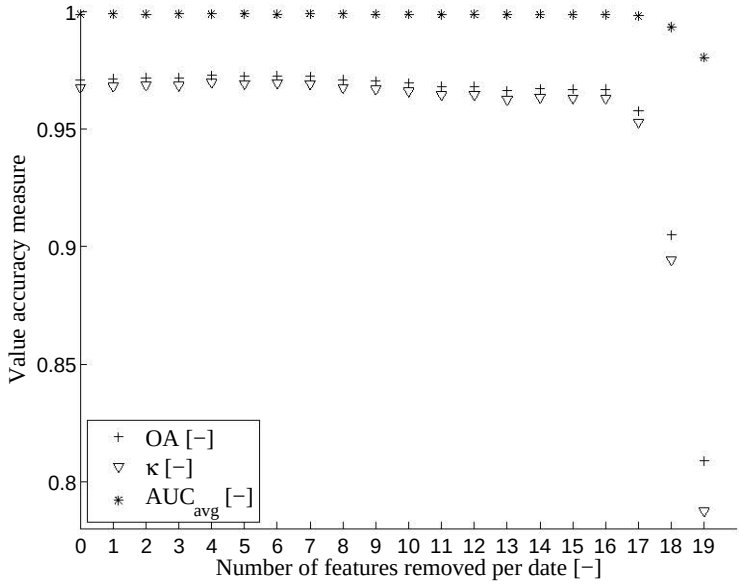


Figure 11.12: Overall accuracy, Cohen's kappa and $\mathrm{AUC}_{\mathrm{avg}}$ as a function of the number of features removed from the single-date data sets. The feature removal is based on the average importance score $C_{j,\mathrm{avg}}$ over all dates (*cfr.* Figure 11.11) with the removed feature being the same on each acquisition date. The numbers on the abscis correspond to specific features from Table 11.3.

alternatives was carried out by means of the McNemar test. A non-significant difference in performance between the complex model and the optimal model was observed for all land cover classes (Table 11.13). On the contrary, the performance of the complex model and the simple model was found to be significantly different for all classes, with the exception of water and forest. Despite the low observed CA for winter wheat and peas, the AUC (0.94 and 0.97, respectively) indicates that the simple model performed much better than random guessing and was fairly able to distinguish winter wheat and peas from other crops (Figure 11.13). The McNemar test between the complex model and the practical model (Table 11.13) pointed out that the practical model performed significantly worse than the complex model for most crops, including beets, peas, spring barley, winter barley, winter wheat and grass. In conclusion, the complex model and the optimal model reached similar levels of accuracy, whereas the practical and the simple model performed significantly worse for several crops.

In order to obtain an integral picture of the difference between the complex land cover model and the reduced alternatives, the probability vector was incorporated in the comparison. Based on this vector, $U = 1 - p_{\max}$ and H were calculated for each pixel. The predicted class, defined as the land cover class corresponding to $p_{\max}$, was compared with the observed land cover and was evaluated as either

Table 11.10: Accuracy measures of four multi-date models, representing a different degree of dimensionality (the complex, optimal, practical and simple model are based on 20, 8, 3 and 1 features, respectively).

Model Type	OA [%]	κ [-]	AUC_{avg} [-]
Complex	97.21	0.969	0.999
Optimal	97.13	0.968	0.999
Practical	95.81	0.953	0.998
Simple	80.79	0.787	0.980

Table 11.11: Class-specific accuracies (CA) of four multi-date models, representing a different degree of dimensionality (the complex, optimal, practical and simple model are based on 20, 8, 3 and 1 features, respectively).

Model	Wa	Fo	Be	Pe	Sb	Oa	Ry	Wb	Ww	Gr
Complex	99.70	99.70	98.89	98.49	95.67	97.28	97.58	94.96	90.42	99.40
Optimal	99.50	99.70	99.09	97.88	95.77	97.18	97.18	94.56	91.13	99.29
Practical	99.50	99.80	96.98	96.57	94.46	97.88	96.57	90.42	88.00	97.88
Simple	99.29	99.50	71.37	64.31	76.01	89.01	84.68	70.87	57.96	94.86

correct or incorrect. The frequency distributions of U and H are shown in Figures 11.14 and 11.15 for the correct and incorrect predictions, respectively. The shape of these distributions is indicative for the strength of the classification model and hence for the quality of the resulting land cover map. Concerning the correct predictions, the shape of the distribution of U was only slightly affected by the complexity of the classification model (Figure 11.14 (a)). The frequency peaks within the $[0, 0.1]$ interval, followed by an exponential decrease towards higher values of U. Only a limited fraction of the correct predictions was associated with an uncertainty U above 0.5. The shape of the entropy distribution (Figure 11.14 (b)) was found to be very similar to the distribution of U, which indicates an inverse relationship between H and p_{max} if a prediction was made correctly. Obviously, the simple model had the lowest prediction strength. This was indicated by a shift in the frequency of correctly classified pixels towards higher values of U and H. On the contrary, not the complex model but the optimal model was attributed the highest prediction strength, followed by the practical model. In case the latter models were used, a higher portion of correct predictions had uncertainties U and H lower than 0.1 and a lower portion within higher uncertainty intervals was observed, as compared to the complex model. The shape of the distribution of the uncertainty measures was found to be completely different for the incorrect predictions (*cfr.* Figure 11.15). Furthermore, larger differences between the four models occurred. The distribution of U showed a maximum between 0.4 and 0.6 in case the complex, the optimal or the practical model was used. This maximum occurred in the $]0.4, 0.5]$ interval if the simple model was applied. For entropy, the frequency peaks within lower uncertainty intervals, ranging from 0.3 to 0.6 and depending on the complexity of the model. Besides the position of the maximum,

Table 11.12: Confusion matrix as resulted from the simple model (based on γ_{hv}). The main confusions are indicated in bold.

	Wa	Fo	Be	Pe	Sb	Oa	Ry	Wb	Ww	Gr
Wa	985	0	1	0	0	0	0	1	1	0
Fo	0	987	0	0	3	0	1	0	3	5
Be	1	0	708	**295**	35	1	3	5	7	13
Pe	0	0	**208**	638	24	0	0	16	3	3
Sb	2	0	27	25	754	93	19	36	32	0
Oa	4	0	0	1	99	883	0	7	5	0
Ry	0	0	10	0	13	0	840	17	**137**	20
Wb	0	0	2	25	27	9	4	703	**189**	1
Ww	0	1	17	5	36	6	**110**	**203**	575	9
Gr	0	4	19	3	1	0	15	4	40	941

also the shape of the distribution showed larger differences between U and H in case of incorrect predictions. This implies that the inverse relationship between $p_{\max}$ and H was less prominent for the incorrect predictions. These findings illustrate that higher portions of incorrect predictions were made with lower uncertainties when the dimensionality of the classification model increased. This is in contrast with the correct predictions, for which increased uncertainties were recorded when the dimensionality of the model became very low.

The results are summarized by the mean of U ($\overline{U}$) and the mean entropy ($\overline{H}$) for the correct predictions on the one hand, and the incorrect predictions on the other hand. Based on Table 11.14, it can be concluded that, if the model (irrespective of the complexity) made a correct prediction, it had limited doubts about the predicted class and confusion with other classes was of minor importance (low $\overline{H}$ and $\overline{U}$). Both the optimal model and the practical model showed lower uncertainties in the correct predictions as compared to the complex model, whereas the simple model resulted in increased uncertainties. On the contrary, incorrect predictions were made with moderate values of $\overline{H}$ and $\overline{U}$. This implies that in case of an incorrect prediction, the multi-date model confused between two or three similar crop types. The complex model and the optimal model were comparable at the level of the incorrect predictions, whereas the simple model showed a substantially lower classification uncertainty. The characteristics of the incorrect predictions made by the practical model can be situated somewhat in between those of the complex model and the simple model. These results confirm the earlier findings from the analysis of the uncertainty distributions.

Table 11.13: Comparison of the complex model (based on 20 features) with the optimal (based on 8 features), the practical (based on 3 features) and the simple model (based on 1 feature) by means of the McNemar test: y, significant difference in performance between the models, n, no significant difference, both at the 0.05 significance level; n_{01} and n_{10} are the error rates from the contingency table.

	Complex (1) versus optimal model (2)									
	Wa	Fo	Be	Pe	Sb	Oa	Ry	Wb	Ww	Gr
McNemar	n	n	n	n	n	n	n	n	n	n
n01	1	0	1	7	9	6	9	13	15	2
n10	2	0	2	8	8	7	9	11	11	2
	Complex (1) versus practical model (2)									
	Wa	Fo	Be	Pe	Sb	Oa	Ry	Wb	Ww	Gr
McNemar	n	n	y	y	y	n	n	y	y	y
n01	2	3	4	7	8	15	18	22	40	0
n10	3	0	24	25	23	9	25	66	67	15
	Complex (1) versus simple model (2)									
	Wa	Fo	Be	Pe	Sb	Oa	Ry	Wb	Ww	Gr
McNemar	n	n	y	y	y	y	y	y	y	y
n01	1	0	3	3	10	14	12	17	39	1
n10	5	2	276	342	205	96	140	256	361	46

Table 11.14: Mean uncertainty ($\overline{U}$) and mean entropy ($\overline{H}$) of the correct and incorrect predictions, as resulted from the complex (based on 20 features), optimal (based on 8 features), practical (based on 3 features) and simple (based on 1 feature) multi-date classification model.

	Complex		Optimal		Practical		Simple	
Prediction	$\overline{U}$	$\overline{H}$	$\overline{U}$	$\overline{H}$	$\overline{U}$	$\overline{H}$	$\overline{U}$	$\overline{H}$
Correct	0.11	0.16	0.08	0.12	0.09	0.13	0.14	0.16
Incorrect	0.51	0.53	0.47	0.49	0.45	0.45	0.38	0.37

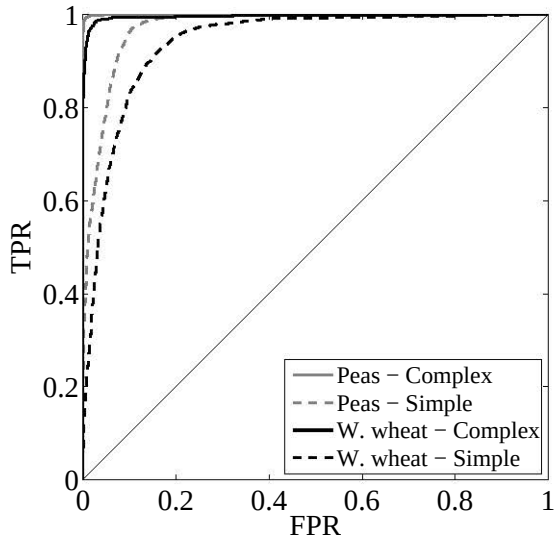
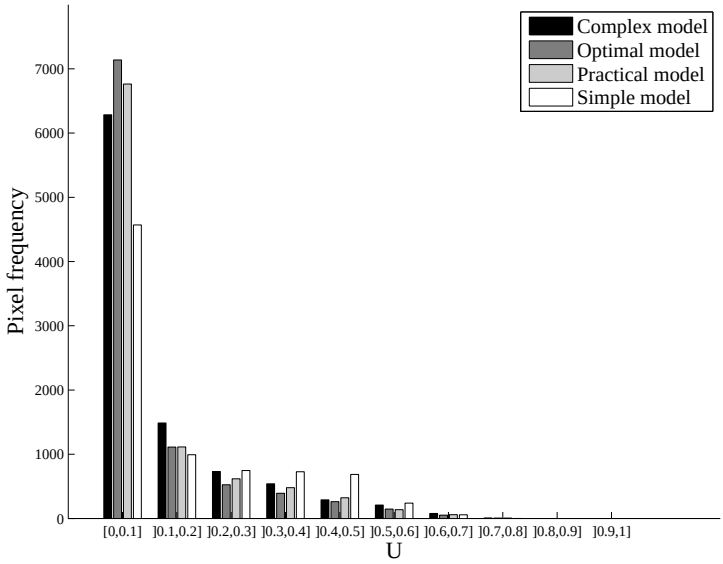
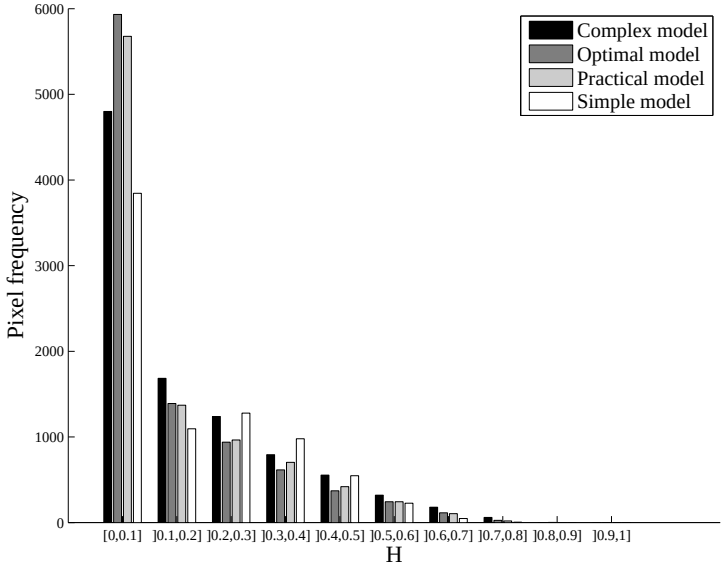


Figure 11.13: ROC curve for the prediction of peas and winter wheat with the complex model (based on 20 features) and the simple model (based on 1 feature), the 0-1 line indicates the curve corresponding with random guessing.

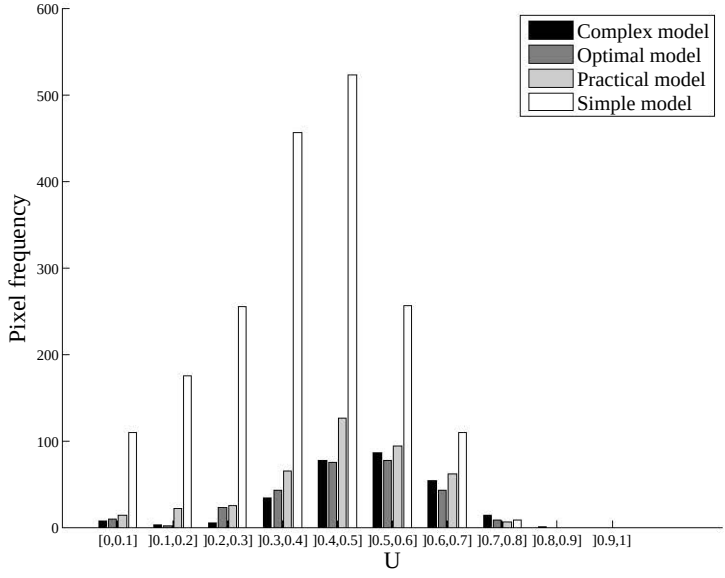


(a)

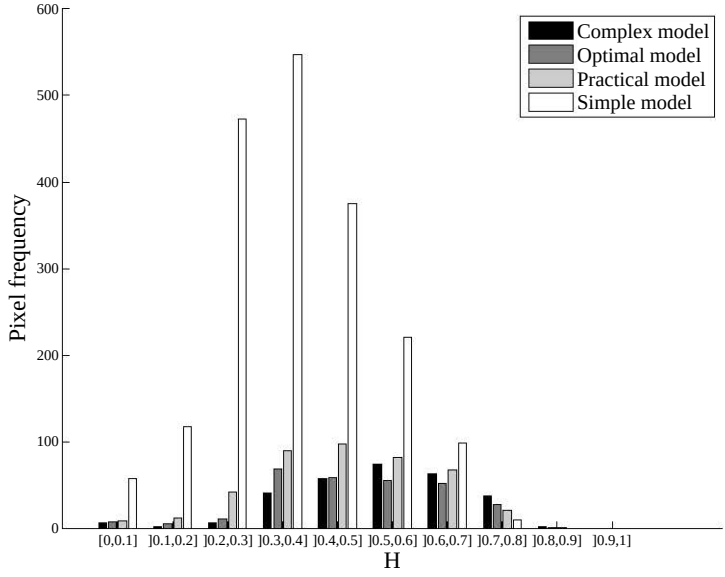


(b)

Figure 11.14: Frequency distribution of $U = 1 - p_{\max}$ (a) and the Shannon entropy H (b) for the correct predictions as resulted from the complex, the optimal, the practical and the simple model.



(a)



(b)

Figure 11.15: Frequency distribution of $U = 1 - p_{\max}$ (a) and the Shannon entropy H (b) for the incorrect predictions as resulted from the complex, the optimal, the pratical and the simple model.

11.6. Conclusion

Multi-date polarimetric SAR has been increasingly used for land cover mapping (Blaes et al., 2005, Skriver, 2012, Skriver et al., 2011) since a single acquisition within the growing season was often found to be insufficient (Stankiewicz, 2006, Frate et al., 2003) to reach the commonly recommended target classification accuracy of 85% (Foody and Atkinson, 2002). However, a disadvantage of the multi-date approach is the high dimensionality of the data set which may hamper the computational efficiency and which may reduce the interpretability of the modelling results. Furthermore, reducing the number of SAR features in the classification process can possibly enhance the quality of the resulting land cover map. Zou et al. (2010) reported a slight increase (0.2-0.5%) in overall accuracy if a selected feature set (based on a ranking in selection metric) was used instead of combining all features at hand.

In the first part of this chapter, Random Forests (a soft classifier) was introduced for land cover classification of an agricultural area based on high-dimensional SAR image data and a method to easily estimate classification uncertainty at pixel-level was presented. This method was applied to polarimetric EMISAR data in order to evaluate the effect of multi-configuration in the data set on the resulting classification uncertainty. In the second part of this chapter, we presented two different strategies to reduce the number of polarimetric SAR features when constructing a multi-date model for land cover classification: (1) an accuracy-oriented feature reduction, in which the preservation of the model accuracy is the main goal, and (2) an efficiency-oriented feature reduction, in which small losses in accuracy are tolerated in favour of less feature requirements. The advantage of the latter method over the former is that it provides the most informative type of features irrespective of the date of SAR retrieval. For both approaches, it was investigated how the feature reduction affected the quality of the resulting land cover map in terms of classification accuracy and classification uncertainty. The methodology for feature reduction was based on the importance score generated by Random Forests.

The conclusions from this study are:
1. Uncertainty assessment is a valuable addition to the established accuracy assessment of classification results. Uncertainty measures $U = 1 - p_{\max}$ and entropy H of the modelled probability vector are two alternative methods to quantify the uncertainty of classification results at pixel-level. Although the former measure shows a rather linear behaviour of the modelled probability of land cover occurrence and the latter a nonlinear behaviour, both measures produced similar uncertainty assessment results in this study. Because entropy gathers information of the entire probability vector, it is to be preferred.

2. The McNemar tests indicated a significant improvement in classification quality

of land cover maps that were obtained by single-configuration models, followed by multi-frequency models and multi-date models to multi-date-frequency models. The acquisition of multi-date-frequency imagery may be challenging, however, the availability of multi-date imagery is foreseen to increase with future planned SAR missions as SENTINEL-1, the RADARSAT constellation and L-band ALOS-II. Multi-frequency models were outperformed by multi-date models, because the imagery obtained at the different dates represented polarimetric SAR signatures of distinct stages in the crop growing season. The relationship between crop development and classification performance should therefore be acknowledged for land cover mapping.

3. Single-date L-band SAR images resulted in a higher classification quality than single-date C-band SAR imagery. Accuracy values were lower for the latter and the uncertainty assessment indicated C-band models to have less confidence in the predicted land cover class. By using a multi-date SAR data set (4 acquisitions), the difference between L- and C-band data with respect to the performance of the classification model was almost entirely reduced.

4. Results of the uncertainty analysis showed a decreasing uncertainty in the correctly classified pixels as the dimensionality of the SAR imagery increased. However, uncertainty values were highly variable, both in time and space. Generally, early-spring acquisitions and the occurrence of winter crops tend to compromise the model uncertainty. The spatial uncertainty analysis revealed boundary effects of mixed pixels between adjacent crop field and patterns of locally increased uncertainty.

5. Further application of classification uncertainty at the pixel level with probabilistic classifiers such as Random Forests is particularly interesting when the land cover maps are used in spatially distributed environmental models. Knowledge on the uncertainty in the land cover data allows for a better interpretation of the results of these models and for uncertainty-weighted post-processing of the classification result.

6. Feature importance generally depended on the time of acquisition. However, some features such as γ_{hv}, ϕ_{hhvv}, γ_{ll} were identified as 'general' features for crop classification as their importance score remained high during the entire growing season. Other features, such as $\overline{\alpha}$, $\gamma_{\mathrm{hv}}/\gamma_{\mathrm{hh}}$ and $\gamma_{++45}/\gamma_{+-45}$, were identified as 'specific' features for crop classification as they showed a high importance score for a specific period within the growing season. However, we recommend caution when extrapolating these results to other periods, frequencies and study sites as the importance ranking might differ accordingly. The features that were designated as 'general' and 'specific' only apply to the L-band EMISAR data that was used in this study.

7. The accuracy-oriented reduction approach resulted in a multi-date model constructed on the eight most important features per date. Since the selected

features were not necessarily the same on all dates, 14 different features (both 'general' and 'specific') were incorporated in the construction of the model. The reduced model was able to preserve the same level of overall accuracy (OA = 97.13%) compared to the nonreduced model (OA = 97.21%). Furthermore, class-specific accuracies were unaffected as well. Moreover, the reduced model produced lower uncertainties in the classification result. More confidence was gained about the correct predictions, which were characterized by a low value of $U = 1 - p_{\max}$ and entropy H.

8. The efficiency-oriented reduction approach resulted in a multi-date model constructed on a subset of 3 features with the highest mean importance score over the entire period. The subset of features (γ_{hv}, ϕ_{hhvv} and γ_{ll}) was the same on all dates and exclusively consisted of 'general' features. The reduced model showed a slight decrease in overall accuracy (OA = 95.81%) compared to the nonreduced model due to a deterioration in the detection of winter wheat and winter barley. The efficiency-oriented reduction resulted, in common with the accuracy-oriented reduction, in lower uncertainties in the model predictions. Additionally, a comparison was carried out between the selected multi-date models and an alternative multi-date model that was constructed with only the most important feature, *i.e.* γ_{hv}, on all dates. Although the latter model was still able to produce an OA of 80.79%, it had serious problems to discriminate peas (CA = 64.31%) and winter wheat (CA = 57.96%). This extremely reduced model showed higher uncertainty values than the nonreduced model.

9. Feature reduction is beneficial for land cover classification as it resulted in a strong dimensionality reduction of the SAR data sets without substantially affecting the accuracy of the classification. Moreover, a decline in classification uncertainty was realized when feature reduction was performed, which is valuable when land cover classification is coupled with spatially distributed hydrological models. Depending on the purpose of the land cover map and thus on the objectives of the feature reduction, an appropriate dimensionality of the model (and hence quality of the land cover map) can be determined. Furthermore, the importance score is found to be a useful tool for feature reduction and provides useful information about the predictive capacity of SAR observables throughout the growing season.

12

Impact of classification uncertainty on the prediction of water fluxes

12.1. Introduction

Spatially distributed hydrological models require information on the spatial distribution of land cover to simulate the water fluxes in a watershed. For each land cover, a set of biophysical parameters is defined in order to describe the land surface characteristics, usually within a land cover look-up table (LC-LUT). As such, the effect of land cover on the hydrological model response is ultimately controlled by the biophysical parameters. The latter play a crucial role within the model because they determine the energy and moisture exchange between the land surface and the atmosphere, such that changes in the vegetation parameters alter the energy budget in the coupled land-atmosphere-transfer scheme and directly affect evapotranspiration, runoff and infiltration. Since hydrological simulations are typically performed under conditions of data scarcity (including predictions in ungauged basins), land cover information is (highly) uncertain and affects the reliability of the simulation result when the uncertainty propagates through the model. From a societal point of view, it is important to quantify this uncertainty because predictions of water fluxes are often used in early-warning systems for natural hazards or for assessing the effect of water resource infrastructures (Montanari et al., 2009).

When a land cover map of the catchment is not readily available, the latter needs to be derived from remotely sensed data through classification analysis (Singh and Woolhiser, 2002). However, one should be aware that the resulting map is an approximation of the complex reality and that substantial discrepancies between the real land cover and its representation may be present (Zhang and Goodchild, 2002). This awareness has been spread among the user community as there is a growing demand to better document the quality of the produced map (Canters et al., 2002). Data quality research is often limited to simple overall measures e.g. Cohen kappa coefficient, while information on the spatial variation in error is lacking. Nevertheless, it is straightforward that an erroneous classification is most likely in areas with a high heterogeneity and for land cover types with a high similarity in

spectral properties or physical characteristics. It is therefore important (i) to obtain more detailed (spatial) information on the uncertainty of the classified map because the true value of the derived map cannot be assessed when this information is lacking, and (ii) to understand how land cover uncertainties affect the hydrological model predictions. As such, an effective support for decision-making systems is provided.

Over the last 25 years, more attention has been given to uncertainty assessment and uncertainty propagation in models driven by remotely sensed data (Heuvelink and Burrough, 2002). For a review, we refer to Crosetto et al. (2000, 2001b,a) and Foody and Atkinson (2002). Despite the growing insight into the effect of different land cover data sets (e.g. Pauwels and Wood, 2000, Wegehenkel et al., 2006) and different spatial resolutions of the land cover map (e.g. Armstrong and Martz, 2008, Pauwels and Wood, 2000) on the environmental model response, research on the effect of classification error is limited and often based on simple accuracy measures (e.g. Kyriakidis and Dungan, 2001, Miller et al., 2007, Livne and Svoray, 2011). In addition, also the impact of biophysical parameter uncertainy on environmental model predictions has been investigated (Breuer et al., 2006, Eckhardt et al., 2003, Liang and Guo, 2003), but was found to be less important than soil physical parameter uncertainty (Liang and Guo, 2003). Although an increasing number of studies on uncertainty propagation is being published, many questions about the effect of land cover information quality on hydrological model predictions remain unanswered. In this chapter, the objectives w.r.t. this research topic are:

1. Evaluation of the uncertainty in water flux predictions due to land cover classification uncertainty. With the uncertainty analysis (UA) we intend (i) to identify the conditions under which land cover uncertainty has the highest impact and (ii) to better support the land cover classification schematization as a function of the modelling purpose. In some cases, a lower quality or less detailed land cover map may be sufficient and resources can be saved.

2. Evaluation of the sensitivity of water flux predictions to small changes in the biophysical parameters. Results of the sensitivity analysis (SA) allow to identify the biophysical parameters for which the hydrological model is most sensitive. By decreasing the uncertainty in these parameters, the reliability of the water flux predictions can be improved.

The hydrological model that has been used in this study to perform the UA and SA is the spatially distributed, physically-based TOPLATS model, which has been applied to the catchment of the Bellebeek, located in Belgium. A brief description of the hydrological model set-up is given in Section 12.2.1. A soft land cover map of the area was derived through classification of multi-temporal SPOT imagery. The methodology to propagate the uncertainty in the classified land cover map through the hydrological model is described in Section 12.2. The method to assess the sensitivity of the model outputs w.r.t. the biophysical parameters is explained

in Section 12.2. Finally, the results of the uncertainty and sensitivity analysis are discussed in Section 12.3.

12.2. Methodology

12.2.1. Hydrological model set-up

The hydrological model TOPLATS was applied to the catchment of the Bellebeek. Therefore, we forced the model with the climatic data for the period January 1, 2006 to December 31, 2006 (*cfr.* Section 7.3). Information on the soil type and on the topography was provided by respectively the STM-FAO and DEM (*cfr.* Section 7.1). Land cover information was provided by:

- LCM-CROP: a detailed land cover map with a level 4 classification scheme (*cfr.* Figure 10.1), indicating 6 general land cover classes (*cfr.* Table 7.2). This map is used as reference land cover map, representing the ground truth. From this map, the labels for the reference data set, which is used to train the Random Forest classifier, are extracted.

- LCM-RF: a soft classified land cover map constructed from three cloud-free SPOT images of the study area acquired on June 13 (DOY 164), June 30 (DOY 181) and July 5 (DOY 186) 2006, with a spatial resolution of 20 m, (*cfr.* Section 7.2) through a Random Forests classification (*cfr.* Section 11.3.1). The classification scheme of LCM-RF was chosen in correspondence with the reference land cover information from LCM-CROP such that the classified land cover map indicates 12 different land covers. There are 5 general classes: urban area (Ur), water area (Wa), broad-leaved forest (Bf), needle-leaved forest (Nf), pasture (Pa) and 7 specific classes: grass (Gr), potatoes (Po), beets (Be), maize (Ma), wheat (Wh), barley (Ba) and other crop species (Oc). In contrast with LCM-CROP, the cropland class (Cr) is no longer present in LCM-RF as the agricultural fields are assigned one of the specific crop classes.

All geographical data of the study area were converted to a grid with a resolution of 50 m. The corresponding soil and vegetation parameters were attributed literature-based values as described in Section 8.2. The calibrated values for the TOPMODEL and routing parameters were used (*cfr.* Tables 8.1 and 8.3). An overview of the time-varying vegetation parameters from the LC-LUT is given in Table 12.1, together with the attributed values. The other parameters from the LC-LUT (Table 6.4), such as the emissivity and the stomatal resistance, were assumed time-invariant.

Table 12.1: Cover-dependent and time-varying (updated at a monthly time step) vegetation parameters for the different vegetation types in the TOPLATS simulations. For the leaf area index LAI, albedo α, momentum roughness $z_{0,m}$, heat roughness $z_{0,h}$ and zero plane displacement d_0, the minimum and maximum value within the growing season are listed.

	LAI [−][a]	α [−][b]	$z_{0,m}$ [m][c]	$z_{0,h}$ [m][d]	d_0 [m][e]	h [m][f]	Growth
Bf	[1.0,5.4]	[0.10,0.15]	0.59	0.06	7.80	10	-
Nf	[4.0,6.3]	[0.09,0.12]	0.70	0.07	9.36	12	-
Gr	[0.1,6.0]	[0.17,0.20]	[0.02,0.17]	[0.002,0.017]	[0.10,0.90]	[0.15,1.35]	-
Pa	[1.5,3.0]	[0.17,0.20]	0.04	0.004	0.34	0.50	-
Cr	[0.5,3.8]	[0.23,0.32]	[0.04,0.19]	[0.004,0.019]	[0.19,0.92]	[0.30,1.20]	March-Nov
Po	[1.5,5.2]	[0.23,0.32]	[0.02,0.06]	[0.002,0.006]	[0.13,0.45]	[0.20,0.70]	April-Sept
Be	[0.1,5.3]	[0.23,0.32]	[0.02,0.08]	[0.002,0.008]	[0.16,0.64]	[0.25,1.00]	April-Sept
Oc	[0.5,3.8]	[0.23,0.32]	[0.02,0.12]	[0.002,0.012]	[0.19,0.92]	[0.30,1.44]	March-Nov
Ma	[0.1,5.4]	[0.23,0.32]	[0.03,0.17]	[0.003,0.017]	[0.26,1.28]	[0.40,2.00]	May-Oct
Wh	[0.1,3.8]	[0.23,0.30]	[0.01,0.09]	[0.001,0.009]	[0.02,0.70]	[0.03,1.10]	Nov-July
Ba	[0.1,3.7]	[0.23,0.30]	[0.01,0.07]	[0.001,0.007]	[0.06,0.58]	[0.10,0.90]	Nov-July

[a] Monthly values are taken from Tian et al. (2004), Breuer et al. (2003), Tekalign and Hammes (2005) and Pauwels et al. (2007).
[b] Monthly values are taken from Kotoda (1986).
[c] Calculated from the canopy height h following Jarvis et al. (1976), Jones (1992), Campbell (1977) or Allen et al. (1988) for forest, bare soil, cropland and grassland, respectively.
[d] Calculated from $z_{0,m}$ following the relation $z_{0,h} = 0.1 \cdot z_{0,m}$ (Allen et al., 1988).
[e] Calculated from the canopy height h following Jarvis et al. (1976), Jones (1992), Campbell (1977) or Allen et al. (1988) for forest, bare soil, cropland and grassland, respectively.
[f] Values for the maximal crop height are taken from Breuer et al. (2003); The crop height equals zero outside the growing season and linearly increases within the growing period (Allen et al., 1988).

12.2.2. Uncertainty analysis

The purpose of the uncertainty analysis is to estimate the probability density function (PDF) of the hydrological model output when the uncertainty in the classified land cover map is taken into account. The performed uncertainty analysis consists of two main steps: (i) an uncertainty modelling step, *i.e.* quantifying the classification uncertainty, and (iii) an uncertainty propagation step, *i.e.* applying the hydrological model with error-corrupted versions of the land cover map. The methods that were used for uncertainty modelling and uncertainty propagation are Random Forests classification and Monte Carlo simulation, respectively. A flowchart of the methodology is given in Figure 12.1.

Uncertainty modelling

The uncertainty in the land cover classification was modelled using the Random Forests classifier, which is an ensemble learning method that generates many tree classifiers and aggregates their results (Breiman, 2001). For a detailed description of the algorithm we refer to Section 11.3.1. By aggregating the predictions of n_{tree} classification trees, each pixel u is attributed a class membership probability vector:

$$\mathbf{p_u} = (p_u(1), p_u(2), ..., p_u(n_{\text{class}})) = \left(\frac{k_u(1)}{n_{\text{tree}}}, \frac{k_u(2)}{n_{\text{tree}}}, ..., \frac{k_u(n_{\text{class}})}{n_{\text{tree}}} \right) \qquad (12.1)$$

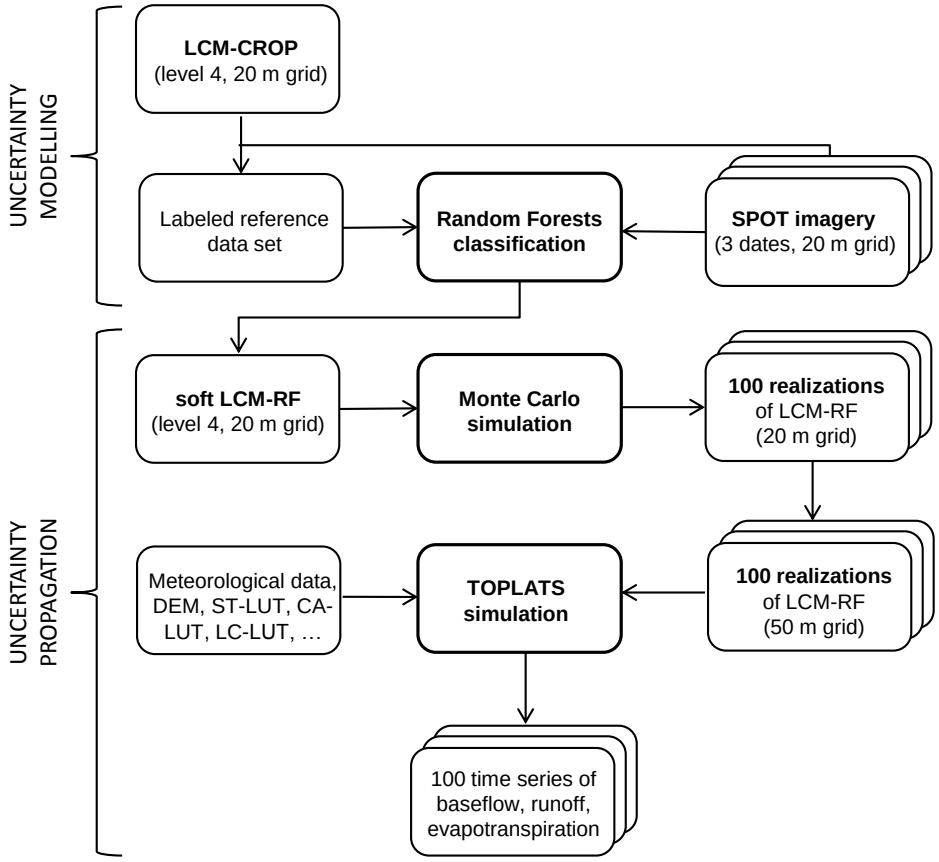


Figure 12.1: Schematic of the uncertainty analysis, including an uncertainty modelling step and an uncertainty propagation step.

with n_{class} the total number of classes (here $n_{\text{class}} = 12$), n_{tree} the total number of trees in the classification process (here $n_{\text{tree}} = 1000$), $k_u(i)$ the number of trees classifying pixel u into class i and $p_u(i)$ the probability that pixel u is classified into class i. The set of probability vectors for each pixel in the land cover map forms the uncertainty model, *i.e.* a stochastic model capable of generating a population of error-corrupted versions of the same pattern in geographical variation (Goodchild et al., 1992). The soft (probabilistic) classification result is usually converted into a hard prediction by applying the majority vote for the ensemble of trees: pixel u is attributed the class that corresponds to the maximum probability $p_{\max}$ from the probability vector $\mathbf{p_u}$. Likewise, the oob predictions of the individual trees are aggregated to estimate the error rate of the ensemble classifier. To obtain an unbiased oob error, the number of trees should be large enough (Liaw and Wiener, 2002, Bylander, 2002). The parameter m that defines the number of variables used to split the nodes was optimized through minimization of the oob error (here

$m = 3$).

While the oob-error and the confusion (or contingency) matrix provide the user with a general indication for the quality of the classified map, the probability vector allows for the calculation of local indices of map quality (Unwin, 1995). In this chapter, the classification quality was evaluated in terms of the entropy H (Shannon, 1948) (*cfr.* Eq. (11.6)). The resulting entropy map allows to identify areas of high and low map quality. This knowledge leads towards a better interpretation of the classification result and towards a higher efficiency in the ground survey. For more information on accuracy and uncertainty measures, we refer to Section 11.3.2.

Uncertainty propagation

By means of a Monte Carlo (MC) simulation, the classification uncertainty was sampled to obtain a set of error-corrupted versions of the land cover map. For a detailed description of the Monte Carlo method, we refer to Section 3.3.1. After sampling, the classification uncertainty was propagated through multiple evaluations of the hydrological model response with the error-corrupted versions (or realizations) of the land cover map. The method was implemented as follows:

1. A random set of n_{MC} realizations of the land cover map LCM-RF was generated based on the stochastic error model $\mathbf{p_u}$ that resulted from the Random Forests classification in the uncertainty modelling step. Before moving to step 2, the resolution of the error-corrupted versions of the land cover map was converted from 20 m to 50 m using a nearest neighbour resampling technique. Aggregation of the land cover map was necessary to assure a feasible computation time but did not affect the original distribution of the classification error (as given by the confusion matrix, not shown).

2. The baseflow Q_b, surface runoff Q_s and evapotranspiration ET were evaluated with TOPLATS for each realization of LCM-RF while keeping all other model inputs and parameters unchanged. Because of the high computational requirement of the hydrological model, it is chosen to limit the MC analysis to 100 model evaluations (hence $n_{\mathrm{MC}} = 100$). Although this number might be insufficient to accurately describe the uncertainty in the water flux predictions, the results of the MC analysis are useful to provide a general indication of the effect of classification uncertainty on the model predictions.

3. An empirical PDF of the water fluxes Q_b, Q_s and ET was constructed based on the set of model simulations obtained in step 2. The mean, the variance and the prediction quantiles were derived from the PDF to represent the uncertainty in the predicted water fluxes.

To generate error-corrupted versions of the classified land cover map (20 m resolution), a field-based simulation approach was used. In this approach, the structural homogeneity of the landscape is preserved during simulation, whereas in the pixel-

based simulation approach the simulated land cover maps are highly fragmented and do not accurately reflect the actual land cover pattern (Canters et al., 2002). A schematic of the pixel-based and field-based simulation approach is shown in Figure 12.2. In the *pixel-based simulation approach*, the land cover is simulated independently for each pixel based on its class membership probability vector. For each pixel, the membership probabilities are matched to a sub-interval from [0,1] and a random number between 0 and 1 is selected from a uniform distribution. The pixel is then attributed the class that corresponds with the sub-interval that includes the random number. As such, the probability distribution is respected across a large number of realizations. In the *field-based simulation approach*, the same allocation mechanism as in the pixel-based simulation is applied, but the random number generation is performed at the field level (instead of the pixel level) in which for all pixels within one field the same random number is used for land cover assignment. A field is defined as a group of pixels with similar probability vectors. Although Canters et al. (2002) suggest to maximally use the information within the probability vector, it is chosen to delineate the fields in this study based on the most likely class as including also the second most likely class added a lot of noise to the land cover pattern.

12.2.3. Sensitivity analysis

The purpose of the sensitivity (SA) analysis is to estimate the sensitivity of the water flux predictions w.r.t. the time-varying vegetation parameters LAI, α, d_0, $z_{0,m}$ and $z_{0,h}$ (*cfr.* Table 6.4). Therefore, a local one-at-a-time (OAT) derivative-based sensitivity analysis was performed to quantify the response of the water fluxes Q_b, Q_s and ET to small changes in the vegetation parameters around their baseline values within the parameter space as defined in Table 12.1. As such, the main sources of prediction uncertainty can be identified (Saltelli et al., 2008b) and the uncertainty in the model prediction can be decreased by increasing the quality of the most sensitive model parameter(s). For more information on derivative-based methods for SA, we refer to Section 4.1. The choice of the SA method is driven by the factors prioritization setting (FP, *cfr.* Chapter 4) and by the highly complex hydrological model. Although we are aware that OAT-SA does not take into account correlations between the vegetation parameters (*cfr.* Section 4.1.2), we opted for this approach because of the ease of implementation (no modifications to the model code are required) and interpretation. The sensitivity of the water fluxes was expressed as a central finite difference approximation of the sensitivity function, using Eq. (4.1). Perturbation of the vegetation parameters was realized by multiplying the nominal parameter value by a user-defined perturbation factor ξ. The value of this factor should be small enough to result in an accurate approximation of the partial derivative, but should not be too large in order to avoid numerical inaccuracies (for more explanation on this topic, see Section 4.1.1). Here,

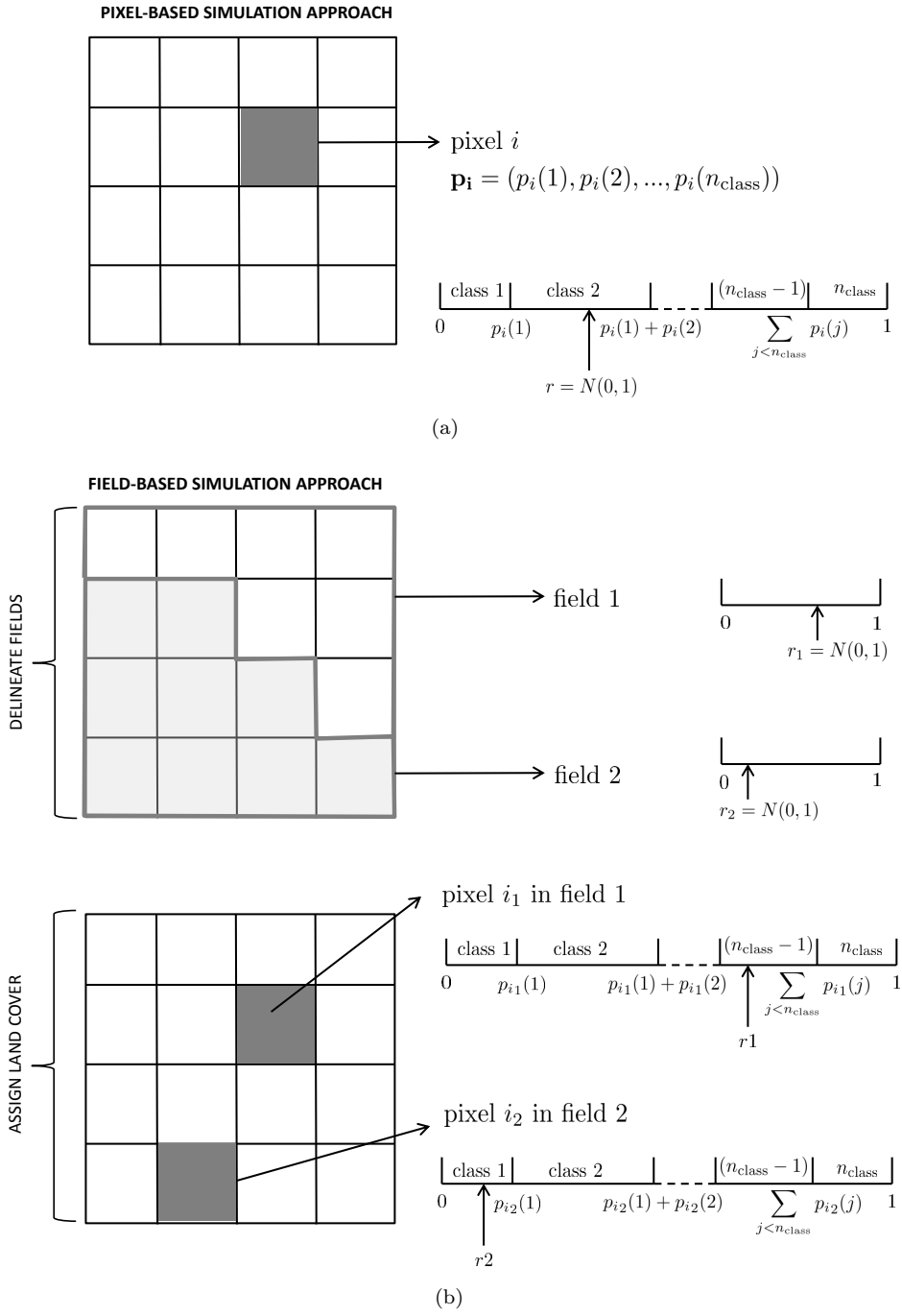


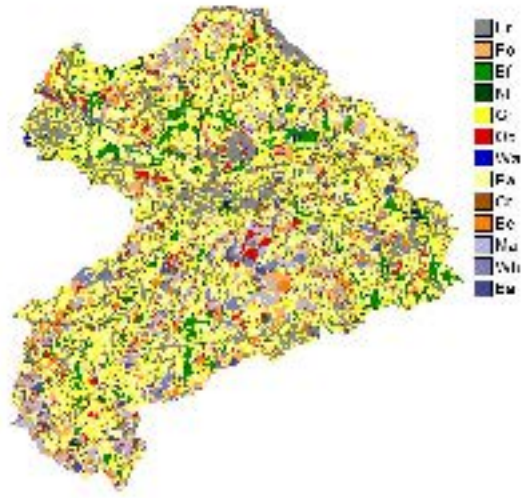
Figure 12.2: Pixel-based (spatially independent) simulation approach (a) and field-based simulation approach (b) to generate error-corrupted versions of the land cover map (modified from Canters et al. (2002)).

a fixed value of $\xi = 10^{-2}$ was chosen to perturb the model parameters based on the results of De Pauw and Vanrolleghem (2006) and Loosvelt et al. (2013). From the sensitivity function, a local dimensionless sensitivity index S was calculated following Eq. (4.2). The value of S indicates the intensity of the hydrological response w.r.t. changes in one of the vegetation parameters and is expressed in a relative manner. This allows to compare the sensitivity among different responses and parameters. The sign of S indicates how the model response reacts to changes in the model parameter: a positive value indicates a codirectional response in which an increase (decrease) of the parameter value results in an increase (decrease) of the model output, whereas a negative sign indicates an inverse model reponse.

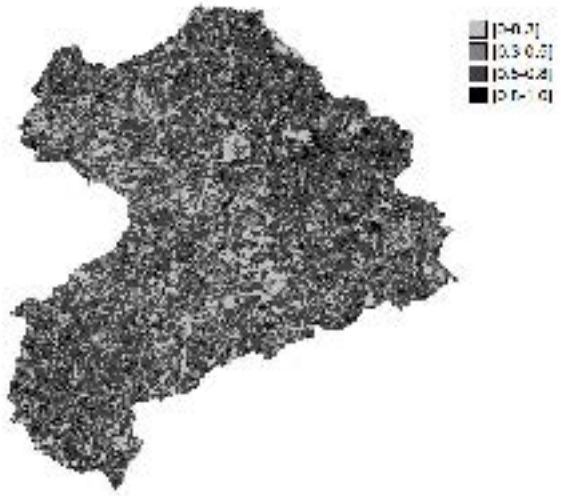
12.3. Results and discussion

12.3.1. Map quality analysis

A set of reference points was extracted from the reference map LCM-CROP by means of a stratified random sampling. For each land cover type (with exception of forest and water), 100 training pixels were selected and the corresponding radiances in the multi-temporal SPOT imagery (acquired on June 13, June 30 and July 5 in 2006, for more details we refer to Section 7.2) were extracted to build a balanced reference data set on which the Random Forest (RF) was trained. The resulting soft classified land cover map LCM-RF provides information about the probability of the land cover distribution at the pixel-level. In Figure 12.3, the most probable land cover distribution (*i.e.* the hard classification result, further referred to as $\text{LCM-RF}_{\text{hard}}$) is shown together with the entropy of the prediction. A visual inspection of the maps reveals that local patches of high quality in the classified land cover map mainly correspond to the classes broad-leaved forest and urban area. Table 12.2 summarizes the uncertainty in the land cover prediction in terms of the class-averaged entropy $\overline{H}$. Entropy values show that the uncertainty in all classified land covers is very high and that only a few classes have an entropy less than 0.5. The highest values were registered for needle-leaved forest and grass. In order to estimate the contribution of each land cover to the global map quality (considering the entire catchment), the entropy values were weighted with the areal cover percentage of each land cover (Table 12.2). Results show that uncertainty in the identification of grass and pasture is the main source of low map quality. This is mainly due to confusion between pasture and grass, as indicated by the high average class-membership probabilities in Table 12.3.



(a)



(b)

Figure 12.3: Hard classified land cover map of the Bellebeek (LCM-RF$_{\text{hard}}$) as obtained through Random Forests (a) and map of the entropy (uncertainty) in the predicted land cover (b).

Table 12.2: Entropy ($\overline{H}$) and areal cover percentage (a) per land cover class in the classified land cover map LCM-RF. The land cover classes include urban area (Ur), water area (Wa), broad-leaved forest (Bf), needle-leaved forest (Nf), pasture (Pa) and 7 specific classes: grass (Gr), potatoes (Po), beets (Be), maize (Ma), wheat (Wh), barley (Ba) and other crop species (Oc).

	Ur	Wa	Bf	Nf	Gr	Pa	Po	Be	Ot	Ma	Wh	Ba
$\overline{H}$	0.365	0.407	0.370	0.664	0.646	0.505	0.539	0.614	0.572	0.443	0.507	0.619
a	0.137	0.002	0.052	0.023	0.276	0.224	0.041	0.040	0.040	0.065	0.054	0.045
$\overline{H} \cdot A$	0.050	0.001	0.019	0.015	0.178	0.113	0.022	0.025	0.023	0.029	0.027	0.028

Table 12.3: Average membership probability vector $\overline{p}$ for the different land covers in the study area, as obtained from the soft classified LCM-RF. The maximum probability is indicated in bold. The land cover classes include urban area (Ur), water area (Wa), broad-leaved forest (Bf), needle-leaved forest (Nf), pasture (Pa) and 7 specific classes: grass (Gr), potatoes (Po), beets (Be), maize (Ma), wheat (Wh), barley (Ba) and other crop species (Oc).

	Ur	Wa	Bf	Nf	Pa	Gr	Po	Be	Ma	Wh	Ba	Oc
Ur	**0.698**	0.011	0.001	0.016	0.007	0.146	0.006	0.007	0.004	0.001	0.042	0.060
Wa	0.027	**0.752**	0.057	0.038	0.010	0.029	0.028	0.001	0.000	0.031	0.019	0.007
Bf	0.001	0.007	**0.688**	0.121	0.033	0.064	0.015	0.000	0.000	0.057	0.005	0.008
Nf	0.034	0.039	0.146	**0.445**	0.054	0.107	0.060	0.001	0.000	0.066	0.029	0.021
Pa	0.014	0.001	0.012	0.003	**0.552**	0.239	0.076	0.014	0.007	0.023	0.027	0.033
Gr	0.105	0.006	0.028	0.023	0.151	**0.450**	0.053	0.019	0.009	0.020	0.063	0.072
Po	0.008	0.002	0.011	0.011	0.122	0.090	**0.519**	0.103	0.067	0.022	0.012	0.033
Be	0.025	0.001	0.003	0.002	0.045	0.097	0.126	**0.456**	0.141	0.002	0.013	0.090
Ma	0.006	0.000	0.000	0.000	0.016	0.051	0.084	0.136	**0.624**	0.000	0.002	0.079
Wh	0.003	0.006	0.066	0.037	0.077	0.081	0.058	0.002	0.000	**0.595**	0.033	0.042
Ba	0.084	0.004	0.013	0.016	0.101	0.151	0.022	0.010	0.005	0.039	**0.484**	0.073
Oc	0.090	0.001	0.003	0.005	0.032	0.095	0.041	0.092	0.122	0.007	0.014	**0.497**

12.3.2. Effect of uncertain land cover on water flux predictions

In this section, it is investigated how and to what extent the uncertainty in the classified land cover map affects the prediction of water fluxes. Therefore, the uncertainty in LCM-RF was propagated through TOPLATS following the methodology as described in Section 12.2.2. This resulted in an ensemble of 100 time series of simulated evapotranspiration ET, runoff Q_s and baseflow Q_b for the year 2006, with a temporal resolution of one hour. The uncertainty in the predicted water fluxes was evaluated both at the catchment level, in which simulations were averaged over the entire study area, and at the local level, in which the spatial variability of the simulated water fluxes was preserved.

Area-averaged effect

The ensemble of monthly-averaged model outputs is shown in Figure 12.4. Despite large uncertainties in the land cover classification (*cfr.* Section 12.3.1), the spread in the simulated water fluxes is (very) small. The effect of classification uncertainty on the area-averaged annual model predictions is summarized in Table 12.4 by means of an uncertainty ratio:

$$UR_{\text{classif}} = \frac{\sum_{t=1}^{N} \left(y^{5\%}(t) - y^{95\%}(t) \right)}{\sum_{t=1}^{N} y^{\text{LCM-RF}_{\text{hard}}}(t)} \tag{12.2}$$

with t the time step index of the simulation, N the total number of time steps in the simulation period, $y^{q\%}(t)$ the model output on time step t at the q-th quantile. The value of the uncertainty ratio can be interpreted as the width of the 90% uncertainty interval (*cfr.* Section 3.2.1) normalized by the time-integrated model output from LCM-RF$_{\text{hard}}$, over a given time interval. The maximal uncertainty in the model output that was introduced by the classification uncertainty is 2.3% and was registered for the simulation of evapotranspiration and baseflow. The uncertainty ratio was the lowest for the simulation of runoff with a value of 1.6%. The prediction obtained with the hard classified land cover map falls within the range of output values as defined by the ensemble of MC simulations and shows a high correlation ($r = 0.98$) with the ensemble average. This implies that, considering the limited uncertainty in the hydrological model output due to land cover classification, an UA through MC simulation is not required, instead, the hydrological model can be applied with the hard classified land cover map to estimate the area-averaged water fluxes.

To analyse the temporal effect of classification uncertainty, UR_{classif} was calculated on a monthly time-basis. Figure 12.5 shows that the model output uncertainties can raise up to 6%, depending on the season and the water flux for which a simulation is intended. For the prediction of runoff and baseflow, large uncertainties in autumn are compensated by low uncertainties in winter and spring. These results are

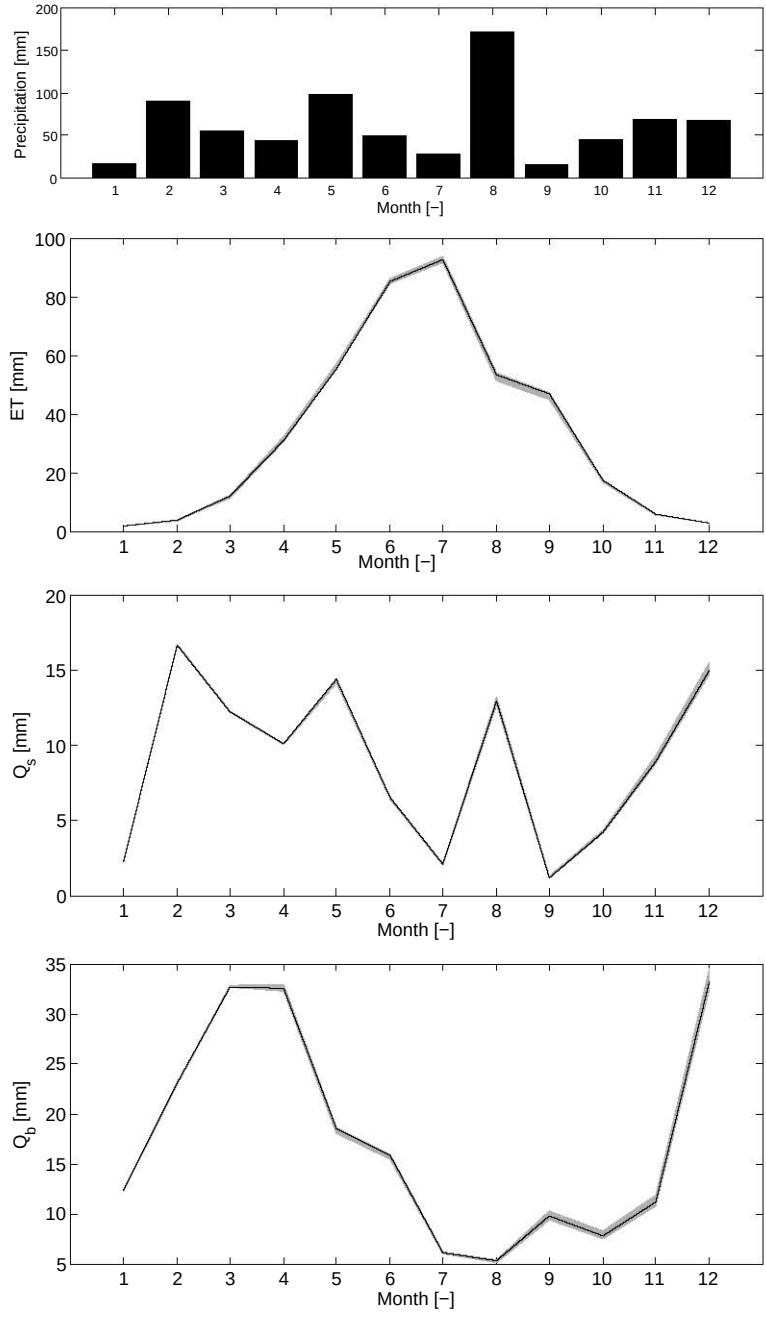


Figure 12.4: Spaghetti plot of the simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b for the Bellebeek catchment during 2006, as resulted from 100 model applications with error-corrupted versions of LCM-RF. The black line indicates the result obtained with LCM-RF$_{\mathrm{hard}}$.

Table 12.4: Evaluation of uncertainty in the water flux predictions due to uncertainty in the classified land cover map.

	5%	50%	95%	LCM-RF$_{\text{hard}}$	UR_{classif}
ET [mm]	403.6	108.6	412.9	409.5	2.3%
Q_s [mm]	105.6	106.4	107.4	107.4	1.6%
Q_b [mm]	206.3	208.5	211.2	214.2	2.3%

consistent with the main source of potential uncertainty, namely confusion between grasses and pasture (*cfr.* Section 12.3.1). The main difference between these land covers concerns the management practice: grazing for pasture and annual mowing (in September) for grasses. After mowing, the leaf area index of the grasses is dramatically reduced and confusing pasture with grasses will have a larger impact on the water flow than in summer when both vegetation types show a high similarity. This trend is not observed in the effect on evapotranspiration, for which a minimal prediction uncertainty was recorded during summer. As the weather circumstances form a boundary condition for the evapotranspiration rate, the effect of land cover confusion is weaker when water availability is limited.

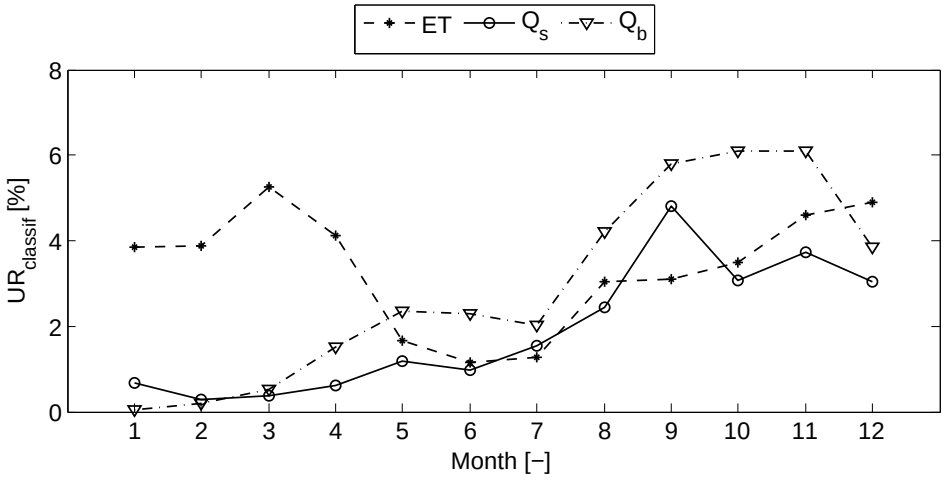
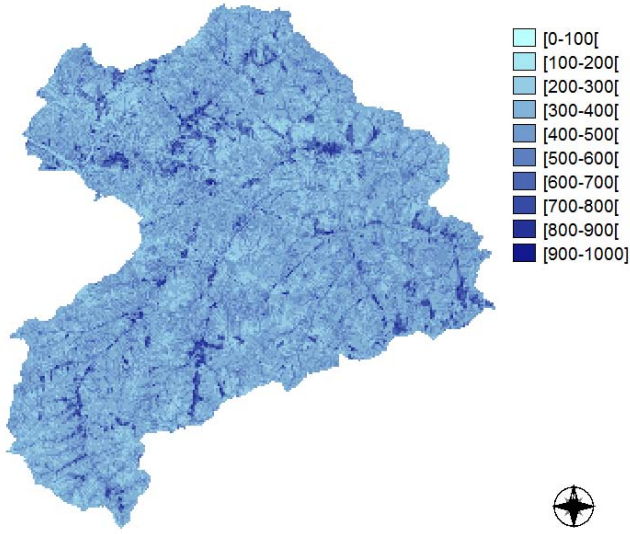


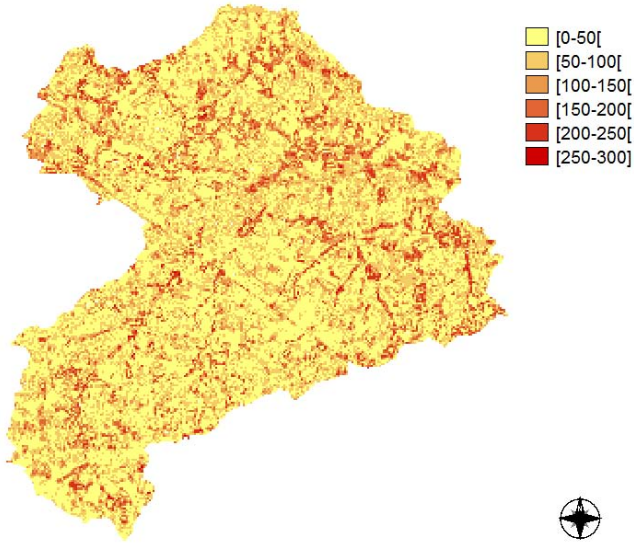
Figure 12.5: Temporal variability of uncertainty in the simulated evapotranspiration ET, surface flow Q_s and baseflow Q_b (for the Bellebeek catchment during the year 2006) due to classification uncertainty.

Location-specific effect

In order to constrain the uncertainty in the predicted water fluxes for future simulations, it is useful to identify the type and location of erroneous land cover classifications that have the highest impact such that the quality of the classified land cover map can be efficiently improved. The analysis of the local effect is restricted to evapotranspiration, as this model output is more frequently the object



(a)



(b)

Figure 12.6: Mean [mm] (a) and standard deviation [mm] (b) of the simulated annual evapotranspiration after propagation of the classification uncertainty.

of spatially distributed simulations.

Figure 12.6 shows a map of the mean simulated *ET* after uncertainty propagation together with a map of the standard deviation on the prediction. For 95% of the pixels, the mean *ET* falls within the range [200-600 mm[with a standard deviation between 0 and 200 mm. High prediction uncertainties (errors exceeding 200 mm) are associated with high values of the mean *ET* (values exceeding 600 mm), and therefore occur (i) at land covers with a high evapotranspiration rate, e.g. broad-leaved and needle-leaved forest (*cfr.* Figure 12.3), and (ii) close to the river network. In contrast with the low area-averaged effect of classification uncertainty, the local error on the simulated *ET* can raise up to 300 mm. This local effect is a result of land cover confusion by the RF classifier and of the deviation in simulated *ET* caused by this confusion. The uncertainty in the hydrological model predictions can be reduced through additional collection of reference land cover in forested areas and close to the river network.

12.3.3. Fast indicative uncertainty assessment

Above results point out that propagation of the classification uncertainty is important when a spatial distribution of the water flux predictions is intended because of the potentially high local effect of classification uncertainty on the hydrological model predictions. However, the Monte Carlo-based uncertainty assessment (MC-UA) as described in Section 12.2.2 is computationally highly demanding such that it would be beneficial in terms of computation time to have an alternative method that is able to produce estimates (further called proxies) of the Monte Carlo-based result (based on a high number of model evaluations) in terms of the mean water fluxes and the associated uncertainty based on a single model evaluation. Two such methods, further called fast indicative uncertainty assessment (FI-UA) methods, were developed: (i) a FI-UA based on the entropy of the land cover classification and (ii) a FI-UA based on uncertainty proxies. In the former FI-UA method, the mean and standard deviation of the MC-based *ET* predictions are approximated by respectively the predicted *ET* with LCM-RF$_{\text{hard}}$ and the entropy in the RF classification. The latter FI-UA method uses probabilistic reference *ET*-fluxes (derived from the model evaluation with LCM-RF$_{\text{hard}}$) in combination with the class membership probabilities (derived from the soft classified LCM-RF) to derive proxies for the mean and standard deviation of the MC-based *ET* predictions.

FI-UA based on classification entropy

In this method, the hydrological model was applied with LCM-RF$_{\text{hard}}$ as input. The simulated *ET* was evaluated for its ability to estimate the mean *ET* from the MC-UA. Although LCM-RF$_{\text{hard}}$ was shown to produce good estimates of the area-averaged water fluxes, it was not always effective to represent the spatial variability

of the simulated water fluxes. This is indicated by the low correspondence between the simulated evapotranspiration with LCM-RF$_{\text{hard}}$, further referred to as ET_{RF}, and the mean simulated evapotranspiration from the MC-UA, further referred to as ET_{MC}, especially for evapotranspiration rates higher than 600 mm (Figure 12.7). The standard deviation on the simulated ET was estimated based on a regression between the classification entropy and the standard deviation from the MC-UA. Although both variables should be logically related, the correlation coefficient was low ($r = 0.78$) (Figure 12.7). This is due to the fact that the entropy only gives information on the magnitude of confusion between two or more land covers, and not on the impact of the confusion on the hydrological model response. It is clear that this method for FI-UA is not very effective to replace the MC-UA because both H and ET_{RF} were insufficiently correlated with respectively the mean and standard deviation of the Monte Carlo-based water flux predictions.

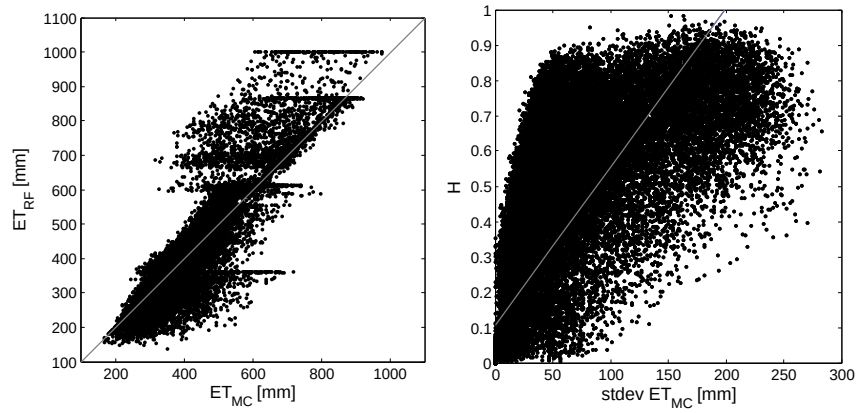
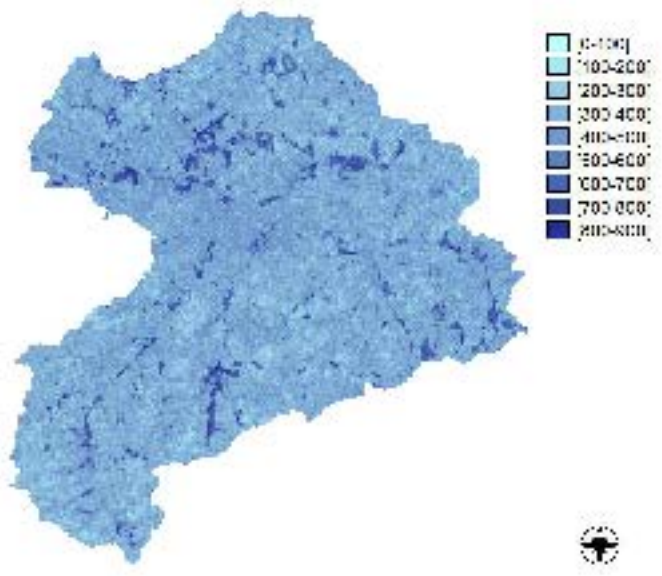


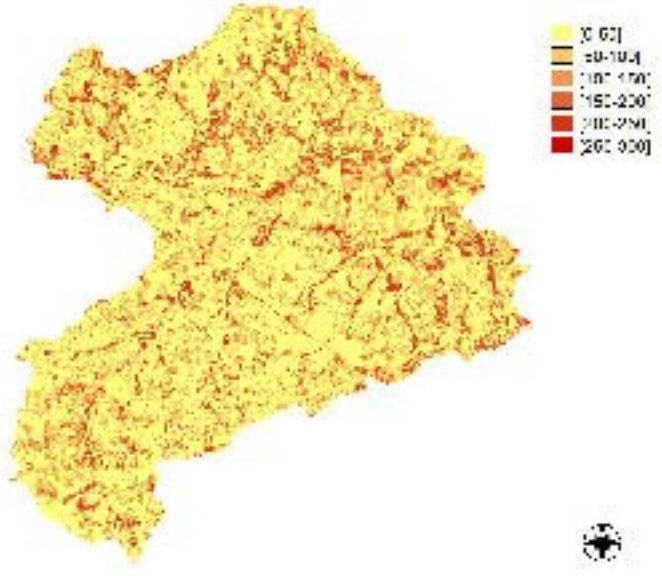
Figure 12.7: Evaluation of the predicted ET with the hard classified LCM (ET_{RF}) and the entropy H in the land cover classification as an indicator for respectively the mean (left) and the standard deviation (right) on ET from the Monte Carlo simulation (ET_{MC}).

Probabilistic reference evapotranspiration

The FI-UA method based on uncertainty proxies requires that a reference ET for each land cover class is defined. The latter can be derived from ET_{RF}. From the results of the MC-UA (Section 12.3.2), it followed that variations in the predicted ET (*cfr.* Figure 12.6) can be explained by variations in the land cover characteristics and by variations in the distance to the river network (or variations in the topography). Therefore, it was chosen to define the reference ET, further referred to as ET_{ref}, as a function of the topographic index λ_T [-] (for a definition we refer to Section 6.1.3). Figure 12.9 shows that an increase in λ_T results in an increase in ET_{RF} until a maximal evapotranspiration rate is reached: for values of λ_T between 5 and 10, the ET_{RF} linearly increases with increasing values of the topindex, whereas for values of λ_T higher than 10, ET_{RF} is constant for pixels falling into the same land cover class. The specified relationship between ET_{RF} and



(a)



(b)

Figure 12.8: Mean (left) and standard deviation (right) of the annual evapotranspiration [mm] as resulted from the fast indicative uncertainty assessment based on uncertainty proxies.

λ_T implies that evapotranspiration on wet pixels is higher than on dry pixels. Six topindex classes ($n_{\text{top}} = 6$) were defined (given in Table 12.5), for which ET_{ref} was calculated for land cover j as the average of ET_{RF} over all pixels with land cover j falling into the respective topindex classes (Table 12.5). The general procedure to determine the reference evapotranspiration rates is given in Algorithm 4.

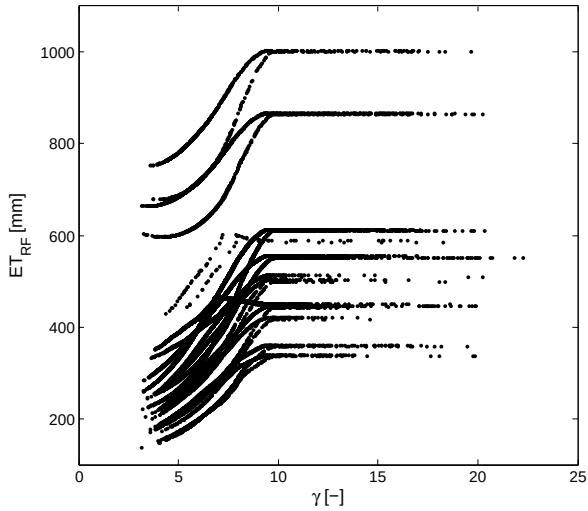


Figure 12.9: Predicted evapotranspiration (ET_{RF}) with the hard classified land cover map as a function of the topindex λ_T.

Algorithm 4: Calculate the reference evapotranspiration

Data: hard classified map LCM-RF$_{\text{hard}}$, including n_{class} classes
Result: set of reference values ET_{ref}

```
begin
    simulate ET_RF with TOPLATS using LCM-RF_hard;
    define n_top topindex classes;
    for each land cover class j ∈ {1, 2, .., n_class} do
        for each topindex class k ∈ {1, 2, .., n_top} do
            calculate ET_ref(j, k) as the mean of ET_RF(u) for all pixels u
            belonging to land cover class j and topindex class k
        end
    end
end
```

FI-UA based on uncertainty proxies

The set of reference values ET_{ref} is combined with the uncertainty in the land cover classification, as given by the uncertainty model $\mathbf{p_u}$ (with $u \in \{1, 2, ..., n_{\text{pix}}\}$), following the procedure as given in Algorithm 5 in order to calculate a proxy for the mean and standard deviation of ET_{MC} for each pixel u. As such, in constrast

Algorithm 5: Fast indicative uncertainty assessment

Data: - uncertainty model $\mathbf{p_u}$ with $u \in \{1, 2, ..., n_{\text{pix}}\}$;
- set of reference values ET_{ref};

Result: - map of proxies for mean of ET_{MC};
- map of proxies for standard deviation on ET_{MC};

begin
  for *each pixel* $u \in \{1, 2, .., n_{\text{pix}}\}$ **do**
    identify the topindex class j;
    generate a series $ET_u(r)$ with $r \in \{1, 2, ..., n_{\text{MC}}\}$ in which $ET_{\text{ref}}(j, k)$ (as can be found in Table 12.5) has a frequency of $p_u(j) \cdot n_{\text{MC}}$ with $j \in \{1, 2, ..., n_{\text{class}}\}$;
    calculate ET_{proxy} as the mean and standard deviation of $ET_u(r)$
  end
end

Table 12.5: Probabilistic reference evapotranspiration ET_{ref} derived from ET_{RF}, obtained after simulation with the hard classified LCM-RF_{hard}, for different user-defined topindex (λ_T) classes.

λ_T [-]	Ur	Wa	Bf	Nf	Pa	Gr	Po	Be	Ma	Wh	Ba	Oc
$< 5[$	371	442	656	747	323	300	283	224	236	196	197	276
$[5 - 6[$	394	481	675	770	361	347	323	253	271	226	223	306
$[6 - 7[$	424	515	699	799	403	396	373	291	309	260	246	348
$[7 - 8[$	442	548	737	841	438	436	408	325	338	283	267	384
$[8 - 9[$	443	553	764	866	458	458	427	341	353	297	277	405
$[9 - 10[$	442	551	773	874	464	465	432	346	358	301	280	412
> 10	441	586	800	897	467	469	433	347	358	303	281	414

with the FI-UA based on entropy, both the magnitude and the impact of land cover confusion is taken into account. The resulting maps of the uncertainty proxies (one for the mean and one for the standard deviation) are shown in Figure 12.8 and highly correspond to the maps as obtained with the Monte Carlo-based uncertainty assessment (*cfr.* Figure 12.6). This is confirmed by the scatterplots in Figure 12.10, indicating a very high correlation ($r = 0.98$) between the proxies and the MC-based results. Hence, the approach for FI-UA based on uncertainty proxies seems a better alternative than the FI-UA based on classification entropy to replace the MC-UA, as to accurately estimate the uncertainty in the model predictions, both the uncertainty in the classification and the impact of land cover confusion need to be incorporated in the FI-UA.

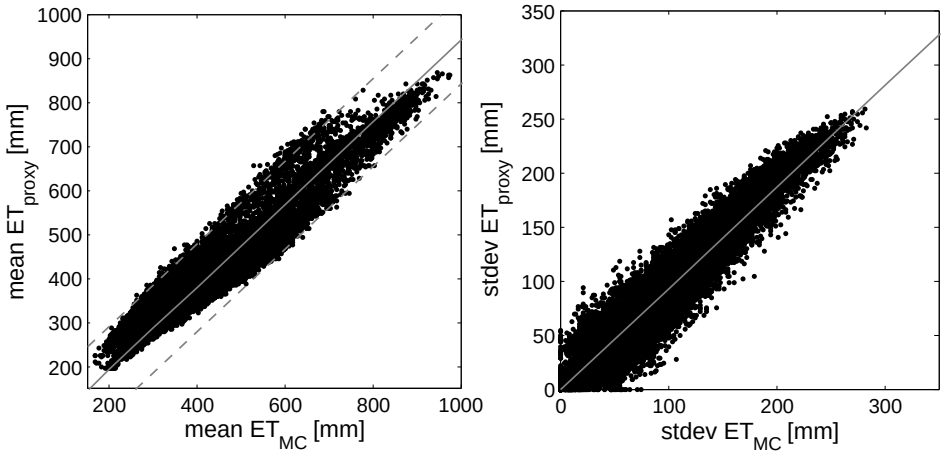


Figure 12.10: Evaluation of the FI-UA based on uncertainty proxies (ET_{proxy}) to approximate the mean (left) and standard deviation (right) of the simulated evapotranspiration (ET_{MC}) that resulted from the Monte Carlo-based uncertainty analysis (MC-UA).

12.3.4. Sensitivity of model predictions to land cover parameters

As the effect of classification uncertainty is ultimately controlled by the biophysical parameters that characterize each of the land covers (within a land cover look-up table), it is important to gain insight in the effect that small changes in the land cover parameters have on the model predictions. To that end, we performed a SA on the time-varying vegetation parameters (*cfr.* Table 12.1) following the methodology as described in Section 12.2.3.

Table 12.6 summarizes the sensitivity of the area-averaged annual evapotranspiration, runoff and baseflow rates when the monthly values of LAI, α, $z_{0,m}$, $z_{0,h}$, d_0 (*cfr.* Table 12.1) were perturbed one-at-a-time (OAT) with a factor 10^{-2}. All model outputs showed the highest sensitivity towards LAI and α, with values of S (*cfr.* Eq. (4.2)) around 0.2. This means that a change of 1% in these parameter values caused a change of 0.2% in the model output. The sensitivity index for $z_{0,m}$, $z_{0,h}$, d_0 was on average ten times lower. Runoff simulations were the least sensitive to changes in the vegetation parameters, but differences with ET and Q_s were minimal. From the results of the SA, it follows that the response in ET to perturbation of the vegetation parameters is opposite to the response in Q_s and Q_b.

The sensitivity was evaluated on a monthly time basis by calculating the sensitivity index S based on the monthly averaged model outputs. The values were ranked into five sensitivity class: (i) class I, negligible sensitivity with $0 \leq |S| < 0.05$ (ii) class II, low sensitivity with $0.05 \leq |S| < 0.20$, (iii) class III: medium sensitivity

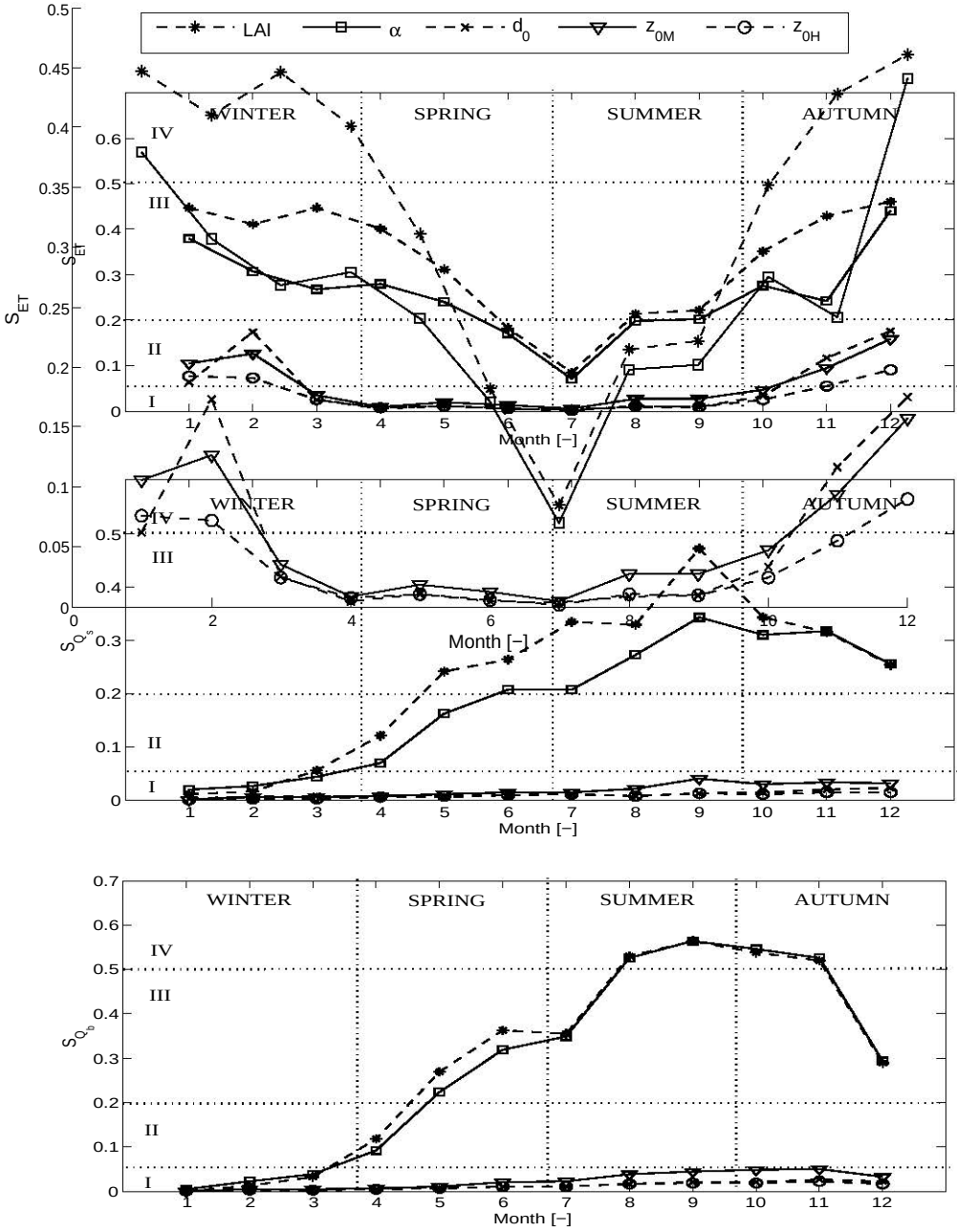


Figure 12.11: Sensitivity of the simulated evapotranspiration, runoff and baseflow to the time-varying vegetation parameters as a function of the month of the year.

Table 12.6: Sensitivity index S of the area-averaged annual evapotranspiration, runoff and baseflow rates to small changes in the monthly values of the leaf area index (LAI), albedo (α), momentum roughness ($z_{0,m}$), heat roughness ($z_{0,h}$) and zero displacement height (d_0).

Model output	LAI	α	$z_{0,m}$	$z_{0,h}$	d_0
ET [mm]	0.223	-0.182	0.019	0.009	0.011
Q_s [mm]	-0.200	0.171	-0.017	-0.008	-0.012
Q_b [mm]	-0.230	0.222	-0.020	-0.010	-0.012

with $0.20 \leq |S| < 0.50$, (iv) class IV, high sensitivity $0.50 \leq |S| < 1.00$, (v) class V, very high sensitivity $|S| \geq 1.00$. Figure 12.11 shows the evolution in model sensitivity during the simulation period (1 year). A clear difference in the sensitivity pattern between ET on the one hand and Q_s and Q_b on the other hand, is observed. The former shows a medium sensitivity (class III) to LAI and α almost all year round and a low sensitivity in July (class II). This is probably due to high evapotranspiration rates in summer that are inevitably associated with high temperatures, such that the influence of vegetative differences diminishes as water availability becomes limited. On the contrary, the sensitivity of ET to $z_{0,m}$, $z_{0,h}$ and d_0 is low to negligible all year round with peak values in winter and autumn. The prediction of Q_s and Q_b is most sensitive to LAI and α, but sensitivities gradually increase from class I in winter to class III or IV in summer. This temporal pattern can be explained by the groundwater regime: in winter the water level is constantly high and the soil becomes saturated, such that the vegetation has little influence on the generated runoff and baseflow. On the contrary, the response of Q_s and Q_b to changes in $z_{0,m}$, $z_{0,h}$, d_0 is negligible, irrespective of the time of the year.

Both the UA and SA provide us with valuable information about the sources of uncertainty w.r.t the land cover information that is most important (and hence require more attention) for the prediction of the water fluxes. This information can be used to reduce the uncertainty in the hydrological model predictions. The uncertainty map (Figure 12.6) revealed that the prediction uncertainty, amongst others, is a function of the distance to the river network. Close to the river, we expect the physiological characteristics of the crop (e.g. crop height, LAI) to be different from a location with the same crop type but more distant from the river. The physiological differences arise from differences in the soil moisture content, which is higher (lower) close to (further from) the river, and from differences in the topography (controlling the incident angle of the solar radiation), which is flatter (steeper) close to (further from) the river. Within-crop variations of the physiological characteristics induce variations in the spectral response, cause confusion in the classification process (as discussed in Section 12.3.1) and ultimately introduce uncertainty in the water flux predictions (as discussed in Section 12.3.2). However, the model parameters that describe the physiological characteristics of the

crop type (*cfr.* Table 12.1) are assumed constant within a given crop type such that within-crop variations of the plant growth are not captured in the model simulation. Incorporating knowledge about the spatial variation of the crop characteristics within the hydrological modelling would hence reduce the prediction uncertainty. The SA pointed out that the leaf area index and the albedo are the most important parameters to characterize the crop type w.r.t. the hydrological response such that including a spatial pattern of these parameters, e.g. as a remote sensing product, in the water flux predictions is expected to efficiently decrease the uncertainty in the water flux predictions.

12.4. Conclusion

A land cover map of the Bellebeek catchment was constructed through classification of multi-temporal SPOT imagery using Random Forests. The uncertainty in the classified land cover map was assessed by means of the class membership probabilities on which the entropy was calculated. Entropy values were generally low for urban areas (0.36 on average) and were the highest for needle-leaved forest and grassland (0.65 on average). The main source of uncertainty in the classified map was the confusion between pasture and grassland. In a next step, 100 realizations of the land cover map were generated following the field-based sampling approach of Canters et al. (2002) and were propagated through TOPLATS in order to simulate evapotranspiration, runoff and baseflow. The uncertainty ratio on the area-averaged water flux simulations was maximally 2.3%. This implies that propagation of the classification uncertainty is not relevant when area-averaged predictions of the water fluxes are intended, but instead water flux predictions are easily obtained from application of the hydrological model with the hard classified land cover map. However, a spatial analysis of the model output uncertainty revealed that the error on the local evapotranspiration rates can rise up to 300 mm, which implies that propagation of the classification uncertainty is advisable when a spatial distribution of the water fluxes predictions is relevant for decision-making or is coupled to other spatially distributed models.

In case propagation of the classification uncertainty is desired, a Monte Carlo-based uncertainty analysis needs to be carried out but this is computationally very demanding, especially when complex spatially distributed hydrological models are involved. Therefore, it would be of practical interest to by-pass the Monte Carlo-based uncertainty assessment (MC-UA) with a fast indicative uncertainty assessment (FI-UA) method that allows to generate proxies of the Monte Carlo-based results in terms of the mean water flux and its associated uncertainty. Two different approaches were evaluated: an FI-UA based on classification entropy was found to be inappropriate to replace the MC-UA because the classification entropy and the simulations of ET with the hard classified land cover map were poorly

correlated with the Monte Carlo-based results. This is due to the fact that the entropy only gives information on the magnitude of confusion between two or more land covers and not on the impact of confusion on the hydrologic model response. Therefore, the prediction result obtained with the hard classified land cover map was combined with the pixel-based class membership probabilities from the soft map LCM-RF to derive uncertainty proxies. Practically, application of the method implies a single evaluation of the hydrologic model with the hard classified map and the construction of a topopgraphic index map. The latter reflects the hydrologic condition for each location and explains the within-class-variability of the simulated evapotranspiration. The model prediction was therefore evaluated as a function of the topography in order to define topindex classes for which the variability in the model prediction is limited. By averaging the model predictions over the pixels with the same land cover class (as given by the hard classification result) and the same topindex class, a set of reference predictions was determined. By applying the probability distribution as given by the soft classification result to the set of reference predictions, a mean and average prediction was determined. With this FI-UA method, the mean and standard deviation on the simulated *ET* from the MC-UA were effectively approximated by their proxies which followed from a very high correlation with the results from the MC-UA. Although the proxies were not able to exactly reproduce the results of the MC-UA, they give a good indication of the effect of classification uncertainty on the prediction result and they provide a more reliable estimate of the water flux prediction than what would be obtained with the hard classified land cover map.

For future model applications, the results from the UA and SA are useful to limit the uncertainty in the water flux predictions. The UA shows that land cover confusion has the highest impact in regions close to the river network such that including a directed collection or sampling of training pixels in these regions will improve the classification result and hence the model prediction. From the SA, it followed that the LAI and the albedo play a crucial role in characterizing the land cover such that improving the quality of these parameter values will also improve the model prediction.

Part IV

Impact of uncertain soil input on hydrological model predictions

13

Introduction

In this part of the dissertation, different aspects of uncertainty modelling w.r.t. soil related inputs are investigated. Usually, the soil information in spatially distributed hydrological models is given by a soil type map and a set of soil parameters that describe the physical soil properties for each of the soil types on the soil map. The soil parameters play a crucial role in hydrological models as they directly control the movement of water and the water balance partitioning. The soil map determines the areal proportions of the different soil types in the watershed and the level of detail (as given by the soil classification scheme) at which the soil information is present.

Although most soil parameters, such as bulk density and hydraulic conductivity, can be measured in the field, hydrological model simulations are typically performed under conditions of data scarcity concerning the soil characteristics because of the time consuming and costly nature of soil data collection at the regional scale. Therefore, the soil parameters need to be determined indirectly, e.g. from textural information. Modellers often rely on relationships between the soil hydraulic parameters (SHPs) and the soil texture by using pedotransfer functions (PTFs), which translate soil related information into the hydrological parameters needed. As such, the uncertainty in the textural information propagates through the hydrological model, as shown in Figure 13.1, and induces uncertainty in the simulated soil moisture.

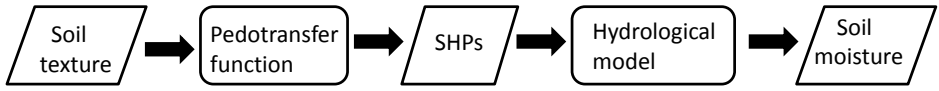


Figure 13.1: Flowchart for simulating the soil moisture in case the soil hydraulic parameters (SHPs) need to be derived from soil textural information.

For two different situations w.r.t. the availability of soil information, the uncertainty is analyzed following the schematic in Figure 13.2:

1. A soil map of the basin is readily available, but measurements of the soil texture or the soil physical parameters are lacking. As the soil type is the only source of information, the exact soil texture (in terms of the particle sizes

clay, sand and silt) is prone to uncertainty due to variability of the textural composition within a given soil type. In Chapter 14, the within-soil-class variability in the SHPs is quantified and the resulting uncertainty in the simulated soil moisture is assessed.

2. A soil map of the basin is readily available together with a set of texture measurements scarcely distributed over the basin (usually one or a few measurements per soil unit). When deriving the soil parameters from the textural information, it is assumed that the texture measurement is representative for the entire soil unit on the soil map. However, the discrepancy between the scale of measurement and the scale of model application might give rise to large uncertainties in the simulated soil moisture. In Chapter 15, a sensitivity analysis of the simulated soil moisture w.r.t. soil texture is performed in order to formulate guidelines for additional data collection such that the uncertainty can be reduced.

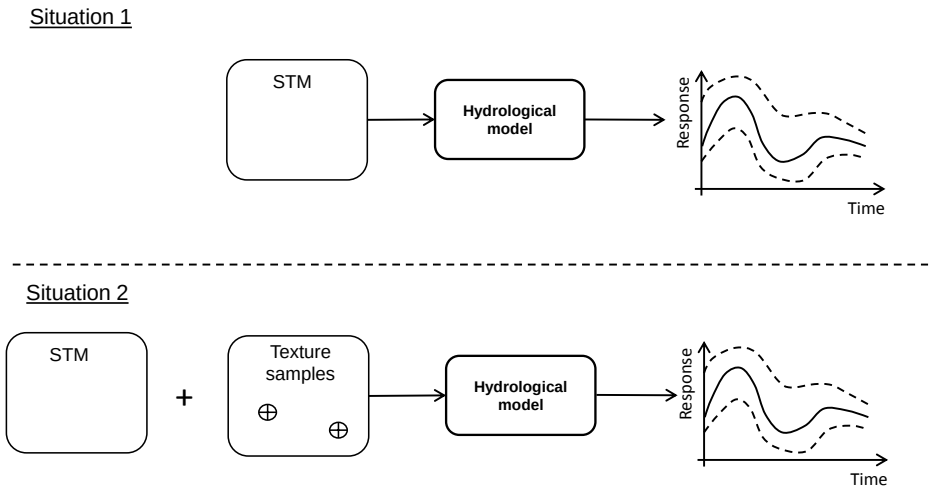


Figure 13.2: A schematic of the different uncertain situations w.r.t. soil information to model the soil moisture for a catchment. In situation 1, soil texture information is available from a soil type map, in situation 2 soil texture information is available from a soil type map and some texture measurements.

14

Impact of soil parameter uncertainty on the prediction of soil moisture

14.1. Introduction

Soil hydraulic parameters (SHPs) play a crucial role in hydrological models as they directly control the subsurface movement of water and the infiltration and runoff process. Furthermore, the SHPs indirectly control the components of the energy balance through the available soil moisture (*cfr.* description of TOPLATS in Chapter 6). There are several ways to determine the soil hydraulic parameters, which can be classified into direct and indirect methods. The direct determination of the SHPs results from laboratory and field measurements. Among the indirect methods are (i) soil class pedotransfer functions (PTFs) (e.g. Rawls et al., 1982, Cosby et al., 1984, Wösten et al., 1999), (ii) continuous PTFs (e.g. Vereecken et al., 1989, Rawls and Brakensiek, 1985, Schaap et al., 1998), and (iii) inverse methods (e.g. Hopmans et al., 2002, Pauwels et al., 2009). Each method comes with certain problems and uncertainties the modeller should deal with. In theory, direct measurement of the soil properties provides the best approximation of the true parameters, but at the regional scale this procedure is not feasible due to its time consuming and costly nature. Furthermore, the scale at which the soil hydraulic parameters are measured is generally non-compatible with the scale at which the hydrological model is applied (Grayson et al., 1992). As a consequence, good estimates instead of direct measurements may be used for many applications. Therefore, researchers have often relied on relationships between SHPs and soil texture by using PTFs, which translate soil related information into the hydrological parameters needed (Bouma, 1989). Since such relations are derived from experimental data under specific conditions, they do not take into account the underlying physical relation, and care should be taken when extrapolating these relations to other regions or ranges of data sets (Cornelis et al., 2001). Soil

This chapter is based on: Loosvelt L., Pauwels V.R.N., Cornelis W.M., De Lannoy G.J.M., Verhoest N.E.C., Impact of soil hydraulic parameter uncertainty on soil moisture modelling, Water Resources Research, 47, W03505, doi:10.1029/2010WR009204, 2011.

class PTFs link the soil class or measured soil texture to a representative set of parameter values. However, one should be aware that soil classes themselves may not represent the best possible way of classifying soils from a hydraulic point of view, as demonstrated by Twarakavi et al. (2010). Irrespective of the classification system used, 'uncertain' parameters are always obtained either due to measurement error, imperfections in the predictive function or variability in texture. The substantial impact of textural variability on the estimation of various soil hydraulic parameters has been shown in the past (e.g. Twarakavi et al., 2009) and is therefore important to take into account in hydrological studies, since ignoring this uncertainty may compromise hydrological modelling by not consistently representing the system behaviour. Former research already dealt with textural variability within soil classes (e.g. Gutmann and Small, 2005, Webb and Lilburne, 2005, Vachaud and Chen, 2002, Bormann, 2008), measurement uncertainty (e.g. Minasny and McBratney, 2002) and PTF uncertainty (e.g. Tietje and Hennings, 1996, Finke and Wösten, 1996) in hydrological modelling.

In practice, simulations are often performed in basins where soil information is limited to soil type maps and knowledge on the exact textural composition is lacking. This forces modellers to rely on class PTFs for estimating the SHPs. Unfortunately, this approach imposes substantial problems with respect to the simulation results: (i) the uncertainty in the SHPs, due to the within-soil-class variability, is not incorporated and (ii) non-region specific PTFs may be used, resulting in 'unreliable' SHPs. In this chapter, both issues are evaluated with respect to the modelled uncertainty in which the main objectives are to assess the uncertainty in the SHPs for simulation in basins with unknown textural composition and to investigate the impact of SHP uncertainty on the model prediction uncertainties. An additional objective is to investigate whether the central point of a soil class can be used to generate a representative SHP set. In the method presented, the SHPs are described by discrete probability distributions obtained by repeatedly applying PTFs to different textural compositions within a soil class. The parameter range within one soil class is then applied in a hydrological model (a description of the model set-up is given in Section 14.2.1), and yields a discrete probability distribution of the simulated soil moisture by means of ensemble predicting. This method differs from former research in that: (i) the construction of the SHP uncertainty does not rely on experimental data but on SHPs predicted with continuous PTFs for different textures within a soil class (the methodology for SHP prediction is described in Section 14.2.2), (ii) all moments of the SHP probability distribution function are determined and (iii) the dependence between the different SHPs is taken into account. In this chapter, the following questions are addressed:

- Do region specific and non-region specific PTFs generate similar SHPs and corresponding uncertainties? (discussed in Section 14.3.1)

- How are the discrete SHP distributions translated into soil moisture uncer-

tainty within a soil class? (discussed in Section 14.3.2)

- What are the implications of the simulated soil moisture distributions for data assimilation? (discussed in Section 14.3.3)
- Can the central texture serve as a proxy for the soil class to generate a best guess soil parameter set? (discussed in Section 14.3.4)

Much attention will be given to the shape of the parameter probability distribution function (PDF), since this was already shown to have a significant influence on the model results (e.g. Benke et al., 2008). It should however be stressed that this chapter only focuses on the uncertainty arising from textural variability within a soil class. The impact of all other sources of uncertainty w.r.t. soil information, including for e.g. uncertainty in the soil map and the intrinsic uncertainty in the PTFs, is left out of consideration such that it is assumed that the input uncertainty dominates the other sources of uncertainty. This statement is supported by the results of Deng et al. (2009) who reported that the PTF intrinsic uncertainty is less important than the input uncertainty and ranges between 2 and 40% (expressed as the coefficient of variation), depending on the soil parameter under consideration. In the context of simulations in basins with unknown textural composition, it would be useful for modellers to be able to make a quick assessment of the mean system behaviour within a given soil class based on one simple representative SHP set, which can be considered as the 'best guess' of the parameter values. Usually, 'average' values of the SHPs are taken from literature and are considered to be representative for the entire soil class, without considering the geographical origin of those data. The central point of a soil class could therefore serve as a more 'reliable' proxy to determine average SHP values for a given soil class, since this allows the use of continuous PTFs.

14.2. Methodology

14.2.1. Hydrological model set-up

TOPLATS was run at the point scale on coordinates 50.89 °N and 4.09 °E (located in the Bellebeek catchment near the meteorological station of Liedekerke (*cfr.* Figure 7.1), for which bare soil conditions were assumed. For this location, the soil corresponds to the USDA class loam (L) (*cfr.* Table 7.1). The texture, however, was not fixed but was varied within the simulation experiment. The model was forced with the climatic data for the period January 1, 2006 to December 31, 2006 (*cfr.* Section 7.3). The soil parameters were derived from the soil texture following the methodology described in Section 14.2.2, the vegetation parameters were attributed literature-based values as described in Section 8.2. For the TOPMODEL and routing parameters, the calibrated values from Tables 8.1 and 8.3 were used. The

simulated soil moisture content for the upper soil layer (5 cm) was validated with in situ soil moisture measurements taken at the simulation point at 2.5 cm depth between May 13 and May 30, 2007 (not shown) (*cfr.* Section 7.4).

14.2.2. Prediction of the soil hydraulic parameters

Numerous pedotransfer functions for predicting the SHPs have been proposed, reviewed and evaluated over the last decade (e.g. Tietje and Tapkenhinrichs, 1993, Wagner et al., 2001, Nemes et al., 2009). However, when relying on PTFs, modellers should be aware of the accuracy and reliability of the PTF used. The accuracy of a PTF is assessed as the correspondence between measured and predicted data for the data set from which a PTF has been developed. This is also referenced as the intrinsic uncertainty of the PTF and has already been shown to result in a substantial SHP prediction error (e.g. McBratney et al., 2002, Soet and Stricker, 2003). The reliability of a PTF can be defined as the correspondence between measured and predicted data for the data set other than the one used to develop the PTF. For a detailed discussion on both characteristics, we refer to Wösten et al. (2001). Incorporation of the intrinsic uncertainty on the PTF was beyond the scope of this study. Nevertheless, the impact of the reliability of a PTF on the SHP uncertainty was evaluated by comparing two PTFs derived from data sets with a different geographical origin. On one hand, the PTFs of Rawls and Brakensiek (1985, 1989), derived from 5320 samples of 1323 soils across the U.S.A., were chosen because they are directly related to the SMRC and the HCC in the hydrological model (*cfr.* Section 6.1.4) and are very frequently used. On the other hand, the PTFs of Vereecken et al. (1989, 1990), derived from a data set of 182 horizons of 40 Belgian soils, were selected since in this study we apply the hydrological model to Belgian weather and soil conditions (*cfr.* Chapter 7). Considering the similarity in soil and climatic features of the region of PTF development and the region of PTF application, the Vereecken et al. (1989, 1990) PTFs were assumed to be more reliable for western European soils as was already demonstrated by Cornelis et al. (2001) and Tietje and Tapkenhinrichs (1993). In this study, the PTFs of Rawls and Brakensiek (1985, 1989) and the PTFs of Vereecken et al. (1989, 1990) can therefore be referred to as respectively non-region specific PTFs and region specific PTFs.

Region specific PTFs

The Vereecken et al. (1989, 1990) regression equations, further referred to as PTF A, were used to predict a [m^{-1}], n [–], θ_s [$m^3 \cdot m^{-3}$] and θ_r [$m^3 \cdot m^{-3}$], which are the parameters of the van Genuchten moisture retention curve (van Genuchten, 1980), and the saturated hydraulic conductivity K_s [$m \cdot s^{-1}$] (Table 14.1). Inputs for the PTFs are the sand content Z [%], clay content C [%], bulk density ρ_b [$g \cdot cm^{-3}$] and organic carbon content OC [%]. Since no measurements were available, the organic carbon content was given a, for Belgian soils, realistic and constant value

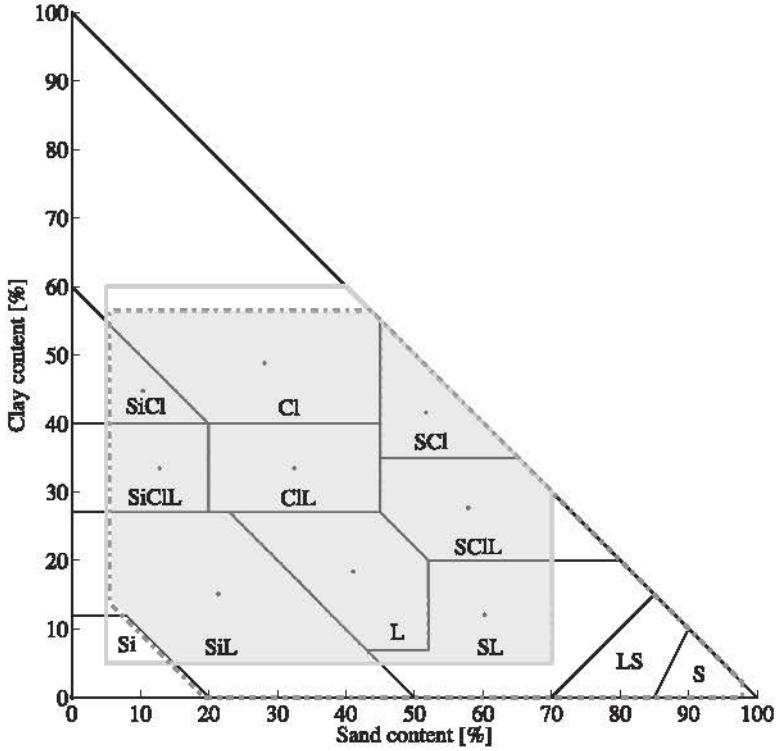


Figure 14.1: USDA texture triangle with indication of the boundary conditions on the PTFs of Rawls and Brakensiek (1985, 1989) (grey full line) and on the PTFs of Vereecken et al. (1989, 1990) (grey dashed line). The area of the particle size distributions considered in the study is shaded and the central point per soil class is indicated with a grey dot.

Table 14.1: Model parameters and regression equations of the region specific (PTF A) and non-region specific pedotransfer functions (PTF B).

PTF A: Vereecken et al. (1989, 1990)		
θ_r	=	$0.015 + 0.005 \cdot C + 0.014 \cdot OC$
θ_s	=	$0.810 - 0.283 \cdot \rho_b + 0.001 \cdot C$
a	=	$exp\,(-2.486 + 0.025 \cdot Z - 0.352 \cdot OC - 2.617 \cdot \rho_b - 0.023 \cdot C)$
n	=	$exp\left(0.053 - 0.009 \cdot Z - 0.013 \cdot C + 0.00015 \cdot Z^2\right)$
K_s	=	$1.1574 \cdot 10^{-5} \cdot exp\,(20.62 - 0.96 \cdot ln(C) - 0.66 \cdot ln(Z) - 0.46 \cdot ln(OC) - 8.43 \cdot \rho_b)$
PTF B: Rawls and Brakensiek (1985, 1989)		
ψ_c	=	$exp\left(4.34 + 0.18 \cdot C - 2.48P - 2.14 \cdot 10^{-3}C^2 - 4.36 \cdot 10^{-2}Z \cdot P - 6.17 \cdot 10^{-1}C \cdot P\right.$ $+1.44 \cdot 10^{-3}Z^2 \cdot P^2 - 8.55 \cdot 10^{-3}C^2 \cdot P^2 - 1.28 \cdot 10^{-5}Z^2 \cdot C$ $\left.+8.95 \cdot 10^{-3}C^2 \cdot P - 7.25 \cdot 10^{-4}Z^2 \cdot P + 5.4 \cdot 10^{-6}C^2 \cdot Z + 0.50 \cdot P^2 \cdot C\right)$
λ	=	$exp\left(0.78 + 1.76 \cdot 10^{-2}Z - 1.06 \cdot P - 5.3 \cdot 10^{-5}Z^2 - 2.73 \cdot 10^{-3}C^2\right.$ $+1.11 \cdot P^2 - 3.09 \cdot 10^{-2}Z \cdot P + 2.66 \cdot 10^{-4}Z^2 \cdot P^2 - 6.11 \cdot 10^{-3}C^2 \cdot P^2$ $\left.-2.35 \cdot 10^{-6}Z \cdot C + 7.99 \cdot 10^{-3}C^2 \cdot P - 6.74 \cdot 10^{-3}P^2 \cdot C\right)$
θ_s	=	$1.16 \cdot 10^{-2} - 1.47 \cdot 10^{-3}Z - 2.24 \cdot 10^{-3}C \cdot P + 0.98P + 9.87 \cdot 10^{-5}C^2 + 3.61 \cdot 10^{-3}Z \cdot P$ $-1.09 \cdot 10^{-2}C \cdot P - 0.96 \cdot 10^{-4}C^2 \cdot P - 2.44 \cdot 10^{-3}P^2 \cdot Z + 1.15 \cdot 10^{-2}P \cdot C$
θ_r	=	$-1.82 \cdot 10^{-2} + 8.73 \cdot 10^{-4}Z + 5.13 \cdot 10^{-3}C + 2.94 \cdot 10^{-2}P - 1.54 \cdot 10^{-4}C^2$ $-1.08 \cdot 10^{-3}Z \cdot P - 1.82 \cdot 10^{-4}C^2 \cdot P^2 + 3.07 \cdot 10^{-4}C^2 \cdot P - 2.36 \cdot 10^{-3}P^2 \cdot C$
K_s	=	$2.78 \cdot 10^{-6} \cdot exp\left(19.52 \cdot P - 8.97 - 2.82 \cdot 10^{-2}C + 1.81 \cdot 10^{-4}Z^2 - 9.41 \cdot^{-3} C^2 - 8.40 \cdot P^2\right.$ $+7.77 \cdot 10^{-2}Z \cdot P - 2.98 \cdot 10^{-3}Z^2 \cdot P^2 - 1.95 \cdot 10^{-2}C^2 \cdot P^2 + 1.73 \cdot 10^{-5}Z^2 \cdot C$ $\left.+2.73 \cdot 10^{-2}C^2 \cdot P + 1.43 \cdot 10^{-3}Z^2 \cdot P - 3.5 \cdot 10^{-6}C^2 \cdot Z\right)$

Notation: θ_s is the saturated soil moisture content [$m^3 \cdot m^{-3}$], θ_r the residual soil moisture content [$m^3 \cdot m^{-3}$], K_s the saturated hydraulic conductivity [$cm \cdot s^{-1}$], λ the pore size distribution [-], ψ_c the bubbling pressure [cm], a is a van Genuchten parameter and is related to the inverse of the air entry suction [cm^{-1}], n is a van Genuchten parameter and is a measure of the pore size distribution [-], ρ_b the bulk density [$g \cdot cm^{-3}$], OC the organic carbon content [%], C the clay content [%], Z the sand content [%] and P the porosity [-]

of 1.5 % (Sleutel et al., 2006). The bulk density, ρ_b, was calculated based on the sand, clay and organic matter content following the procedure as described by Saxton and Rawls (2006). With these parameters, the van Genuchten retention curve was constructed from which the Brooks and Corey parameters, ψ_c and λ, were indirectly derived through fitting the Brooks and Corey model (Brooks and Corey, 1964) to the van Genuchten retention curve.

Since a regression is only valid for the conditions for which the relation was developed, the application of PTF A is limited to textures having a clay content less then 56.46 %, a silt content less then 80.07 % and a sand content between 5.60 % and 97.80 % (Figure 14.1). One should however be aware that the real textural data limitations may even be more stricter.

Non-region specific PTFs
Rawls and Brakensiek (1985, 1989) presented one of the few multivariate PTFs for predicting the parameters of the Brooks and Corey model (*cfr.* Section 6.1.4) given the sand content Z [%], the clay content C [%] and the porosity P [-] (Table 14.1). The porosity was calculated from the bulk density, ρ_b [$\mathrm{g \cdot cm^{-3}}$], and the particle density, ρ_s [$\mathrm{g \cdot cm^{-3}}$] as follows:

$$P = 1 - \frac{\rho_b}{\rho_s} \tag{14.1}$$

The particle density, ρ_s, was corrected for the organic matter content, for which a density of $1.4 \cdot 10^{-3}\,\mathrm{g \cdot m^{-3}}$ (Kaiser and Guggenberger, 2003, Mayer et al., 2004) was assumed. The bulk density, ρ_b, was calculated following the procedure as described by Saxton and Rawls (2006). Application of the Rawls and Brakensiek (1985, 1989) PTFs is limited to textures with a clay and sand content, respectively between 5 % and 60 %, and between 5 % and 70 % (Figure 14.1). For the remainder of this chapter, the Rawls and Brakensiek model is referred to as PTF B.

Since both PTFs A and B were applied, all calculations are restricted to the overlapping validity range as shown by the shaded area in Figure 14.1. The soil class silt (Si) was excluded from the calculations as the area satisfying the limits on the validity of the PTFs was very small. Also other classes were truncated, but as a considerable portion of the class was still covered, they were included in the study and the results only apply for the areas covered in the class considered. The resulting USDA texture classes that were considered in this study, are silty clay (SiCl), clay (Cl), silty clay loam (SiClL), clay loam (ClL), sandy clay (SCl), sandy clay loam (SClL), sandy loam (SL), loam (L) and silt loam (SiL).

14.2.3. Ensemble predicting

In ensemble predicting, the model is run several times with varying parameters to end up with a set of predictions at each time step. This set of predictions is called

an ensemble. As such, ensemble prediction can be adopted as a method to estimate the uncertainty in the model prediction due to uncertainty in the model parameters. Because the ensemble contains an overwhelming amount of information that is difficult to analyze in detail, the main features are often summarized into statistics. These statistics are time-dependent and require the following assumptions: (i) a relatively large set of ensemble members is used in the calculation, and (ii) each ensemble member is an equally likely realization of the output PDF. The first moment of a PDF at time t is given by the ensemble mean:

$$\mu_t = \frac{1}{n_\mathrm{e}} \sum_{i=1}^{n_\mathrm{e}} y_i(t) \tag{14.2}$$

where n_e is the number of ensemble members and $y_i(t)$ is the model output on time t from the i-th member of the ensemble. The second moment of a PDF at time t is given by the ensemble spread:

$$\sigma_t^2 = \frac{1}{n_\mathrm{e}} \sum_{i=1}^{n_\mathrm{e}} \left(y_i(t) - \mu_t\right)^2 \tag{14.3}$$

The standardized third and fourth moment of the PDF are given by the skewness and kurtosis, respectively:

$$(Skew)_t = \frac{1}{n_\mathrm{e}} \sum_{i=1}^{n_\mathrm{e}} \left(\frac{y_i(t) - \mu_t}{\sqrt{\sigma_t^2}} \right)^3 \tag{14.4}$$

$$(Kurt)_t = \frac{1}{n_\mathrm{e}} \sum_{i=1}^{n_\mathrm{e}} \left(\frac{y_i(t) - \mu_t}{\sqrt{\sigma_t^2}} \right)^4 - 3 \tag{14.5}$$

It is assumed that large ensemble sizes were used, such that the division by n_e instead of $(n_\mathrm{e} - 1)$ in Eq. (14.3) or $(n_\mathrm{e} - 2)$ in Eqs. (14.4)-(14.5) does not result in a biased estimate of the higher moments. The skewness and kurtosis are both descriptors of the shape of a probability distribution. The former is a measure for the asymmetry, whereas the latter is a measure for the peakedness of the distribution. In case skewness is absent ($Skew = 0$), the distribution is symmetric. A positive (negative) skewness indicates the presence of a long right (left) tail, so few high (low) values are included. A distribution is said to be Gaussian or normal if both the skewness and the kurtosis (the latter with a correction of 'minus 3') are zero.

In this study, ensemble prediction is used to estimate the uncertainty in the soil moisture predictions by TOPLATS due to uncertainty in the soil hydraulic parameters. The basic assumptions are: (i) no measurements of any soil characteristics are available, and (ii) the only soil related information is a soil type map. It should be noted, however, that incorporation of the uncertainty in the soil type map was

beyond the scope of this study. The SHPs hence need to be derived from the soil class, which was defined according to the USDA classification system (*cfr.* Figure 7.2). However, each soil class contains a range of particle size distributions, given by its sand (2 - 0.05 mm), silt (0.05 - 0.002 mm) and clay (< 0.002 mm) content, such that, within a soil class, the SHPs are variable when knowledge on the exact textural composition is lacking. In order to account for this variability, the USDA texture triangle was uniformly sampled along a regular grid with a spacing of 1 %, excluding the samples that fall outside the valid area for PTF application (Figure 14.1). The resulting number of samples per soil class is given in Table 14.2.

Table 14.2: Number of synthetically generated samples in the 9 truncated USDA soil classes considered in this study.

Texture Class	Samples
SL	328
SClL	373
SCl	231
L	392
ClL	364
SiL	622
SiClL	210
SiCl	106
Cl	574

For each sampled texture within a given soil class, the soil hydraulic parameters were predicted by means of continuous pedotransfer function (*cfr.* Section 14.2.2) from which a discrete probability distribution of each SHP was constructed. The PDFs of the SHPs were used to generate a number of hydrological model forecasts (equal to the number of texture samples given in Table 14.2) by drawing a combination of parameters from the distributions, based on random sampling. The PDFs of the SHPs were not sampled individually to preserve the correlation between the SHPs. In order to assess the uncertainty in the model output, an ensemble of model runs was performed, for which each ensemble member only differed from another member through its SHPs. Such an ensemble was constructed for each selected USDA soil class (*cfr.* Section 14.2.2) by attributing the SHPs of each texture sample within the soil class (Table 14.2) to the point for which TOPLATS was run (*cfr.* Section 14.2.1). Each model run resulted in a time series of simulated daily soil moisture content, θ_{daily} [$m^3 \cdot m^{-3}$], within the upper 5 cm of the soil. On each time step, a PDF of the simulated soil moisture was derived from the ensemble of simulations and was summarized into its statistics using Eqs. (14.2)-(14.5). A general description of the ensemble prediction procedure is given in Algorithm 6.

Algorithm 6: Modelling soil moisture uncertainty through ensemble predicting

Data: soil map, indicating n_{class} soil classes
Result: uncertainty in the simulated soil moisture θ

```
begin
    sample n textures T = (C, Z, L) on the texture triangle;
    for each soil class j ∈ {1, 2, ..., n_class} do
        select the samples that fall into class j;
        for each texture T in class j do
            estimate the SHPs by means of a PTF (Table 14.1);
            simulate θ using TOPLATS (Section 14.2.1);
        end
        construct the PDF of the SHPs;
        calculate the statistics of the PDF (Eqs. (14.2)-(14.5));
        for each time step t ∈ {1, 2, ..., N} do
            construct the PDF of θ_t;
            calculate the statistics of the PDF (Eqs. (14.2)-(14.5));
        end
    end
end
```

14.3. Results and discussion

14.3.1. Within-soil-class variability of the SHPs

On the texture triangle, 5000 textures were sampled following the methodology described in Section 14.2.3. For these sampled textures, the SHPs were predicted with either PTF A or PTF B (using the equations in Table 14.1), and a discrete probability distribution for each of the SHPs was constructed per soil class and was characterized by its first four moments (mean, spread, skewness and kurtosis). The set of SHPs obtained with PTF A and PTF B are referenced as respectively, SHPs-A and SHPs-B.

Mean and standard deviation

The mean SHPs-A and SHPs-B for each soil class are given in Tables 14.3 and 14.4 together with the corresponding standard deviation, σ_{unit}, which is expressed per unit area on the texture triangle (defined as a square with a range of 1 % clay and 1 % sand) in order to avoid poor standard deviation statistics by truncating the soil class. For θ_r and ψ_c, PTF A gives consistently higher values in comparison to PTF B, while for λ the opposite is observed. For θ_s and K_s, differences between PTF A and PTF B are not consistent. Likewise, the standard deviation of the SHPs within a soil class seems to be highly dependent on the PTF used. In general,

PTF B results in larger standard deviations (*cfr.* Tables 14.3-14.4). These findings demonstrate the large disagreement in predicted SHPs between different PTFs, as was already reported by Soet and Stricker (2003), Cornelis et al. (2001) and others.

Figure 14.2 illustrates the relationship between the textural composition and the predicted values of K_s and ψ_c. It is clear that both SHPs are affected by the texture in a different way with a large difference between the results of the two PTFs. Depending on the soil class, the within-soil-class variability of K_s and ψ_c is dominated by a variation in either the sand or clay content. The sensitivity of the SHPs to those variations is indicated by the density of the contour lines. As can be seen from Figure 14.2, the sensitivity towards texture of the SHPs largely depends on the choice of the PTF. In general, an increase in the sensitivity can be observed with an increasing clay content. The differences in variability and sensitivity of the SHPs in relation to soil texture can mainly be attributed to a difference in geographical origin of the data sets used to construct the PTFs, since the origin determines the soil and climatic conditions (*cfr.* Section 14.2.2). Furthermore, the information content of both data sets may differ with respect to the number of soil samples, the variability in sampled textures, the methods used to measure the SHPs, etc. All these differences are reflected in the resulting regression equation (Table 14.1) and continue to exist in the behaviour of the SHPs in relation to the soil texture. For example, θ_s and θ_r are insensitive towards the sand content according to PTF A, whereas the same parameters vary with the sand content in PTF B. This raises questions about the general applicability of the standard PTFs (often non-region specific) as it seems that a crucial soil variable (e.g. soil structure) to allow for non-region specific application is missing in the regression.

Shape of the parameter probability density function

The skewness ($Skew$) and kurtosis ($Kurt$) were calculated for each PDF of the SHPs and are listed in Tables 14.5 and 14.6 for PTFs A and B, respectively. From these tables, it can be seen that the PDF of most SHPs does not represent a normal distribution. Assuming a normal or lognormal distribution (*i.e.* only defining the first 2 moments) for the SHPs would therefore imply a loss of information. It seems that the shape of the PDF is not only a function of the SHP itself, but that the shape also depends on the soil class and the choice of the PTF. This is illustrated in Figure 14.3, showing the PDF of K_s and ψ_c for the soil class ClL. This class is chosen because of its central position on the soil texture triangle (no extreme textures). It is clear that the PTFs result in both a different range of SHP values and a different shape of the SHP distribution. Therefore, it can be stated that the selection of the PTF is a crucial step in assessing the uncertainty in the SHPs. If region specific PTFs are available, they are preferable since they are assumed to generate more reliable SHPs than non-region specific PTFs. Furthermore, the higher-order moments are required to accurately describe the PDF of the SHP, since assuming a normal distribution may not be representative for the SHP at

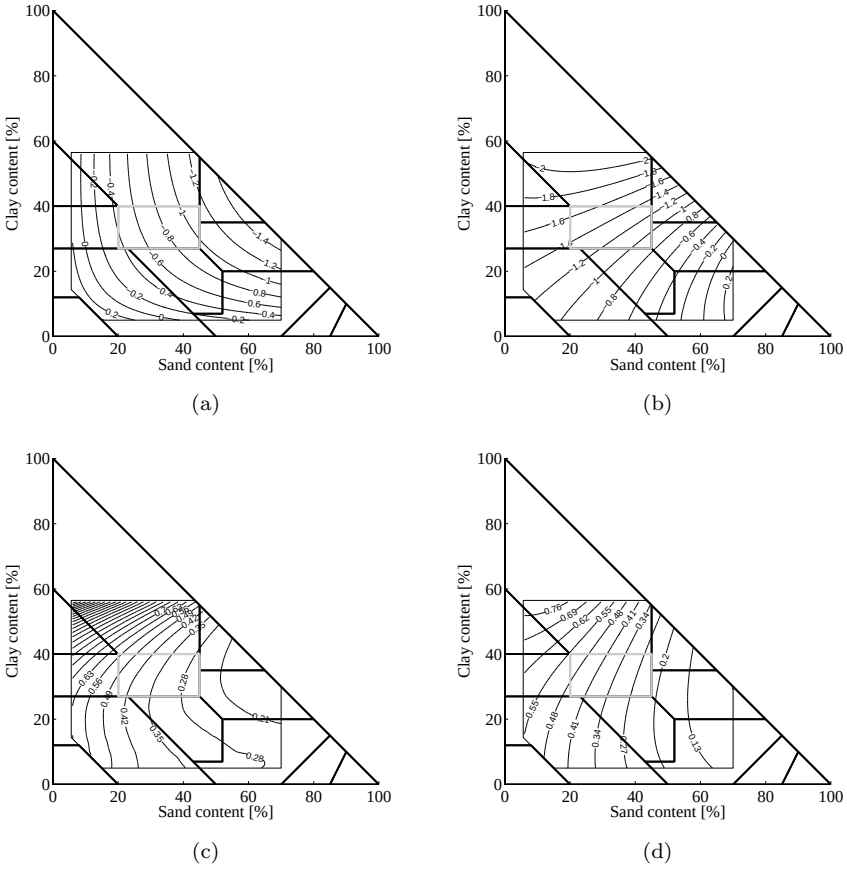
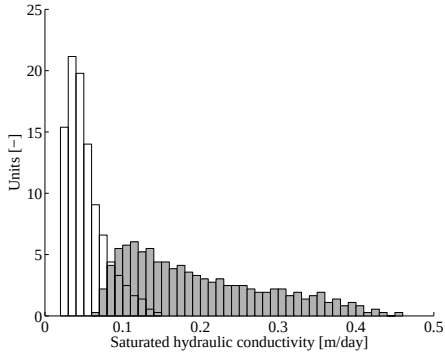
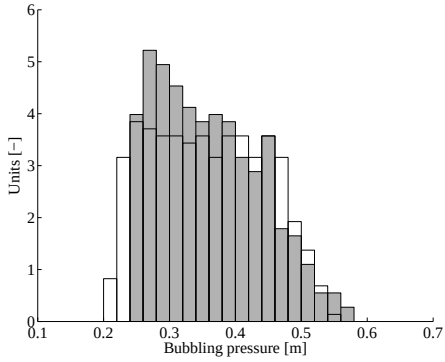


Figure 14.2: Top panel: contour plot of the logarithm ($\log_{10}$) of the saturated hydraulic conductivity K_s [$\mathrm{m} \cdot \mathrm{day}^{-1}$] predicted with PTF A (a) and PTF B (b). Bottom panel: contour plot of the bubbling pressure ψ_c [m] predicted with PTF A (c) and PTF B (d). The soil class clay loam is highlighted in grey.



(a)



(b)

Figure 14.3: Probability distribution of (a) the saturated hydraulic conductivity, K_s [$\mathrm{m} \cdot \mathrm{day}^{-1}$] and (b) the bubbling pressure ψ_c [m] predicted with PTF A (grey) and PTF B (white) for the soil class clay loam.

Table 14.3: Mean (μ) and standard deviation per unit area on the texture triangle (σ_{unit}) of the SHPs as predicted from PTF A for the 9 selected soil classes of the USDA texture triangle.

Texture	$\theta_s[m^3 \cdot m^{-3}]$		$\theta_r[\mathrm{m}^3 \cdot \mathrm{m}^{-3}]$		$\psi_c[\mathrm{m}]$		λ[-]		K_s $[\mathrm{m} \cdot \mathrm{day}^{-1}]$	
	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}
SL	0.45	1.16E-05	0.12	8.87E-05	0.27	1.13E-04	0.69	7.27E-04	0.28	7.03E-04
SClL	0.46	2.52E-05	0.20	6.56E-05	0.20	8.13E-05	0.28	1.82E-04	0.07	6.89E-05
SCl	0.48	5.59E-05	0.27	1.27E-04	0.24	3.18E-04	0.15	1.42E-04	0.05	5.84E-05
L	0.47	2.73E-05	0.15	8.04E-05	0.3	9.30E-05	0.46	4.14E-04	0.26	2.88E-04
ClL	0.49	3.44E-05	0.22	6.21E-05	0.36	2.39E-04	0.22	1.14E-04	0.20	2.81E-04
SiL	0.48	2.46E-05	0.13	5.50E-05	0.45	1.40E-04	0.57	3.02E-04	0.93	8.39E-04
SiClL	0.52	4.95E-05	0.22	1.08E-04	0.61	4.74E-04	0.25	2.15E-04	0.78	1.80E-03
SiCl	0.54	1.06E-04	0.28	1.76E-04	0.94	2.30E-03	0.17	1.94E-04	0.88	3.1E-03
Cl	0.52	3.69E-05	0.30	4.44E-05	0.73	6.66E-04	0.14	4.32E-05	0.26	4.34E-04

Table 14.4: Mean (μ) and standard deviation per unit area on the texture triangle (σ_{unit}) of the SHPs as predicted from PTF B for the 9 selected soil classes of the USDA texture triangle.

Texture	$\theta_s[\mathrm{m}^3 \cdot \mathrm{m}^{-3}]$		$\theta_r[\mathrm{m}^3 \cdot \mathrm{m}^{-3}]$		$\psi_c[\mathrm{m}]$		λ[-]		K_s $[\mathrm{m} \cdot \mathrm{day}^{-1}]$	
	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}	μ	σ_{unit}
SL	0.44	3.49E-05	0.07	6.35E-05	0.13	1.29E-04	1.79	5.54E-04	0.97	1.70E-03
SClL	0.44	4.51E-05	0.10	2.12E-05	0.13	1.28E-04	1.22	5.14E-04	0.47	1.00E-03
SCl	0.49	1.34E-04	0.11	7.32E-06	0.19	2.82E-04	0.76	6.98E-04	0.08	3.30E-04
L	0.46	3.61E-05	0.08	4.26E-05	0.24	1.67E-04	1.58	4.53E-04	0.22	3.35E-04
ClL	0.50	5.55E-05	0.10	1.71E-05	0.36	2.61E-04	1.15	3.83E-04	0.05	7.64E-05
SiL	0.49	2.84E-05	0.06	3.20E-05	0.41	1.53E-04	1.62	2.34E-04	0.10	1.07E-04
SiClL	0.54	7.89E-05	0.10	3.40E-05	0.56	2.47E-04	1.23	4.84E-04	0.03	3.29E-05
SiCl	0.58	1.89E-04	0.11	2.38E-05	0.67	5.06E-04	1.00	6.22E-04	0.01	2.48E-05
Cl	0.57	5.92E-05	0.11	3.44E-06	0.53	2.79E-04	0.78	2.43E-04	0.02	1.44E-05

Table 14.5: Skewness ($Skew$) and kurtosis ($Kurt$) of the SHPs as predicted from PTF A for the 9 selected soil classes of the USDA texture triangle.

Texture	θ_s[m^3 · m^{-3}]		θ_r[m^3 · m^{-3}]		ψ_c[m]		λ[-]		K_s [m · day^{-1}]	
	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$
SL	-0.10	-0.59	0.11	-1.27	0.12	-0.36	0.19	-1.21	1.09	0.20
SClL	0.32	-0.79	-0.05	-1.11	-0.13	-0.94	0.49	-0.82	0.38	-0.73
SCl	0.07	-0.58	0.57	-0.61	1.17	1.62	-0.23	-0.99	0.52	-0.32
L	0.64	-0.43	-0.27	-1.00	0.46	0.31	0.88	-0.17	0.97	0.84
ClL	0.18	-0.74	4.55E-14	-1.21	0.46	-0.67	0.24	-1.05	0.68	-0.53
SiL	0.30	-0.86	0.18	-1.08	0.37	-0.93	0.29	-1.08	0.96	0.19
SiClL	0.13	-0.63	3.63E-14	-1.21	0.42	-0.40	0.25	-1.04	0.78	-0.43
SiCl	0.18	-0.60	0.57	-0.62	1.00	0.74	-0.28	-0.81	0.27	-1.07
Cl	0.37	-0.63	-0.17	-1.16	1.17	0.88	0.28	-0.89	2.19	5.72

Table 14.6: Skewness ($Skew$) and kurtosis ($Kurt$) of the SHPs as predicted from PTF B for the 9 selected soil classes of the USDA texture triangle.

Texture	θ_s[m^3 · m^{-3}]		θ_r[m^3 · m^{-3}]		ψ_c[m]		λ[-]		K_s [m · day^{-1}]	
	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$	$Skew$	$Kurt$
SL	0.01	-0.55	-0.11	-1.23	0.63	0.08	-0.29	-1.10	0.62	-0.71
SClL	0.28	-0.73	-0.42	-0.85	0.42	-0.841	2.40E-03	-1.06	1.26	1.04
SCl	0.51	-0.32	-1.32	1.18	0.14	-0.68	-0.16	-0.91	1.58	2.67
L	0.80	-0.05	-0.59	-0.68	0.78	-0.06	0.20	-1.08	0.65	-0.50
ClL	0.12	-0.62	-0.43	-0.85	0.12	-1.08	-0.15	-0.97	1.21	1.20
SiL	0.31	-0.70	-0.06	-1.11	-0.08	-0.94	-0.10	-0.94	1.52	2.23
SiClL	0.12	-0.73	-0.27	-1.11	-0.14	-0.74	-0.06	-1.15	0.53	-0.49
SiCl	0.43	-0.51	0.12	-1.07	-0.03	-0.49	-0.30	-0.81	0.01	-0.99
Cl	0.10	-0.75	-0.39	0.82	0.09	-1.08	-0.31	-0.60	1.52	2.34

hand.

14.3.2. Uncertainty in the soil moisture predictions

The above constructed discrete probability distributions of the SHPs were propagated through the hydrological model following the methodology as described in Section 14.2.3. It was evaluated how the shape of the parameter PDFs was translated into the PDF of the modelled soil moisture. A comparison of the model prediction uncertainty resulting from SHPs-A and SHPs-B was carried out.

The resulting ensemble of θ_{daily} is illustrated for the classes SL (328 ensemble members) and SiCl (106 ensemble members) in Figure 14.4 with a spaghetti plot of the different ensemble members. These two classes were chosen since they represent two extremes on the texture triangle (Figure 14.1). A window of 70 days, starting from day of year (DOY) 200 (*i.e.* July 19) up to DOY 270 (*i.e.* September 27), was selected to properly visualise the behaviour of each ensemble member. From Figure 14.4, it can be seen that (i) there is a substantial impact of the within-soil-class variability of the SHPs on the modelled soil moisture and that (ii) there are large differences with respect to the predicted value, the dynamics and the range of the ensembles between PTFs A and B. For the soil class SL it is apparent that all ensemble members resulting from SHPs-B predict lower soil moisture values than those resulting from SHPs-A. Additionally, the use of SHPs-B results in a much higher density of the ensemble members, whereas SHPs-A produce a more uniform distribution of the ensemble members. On the other hand, the PTF also affects the dynamics in the predicted soil moisture, as can be seen from the spaghetti plot of SiCl (Figure 14.4). Despite the better correspondence in average simulated soil moisture, SHPs-B cause a faster wetting and drying of the soil as compared to the dynamic in θ_{daily} resulting from SHPs-A. The intensity of the fluctuations in soil moisture is related to the shape of the SMRC and the HCC (*cfr.* Section 6.1.4). For each set of SHPs in the ensemble, both curves were constructed and are shown in Figure 14.5. It is clear that the SMRC is more nonlinear when SHPs-B were used instead of SHPs-A. The difference in SMRC between the region and non-region specific PTFs is the most apparently for SiCl (mainly due to the discrepancy in θ_r, Tables 14.3-14.4).

All ensemble members were averaged in time in order to obtain one single value per PTF application, referred to as the average simulated soil moisture content θ_{avg} [$m^3 \cdot m^{-3}$], for each sample on the texture triangle. This model output was reprojected to the texture triangle to result in a contour plot as shown in Figure 14.6. In this way, the sensitivity of the model output to soil texture, as a result of the joint sensitivity of the SHPs to texture, was deduced. As can be seen from Figure 14.6, θ_{avg} resulting from SHPs-A decreases with decreasing clay content and is nearly unaffected by the sand content. When using SHPs-B, however, both

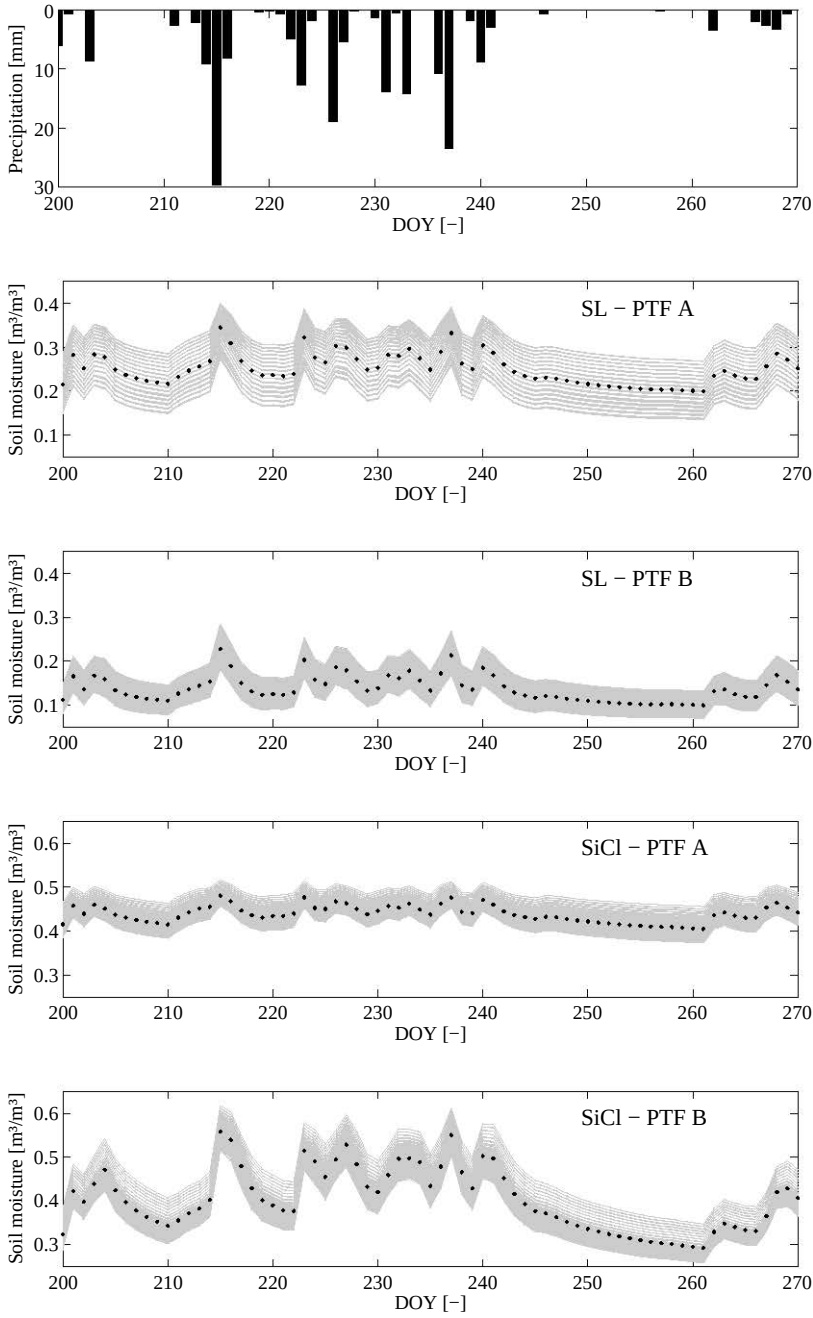


Figure 14.4: Spaghetti plots of the daily simulated soil moisture content (2006), θ_{daily} $[m^3 \cdot m^{-3}]$, for sandy loam (SL) and silty clay (SiCl) with the SHPs predicted from PTF A and PTF B. The ensemble mean is indicated with black dots.

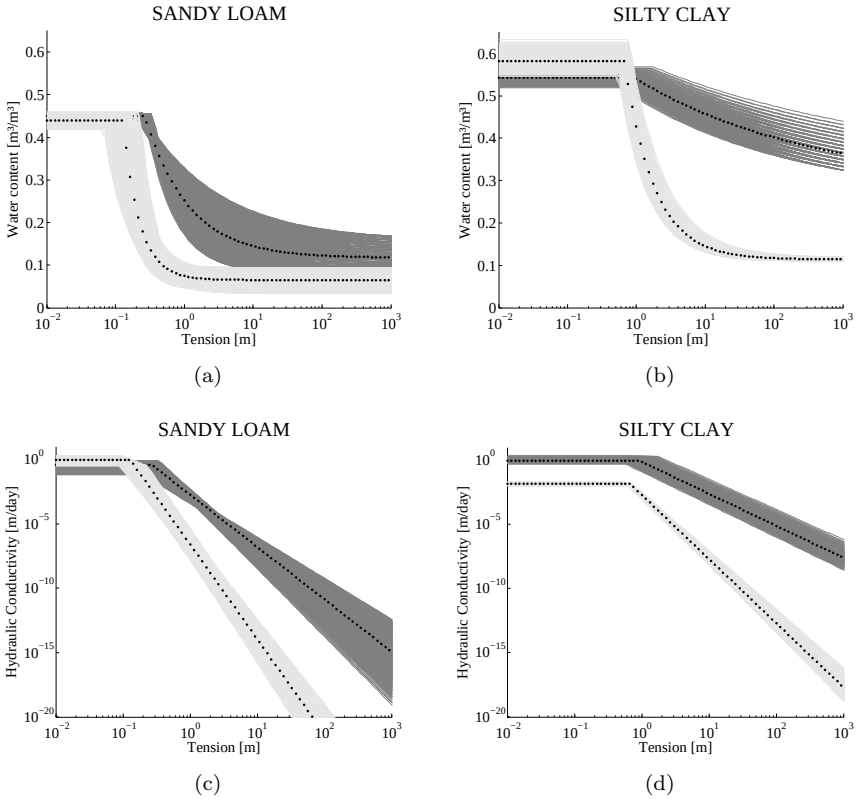


Figure 14.5: Left panel: ensemble (328 members) of soil moisture retention curves (a) and hydraulic conductivity curves (c) for the soil class sandy loam resulting from application of PTF A (dark grey) and PTF B (light grey). Right panel: ensemble (106 members) of soil moisture retention curves (b) and hydraulic conductivity curves (d) for the soil class silty clay resulting from application of PTF A (dark grey) and PTF B (light grey). The ensemble mean is indicated with black dots.

a decrease in clay content and an increase in sand content cause a reduction of θ_{avg}. From the density in contour lines, it can be derived that the within-soil-class variability of θ_{avg} increases for a decreasing clay content in case SHPs-A were used, whereas the opposite is observed when using SHPs-B. These results again reflect the empirical relationship between the SHPs and the textural composition (Table 14.1). For example, the insensitivity of θ_s and θ_r towards the sand content strongly determines the dominant influence of the clay content on the simulated soil moisture as resulted from PTF A.

The statistics of the ensembles of θ_{daily} were further analyzed through their moments. For each day t, the ensemble mean, μ_t, the ensemble spread, σ_t, the skewness, $Skew_t$, and the kurtosis, $Kurt_t$ were calculated following Eqs. (14.2)-(14.5).

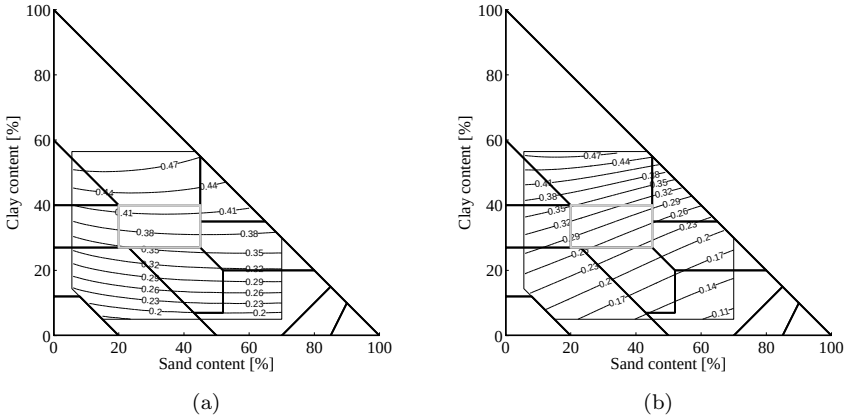


Figure 14.6: Contour plot of the average simulated soil moisture content, θ_{avg} [$m^3 \cdot m^{-3}$], using the SHPs as predicted from PTF A (a) and PTF B (b). The soil class clay loam is highlighted in grey.

Ensemble mean and spread

The temporal behaviour of the ensemble mean and the ensemble spread for θ_{daily} is illustrated for the classes SL and SiCl in Figure 14.7, which clearly demonstrates the impact of the PTF on the fluctuations in mean predicted soil moisture. As explained earlier, these differences in fluctuation can be attributed to a different shape of the SMRC (Figure 14.5). Also the evolution of the ensemble spread is affected by the choice of the PTF, but its extent depends on the soil class considered. Both the fluctuations in the ensemble mean and spread are closely related to the precipitation regime. When rainfall events occur, the water tension decreases and drops below the bubbling pressure after a while. When $\psi = \psi_c$, a turning point is clearly visible in both the SMRC and HCC (Figure 14.5). At this point, the soil moisture equals θ_s and the hydraulic conductivity is K_s. If $\psi < \psi_c$, there is a substantially lower spread on the HCC as compared to the zone where $\psi > \psi_c$. A rainfall event is therefore generally accompanied with a decrease in ensemble

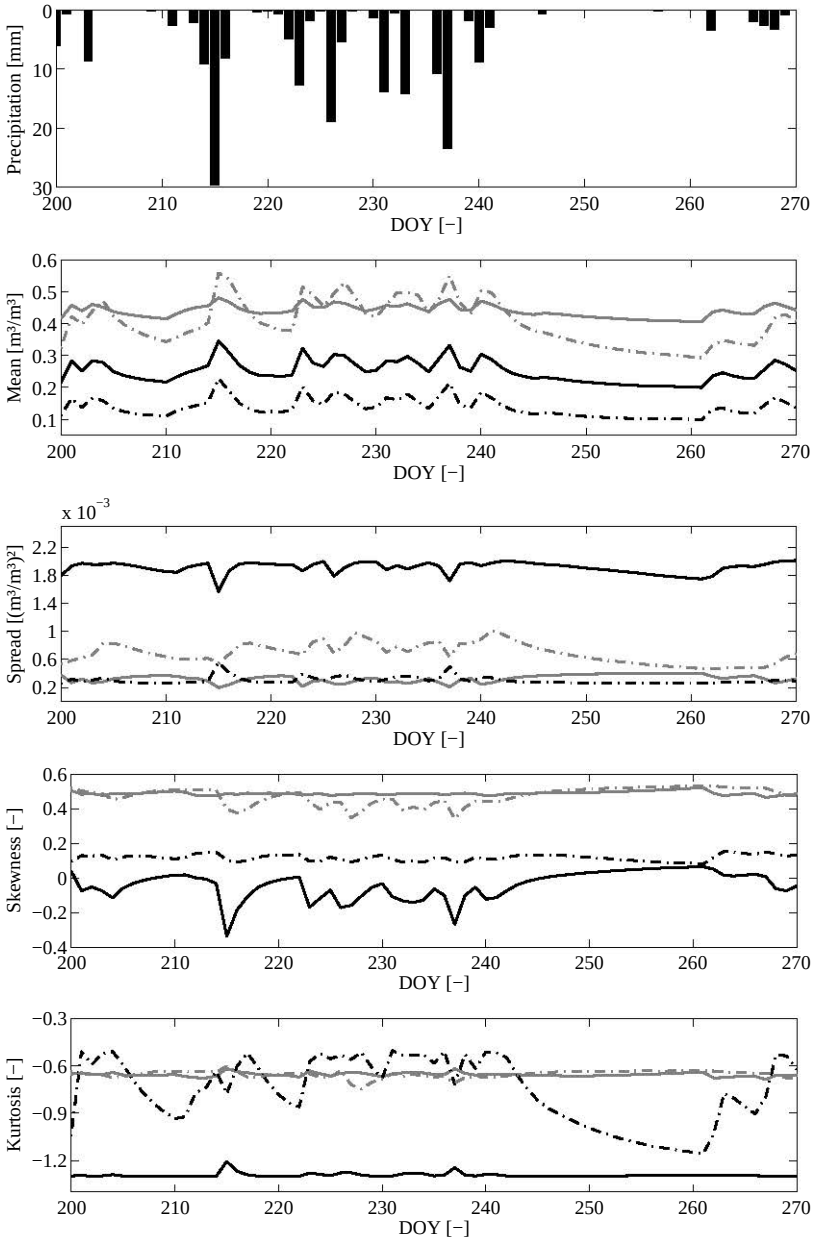


Figure 14.7: Time series of the ensemble mean, ensemble spread, ensemble skewness and ensemble kurtosis for the simulated soil moisture content (2006), θ_{daily} $[\mathrm{m}^3 \cdot \mathrm{m}^{-3}]$. The results are given for the soil classes sandy loam (black) and silty clay (grey) using the SHPs as predicted from PTF A (solid line) and PTF B (dotted line).

spread.

For θ_{avg}, the mean and standard deviation per unit area on the texture triangle, were calculated for each soil class and are listed in Table 14.7. This table shows that θ_{avg} obtained with SHPs-A is consistently higher than when SHPs-B were used. Nevertheless, the difference is the smallest for the soil classes with a low sand content (SiL, SiClL, SiCl and Cl). In general, the standard deviation on θ_{avg} is slightly higher when the SHPs were predicted with PTF B. This could be expected because of the higher standard deviation associated with SHPs-B as compared to SHPs-A (Section 14.3.1). The consistent underestimation of the soil moisture obtained with the non-region specific PTF (PTF B) as compared to the region specific PTF (PTF A), indicates that this is not purely by chance. It can therefore be argued that there is a fundamental difference between the region and non-region specific PTFs.

Ensemble skewness and kurtosis

As already mentioned in Section 14.3.1, the skewness and kurtosis characterize the shape of a distribution. From the evolution of both statistics in time for θ_{daily} (Figure 14.7), it can be seen that often there is a large variation in the shape of the PDF through time, but that the evolution depends on the PTF used and on the soil class considered. In general, the distribution strongly deviates from normality with either a too flat or too skewed distribution. The evolution in the shape of the PDF can be linked to the precipitation regime, but the fluctuations are not always as strong as compared to those in the ensemble mean. For the SiCl soil, the skewness and kurtosis are more stable than for the SL soil and implies a steadier evolution of the soil moisture distribution. A possible explanation for this can be found in the shape of the HCC. The latter is less nonlinear for SiCl than for SL (Figure 14.5).

The skewness and kurtosis for θ_{avg} are given in Table 14.7, for both the model run with SHPs-A and SHPs-B. It can be seen that the distribution of θ_{avg} tends to result in a positively skewed and flattened distribution when using SHPs-B, whereas for the result obtained with SHPs-A, the distribution is either positively or negatively skewed depending on the soil class. Figure 14.8 shows the discrete distribution of θ_{avg} for the soil class ClL (highlighted in Figure 14.6). This class is chosen because of its central position on the texture triangle. It can be seen that the choice of the PTF for predicting the SHPs has a substantial impact (i) on the absolute value of the predicted soil moisture, (ii) on the range in simulated soil moisture and (iii) on the shape of the soil moisture distribution. Selecting the most reliable PTF is therefore a crucial step when assessing the uncertainty in the simulated soil moisture since differences between region and non-region specific PTFs remain after propagating the SHP uncertainty through the hydrological model.

Table 14.7: Mean (μ), standard deviation per unit area on the texture triangle (σ_{unit}), skewness ($Skew$) and kurtosis ($Kurt$) of the average simulated soil moisture content θ_{avg} (within the upper 5 cm of the soil during the year 2006) for the 9 selected USDA classes using the SHPs as predicted from PTF A and B.

Texture	θ_{avg} [$m^3 \cdot m^{-3}$] with SHPs from PTF A				θ_{avg} [$m^3 \cdot m^{-3}$] with SHPs from PTF B			
	μ	σ_{unit}	$Skew$	$Kurt$	μ	σ_{unit}	$Skew$	$Kurt$
SL	0.25	1.644E-04	-0.03	-1.30	0.13	6.24E-05	0.13	-0.68
SClL	0.36	6.96E-05	-0.26	-1.02	0.20	7.79E-05	0.25	-0.78
SCl	0.42	9.53E-05	0.48	-0.72	0.30	2.51E-04	0.71	-0.22
L	0.30	1.22E-04	-0.52	-0.77	0.19	9.31E-05	0.16	-0.88
ClL	0.39	6.48E-05	-0.14	-1.15	0.28	9.26E-05	0.07	-0.73
SiL	0.25	8.54E-05	0.05	-1.11	0.21	6.83E-05	0.07	-0.93
SiClL	0.38	1.33E-04	-0.12	-1.10	0.32	1.32E-04	0.02	-1.03
SiCl	0.44	1.76E-04	0.49	-0.66	0.40	2.53E-04	0.48	-0.65
Cl	0.46	3.79E-05	-0.10	-1.05	0.40	8.63E-05	-0.22	-1.01

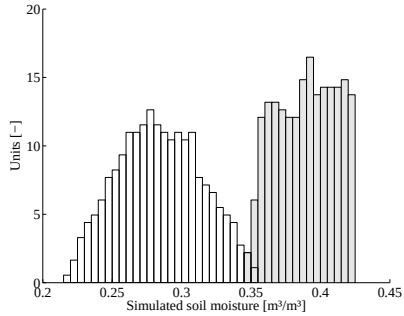


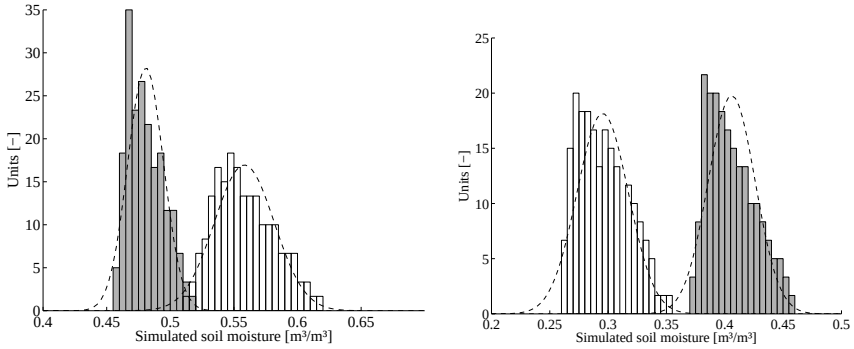
Figure 14.8: Probability distribution of the average simulated soil moisture content, θ_{avg} [$m^3 \cdot m^{-3}$], using the SHPs as predicted from PTF A (grey) and PTF B (white) for the USDA class 'clay loam'.

14.3.3. Implications for data assimilation

The above findings provide some suggestions for data assimilation applications as the soil moisture that resulted from the hydrological model application in the previous section can be assimilated using e.g. measurements of soil moisture or radar images. By applying data assimilation, an optimal estimate of the state and associated uncertainty can be assessed under an appropriate evaluation of errors (model error, initial conditions, parameter, measurements). The optimal estimate and its accuracy may be obtained from the PDF of the state which is presented as the complete solution to the estimation problem, since it embodies all available information. Prior to the application of a data assimilation algorithm, one should check for the boundary conditions (e.g. Gaussian error or nonlinearity of the system) relevant for the algorithm. In this way, an optimal data assimilation method can be selected and errors associated with the implementation of a suboptimal data assimilation algorithm can be assessed. For a review on the most commonly data assimilation methods used in earth science and hydrology we refer to Reichle (2008), Liu and Gupta (2007), McLaughlin (2002) and Ni-Meister (2008).

If the commonly applied Kalman Filter is used, state estimation through data assimilation typically assumes Gaussian forecast errors. However, as can be observed in Figure 14.7, the state variable θ_{daily} is non-normally distributed throughout the entire simulation period. In addition, due to the nonlinear behaviour of the system, fluctuations in the shape of the state's PDF occur when propagating through the model, especially when rain events or long drought periods occur. Similar conclusions were reported by Zhou et al. (2006) who observed a skewness in the soil moisture for very dry or wet conditions and by Reichle et al. (2002) who found that the distribution of the soil moisture only became skewed for dry conditions, whereas for other conditions, the state's PDF was found to be typically symmetric. Suppose that the state's PDF is approximated by a normal distribution, then

a lot of information on the modelled state would be lost since the PDF cannot be fully represented by its mean and variance. This is illustrated in Figure 14.9, showing the probability distribution of θ_{daily} in a SiCl soil on a wet day (DOY 215) and on a dry day (DOY 260). The corresponding normal distribution is plotted on top of the empirical distribution. From this figure, it can be seen that due to the skewness, the density in low soil moisture values is underestimated when assuming a normal distribution, especially when dealing with dry conditions. Nevertheless, it is difficult to generalize the impact of the Gaussian error assumption considering the profound influence of the parameter prediction method on the model output distribution. Consequently, nonlinear filters accounting for non-normal distributions and nonlinear systems should be used (e.g. particle filters). Applying methods relying on the normal error assumptions in such case may compromise the performance of the assimilation algorithm.



(a) (b)

Figure 14.9: Probability distribution of the simulated soil moisture content, θ_{daily} $[m^3 \cdot m^{-3}]$ using the SHPs predicted from PTF A (grey) and PTF B (white), in a silty clay soil on a wet day (DOY 215) (a) and on a dry day (DOY 260) (b). The dotted black line shows the corresponding normal distribution.

14.3.4. Evaluation of the central point as a proxy for the soil class

It was examined whether the central point (CP), corresponding to the centre of gravity of a soil class, could serve as a representative point to (i) generate a 'best guess' SHP set and (ii) to generate the average behaviour in modelled soil moisture content. The resulting SHPs and the modelled soil moisture were compared to the 'common practice' of using average soil class SHPs from literature.

Representative SHPs

For each of the 9 selected soil classes, the CP was determined with respect to the limits on the PTFs (Figure 11.4). The textural information of the CP was used as

input for PTFs A and B, and resulted in two sets of SHPs (Figure 14.10). The latter were compared to the mean predicted SHPs of the corresponding soil class (which can be found in Tables 14.3 and 14.4). Given their high correlations (not shown), the CP can be considered as a good proxy for the average SHPs per soil class. Furthermore, the SHPs of the CP are mostly located near the centre of the total range in predicted SHPs. In order to assess the impact of the PTF limits on the performance of the CP, the CP corresponding to the full area on the texture triangle of the soil classes falling partly or completely outside the valid area was determined and its corresponding SHPs were calculated and compared to the mean SHPs of the full soil class. Again, the CP was found to be representative for the soil class (not shown), from which it can be assumed that irrespective of the area on the texture triangle covered by the soil class, the CP performs well to generate a representative set of SHPs. Nevertheless, one should be aware that due to the complexity and the nonlinearities in the relationships between the SHPs and soil texture, the CP may not correspond to the real central point in terms of hydraulic properties.

An alternative and more common way to determine a 'best guess' SHP set, is to take average values from literature. Frequently used literature values are those reported by Rawls et al. (1982) and Meyer et al. (1997) (based on data sets from the USA) and are shown in Figure 14.10. It can be seen that the literature values often fall outside the range of SHPs predicted with the continuous PTFs, especially for θ_s there is a large deviation. For some SHPs, like θ_r and K_s, the literature values are in better agreement with SHPs-B than with SHPs-A, while for other parameters, like λ, the opposite is observed. Also note the large disagreement among the literature sources themselves, for example the value of K_s for SiCl. These observed differences highlight the relevance of the reliability of both continuous and class PTFs for determining a representative set of SHPs. In this context, the CP can serve as a tool to construct a more reliable average set of SHPs as compared to those derived from class PTFs. In literature, there is a wider choice of continuous PTFs compared to class PTFs. Most of the PTFs that were developed for a specific region of interest are continuous.

Average soil moisture behaviour

TOPLATS was run with the two parameter sets of the CP, as obtained with both PTFs A and B. The resulting average soil moisture content, θ_{avg}, is shown in Figure 14.10 and was compared to the mean prediction of θ_{avg} from Table 14.7. This comparison pointed out that the CP is a good proxy for the mean θ_{avg} within a given soil class. From Figure 14.10, it can be seen that the model output with the CP tends to be situated near the central part within the total range of model outputs. For the CP of the classes SiL, SiClL, SiCl and Cl, the difference in generated soil moisture between the region and non-region specific PTF is relatively small given the large difference in their predicted SHPs. For the other classes, the discrepancy between both model outputs is large, with only a small overlap in the

range of predicted soil moisture. The average simulated soil moisture obtained with the literature-based SHPs is also shown in Figure 14.10. By analogy with the CP, it can be said that in contrast with the huge differences in the SHPs, the difference in model output is only minor. For many soil classes, the SHPs of Meyer et al. (1997) and Rawls et al. (1982) yield a similar soil moisture content and more or less fall within the range of predicted SHPs-B. These findings put the physical meaning of SHPs into perspective, and tend to characterize the SHPs rather as fitting parameters than as physical parameters. Nevertheless, large differences among the parameter prediction methods remain present, especially for the soils with a high sand content.

To assess whether the CP also gives a good representation of the average temporal behaviour in soil moisture, a subset of the time series of θ_{daily} for the CP of the soil class ClL (chosen because of its central position on the texture triangle) is shown in Figure 14.11, together with the corresponding ensemble mean. It can be seen that the application with the SHPs of the CP is in good agreement with the ensemble mean and is able to capture all fluctuations in the mean soil moisture content. Note that not all soil classes show the same perfect pattern, though the match between the CP and the ensemble mean appears to be very good. Irrespective of the PTF used, the CP seems to be able to represent the average behaviour in simulated soil moisture within a given soil class. The time series of θ_{daily} obtained with the SHPs from Meyer et al. (1997) and Rawls et al. (1982) are quite different from those obtained with the CP. The former more or less show the same behaviour in θ_{daily} as obtained with SHPs-A, but are shifted towards the lower soil moisture values (Figure 14.11). Despite the similar mean θ_{avg}, the literature-based SHPs are unable to generate a similar evolution in θ_{daily} as obtained with SHPs-B. The former result in a more gradual wetting and drying of the soil with less extremes in the high and low soil moisture values. These findings underline the need for a reliable estimation of the SHPs in hydrological modelling since they have a high impact on the modelled soil moisture and since standard values for the SHPs may be inappropriate.

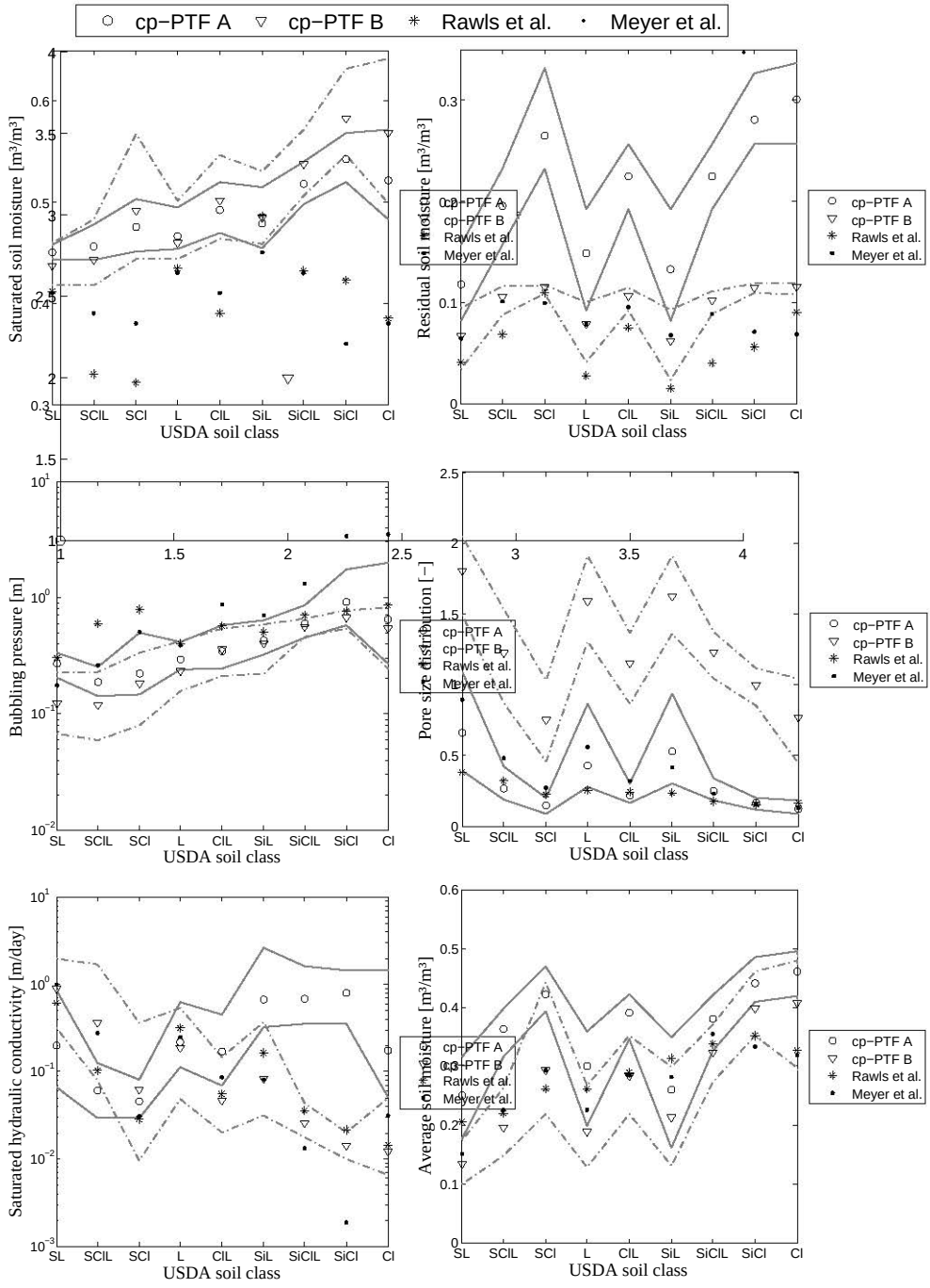
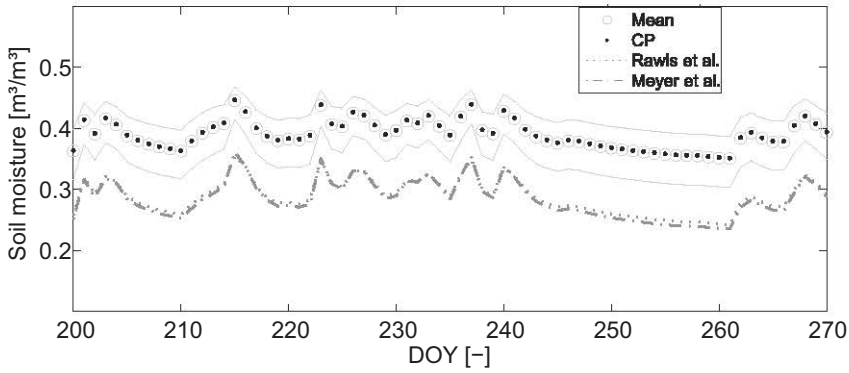
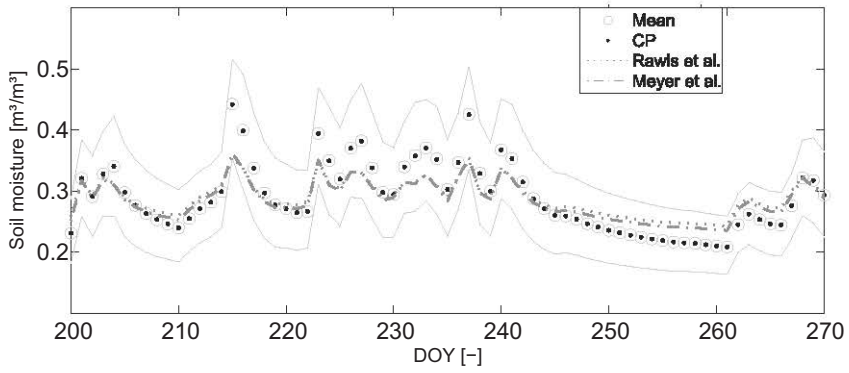


Figure 14.10: Range of the SHPs and θ_{avg} within each USDA class as predicted from PTF A (solid line) and PTF B (dotted line) with indication of the SHP for the central point (CP). The results are compared to standard SHPs as reported by Rawls et al. (1982) and Meyer et al. (1997).



(a)



(b)

Figure 14.11: Time series of the ensemble mean of the simulated soil moisture content, θ_{daily} [$m^3 \cdot m^{-3}$] resulting from PTF A (a) and PTF B (b) for the soil class clay loam with indication of the range in model output for comparison with the model run for the central point and with the SHPs taken from by Rawls et al. (1982) and Meyer et al. (1997).

14.4. Conclusion

The incorporation of uncertainty in the soil parameters is important for a correct representation of the hydrological model response. Unfortunately, modelling of uncertainty is not a standard practice in hydrological modelling (mainly due to limited computational resources) and there is a lack of framework for assessing parameter uncertainty and propagating the uncertainty through the model.

In this chapter, we dealt with two uncertainty issues related to simulations in basins where soil information is limited to a soil type map: (i) within-soil-class variability of the SHPs, and (ii) the reliability of the PTF (region specific versus non-region specific) used to estimate the SHPs. A methodology was presented to quantify the uncertainty in the SHPs arising from the textural variability within soil classes and to propagate the SHP distribution through a hydrological model with conservation of its higher order moments. The intention of the proposed method was not to provide the 'true' SHP uncertainty, but to give a 'reliable' estimate of the uncertainty based on the application of continuous region specific PTFs. The method was illustrated by means of a synthetic experiment. It was shown that the resulting PDFs of the SHPs were non-normally distributed, irrespective of the PTF used. However, a non-region specific PTF resulted in a substantial different SHP uncertainty in comparison with a region specific PTF, in terms of (i) the absolute SHP-values, (ii) the range in predicted SHPs and (iii) the shape of the distribution. These differences could be attributed to, amongst other factors, a different number of samples in the data set, the geographical origin of the soil data (determining the climatic and soil conditions), the methods used to measure the SHPs, etc. Selecting the PTF is therefore a crucial step when assessing the uncertainty in the SHPs. If region specific PTFs are available, they are preferable over standard (often non-region specific) PTFs. Anyhow, PTFs remain poor predictors as long as the structure of the soil is not incorporated in the regression. The PDFs of the SHPs were applied in a hydrological model and were found to result in non-normal distributions of the simulated soil moisture content. Especially under very dry conditions, this non-normality was found to be strong. A representative characterization of the uncertainty hence requires that higher-order moments of the distribution are taken into account. Towards the application of data assimilation algorithms this implies that nonlinear filters accounting for non-normal error distributions should be used since neglecting the higher-order moments of the state's PDF may compromise the assimilation result. In general, a substantial impact of the within-soil-class variability of SHPs on the modelled soil moisture was observed. However, there were large differences between the region and non-region specific PTFs with respect to the modelled soil moisture uncertainty. These differences could be linked to the shape and the spread of the SMRC and HCC. The consistent underestimation of the soil moisture obtained with the non-region specific PTFs as compared to the region specific PTFs, suggested a

fundamental deficiency in the regression equation to account for differences between soils having a different geographical origin. It could therefore be argued that the choice of the PTF is a crucial step when assessing the uncertainty in the soil moisture since differences between PTFs remain present after propagating their SHPs through a hydrological model. Furthermore, an accurate computation of the PDF (including elements of nonlinearity and non-normality) is important if the uncertainty in the soil moisture content needs to be modelled accurately.

In analogy to the common practice, a 'best guess' SHP set was generated in order to obtain a quick assessment of the mean system behaviour, based on the soil type information. Therefore, the CP, corresponding to the centre of gravity of a soil class, was selected to serve as representative texture sample for a given soil class. The SHPs of the CP were found to show a high correspondence with the mean SHPs of the soil class. Furthermore, the CP was found to be able to represent the average behaviour in modelled soil moisture content. Literature values for the SHPs reported by Rawls et al. (1982) and Meyer et al. (1997) could either not capture the average soil moisture or the fluctuation in the soil moisture behaviour as obtained with the best guess SHPs. Based on these findings, it could be concluded that the CP is a useful tool to generate a more reliable behaviour in mean soil moisture as compared to the standard class PTFs, since the CP allows for the use of continuous PTFs. In literature, there is a wider choice of the latter, such that they are often available for the specific region of interest.

In the future, more attention should be paid to appropriately address the SHP uncertainty due to within-soil-class variability, since this generates a substantial amount of uncertainty in the modelled soil moisture. One should aim at developing methods that allow for a fast and accurate assessment of the uncertainty since current methods are time consuming and hence not feasible in practice. However, to obtain a reliable estimate of the uncertainty in the hydrological model output, a reliable PTF needs to be selected and the output PDF need to be fully characterized (including the higher-order moments). When used in a data assimilation framework, it would be useful to assess the impact of the higher-order moments on the assimilation result.

15

Sensitivity of soil moisture to changes in soil texture

15.1. Introduction

In hydrological studies, modellers are sometimes confronted with multivariate data that carry only relative information of which the components represent parts of a whole. Such type of data is called compositional or closed data as the components always sum to a constant, e.g. 1 or 100%. A typical example is the sedimentary particle size distribution of which the closed character implies that the components are not free to vary independently such that if one of its components (e.g. clay) decreases (increases), at least one of the others (e.g. silt or sand) must increase (decrease). Because of this particular property, the application of standard statistical methods to compositional data is hampered and many of the results are invalid because the methods are inappropriate to analyse this type of data. Problems in the analysis of compositional data have been discussed since the end of the twentieth century by a number of authors (e.g. Aitchison, 1986, Aitchison and Egozcue, 2005).

In case a sensitivity analysis (SA) on a multivariate input is intended, the variables can be varied simultaneously based on their underlying probability distribution (e.g. Gwo et al., 1996), or they can be varied one-at-a-time (OAT) (e.g. Ferreira et al., 1995). For more details on the opportunities and threats of OAT-SA, the reader is referred to Section 4.1.2. In case the multivariate input is compositional, the different components of the input are related through the closure balance, and consequently an OAT-SA on its individual components is not justified, but instead all components should be varied simultaneously in order to preserve the closed character of the data. Despite the need to deal with this type of data in hydrological models, limited research on sensitivity analysis involving compositional

This chapter is based on: Loosvelt L., Vernieuwe, H., Pauwels, V.R.N. De Baets B. and Verhoest N.E.C., Local sensitivity analysis for compositional data with application to soil texture in hydrologic modelling, *Hydrology and Earth System Sciences*, 17, 461-478, 2013.

model inputs has been reported to date. Often, the methods applied do not or only partly respect the characteristic properties of compositional data. For example, Bormann (2007) defined a neighbourhood sensitivity for soil texture by applying a fixed change of 1% in the portion of clay or silt while keeping the portion of silt, respectively sand fixed, although a simultaneous change in all of its portions would have been expected.

The main objective of this chapter is to develop a sensitivity analysis method that allows for quantifying the sensitivity of a model output w.r.t. a specific input scenario in case the model input consists of compositional data. To that end, the finite difference technique has been adopted and modified to deal with the closed character of the inputs. The method comprises the calculation of an omnidirectional local sensitivity index that indicates the average impact on the model output when perturbing the compositional model input in different directions around a given point. Since the results of the derivative-based method depend on the magnitude of perturbation (Breshears et al., 1992), especially in case the model shows strong nonlinear relationships and correlations (Saltelli et al., 2000), the method also includes a procedure to optimize the perturbation factor. Subsequently, the SA method is applied to the hydrological model TOPLATS and is used to evaluate changes in the simulated soil moisture content w.r.t. small local changes in soil texture, of which the composition was varied across the entire input domain, as defined by the soil texture triangle. On the basis of this generated local sensitivity index, we aim at locating regions in the texture triangle to which the modelled soil moisture is most sensitive.

In addition to constructing and applying the SA method, another objective of this chapter is to gain more insight into the behaviour of the hydrological model, and more specifically with regard to the role of soil texture therein. Information on soil texture is essential for the operation of a hydrological model since it is used to estimate the soil hydraulic parameters from the soil hydraulic model. Because soil texture is often measured at a number of sparsely distributed locations within the study area, all locations falling into the same soil type as the one of the sampled location (*cfr.* information on the soil map) are attributed the same hydraulic properties within the hydrological model. This discrepancy between the scale of measurement (spacing, *cfr.* scale triplet (Blöschl and Sivapalan, 1995)) and the scale of model application (grid resolution) raises doubts about the suitability of the measured input value as the most probable (Barth et al., 2001) since it may give rise to large uncertainties in the model output. In this perspective, the presented sensitivity analysis offers the possibility to reduce this type of model output uncertainty by formulating guidelines for additional data collection as a function of the measured soil texture.

The hydrological model set-up used in this study is described in Section 15.2. The basic principles of operations on compositional data are briefly described in Section

15.3. These operations are used to develop a technique to locally perturb soil texture in Section 15.4. Finally, the results are presented and discussed in Section 15.5.

15.2. Hydrological model set-up

TOPLATS is applied at a point location (with coordinates 50.89 °N and 4.09 °E) in the catchment of the Bellebeek in order to simulate the soil moisture content of the upper soil layer (5 cm) during the period January 1, 2006 to December 31, 2006, using an hourly time step. For the simulation point, the vegetation parameters were attributed literature-based values (*cfr.* Section 8.2), the TOPMODEL and routing parameters were attributed the calibrated values (*cfr.* Tables 8.1 and 8.3). The soil and land cover type registered at the simulation point are loam and bare soil, respectively. The texture, however, was not fixed but was varied within the sensitivity analysis. The meteorological data collected for the Bellebeek catchment during the year 2006 were used to force the model. Furthermore, *in situ* soil moisture measurements (at 2.5 cm depth) taken between May 13 and May 30, 2007 were used to validate the model (*cfr.* Section 7.4).

In TOPLATS, the soil moisture retention curve (SMRC) and the hydraulic conductivity curve (HCC) are modelled through the closed-form analytical equations of Brooks and Corey (1964) (*cfr.* Section 6.1.4). The shape of both curves is determined by the soil hydraulic parameters (SHPs): the residual soil moisture content θ_r [$m^3 \cdot m^{-3}$], the saturated soil moisture content θ_s [$m^3 \cdot m^{-3}$], the bubbling pressure ψ_c [m], the pore size distribution index λ [$-$] and the saturated hydraulic conductivity K_s [$m \cdot s^{-1}$]. When field measurements of the SHPs are not available, they are estimated based on soil textural information (soil type or particle size distribution) through application of either class or continuous pedotransfer functions (PTFs). Numerous PTFs have been proposed, reviewed and evaluated over the last decade (e.g. Tietje and Tapkenhinrichs, 1993, Wagner et al., 2001, Nemes et al., 2009), but the accuracy and reliability of the PTFs are highly variable (*cfr.* Chapter 14) and mainly depend on the similarity of the soil and climatic features between the region of PTF development and the region of PTF application.

In this study, the continuous PTFs of Rawls and Brakensiek (1985, 1989) (Table 14.1) were applied to estimate the SHPs for the simulation point based on the sand content Z [%], the clay content C [%], and the soil porosity P [-]. The latter was calculated following Eq.(14.1). The particle density was corrected for the presence of organic matter, for which a content of 3 % (Sleutel et al., 2006) and a density of 1.45 $g \cdot cm^{-3}$ (Kaiser and Guggenberger, 2003, Mayer et al., 2004) was assumed. The bulk density was calculated following the procedure as described by Saxton and Rawls (2006). When applying the PTFs of Rawls and Brakensiek (1985, 1989), one should bare in mind that these PTFs were actually developed for textures with

a clay content between 5 % and 60 % and a sand content between 5 % and 70 % (*cfr.* Figure 14.1 for the boundary conditions on the PTF).

15.3. Compositional data

15.3.1. Basic concept and operations

Compositional or closed data are multivariate data, represented by positive real vectors of which the components sum up to a constant κ_C. The components of the vector show the relative weight or importance of the parts in a total, which means that compositional data carry only relative information. A typical example of compositional data is soil texture, which provides information on the relative portion of sand, clay and silt in a given soil sample, and of which the closed character implies that changing one portion causes the other portions to change as well, such that the sum of the portions remains equal to 100%. The set of all possible compositions $\mathbf{x}$ with D components forms a simplex sample space, denoted as S^D, and is defined as:

$$S^D = \{\mathbf{x} = (x_1, x_2, ..., x_D) \,|x_i \geq 0,\ i = 1, 2, ..., D;\ \sum_{i=1}^{D} x_i = \kappa_C > 0\} \qquad (15.1)$$

where x_i is the i-th part of composition $\mathbf{x}$, and κ_C is the closure constant of which the value is generally 1 (proportions) or 100 (percentage). For the specific problem setting in this study, the sample space is a simplex with $\kappa_C = 100$ and $D = 3$, as the soil texture encloses three different parts that sum up to 100%. In the simplex, the composition $\mathbf{p_0}$ with coordinates $\left(\frac{100}{3}, \frac{100}{3}, \frac{100}{3}\right)$ is called the baricenter and can be conceived as the origin of the sample space.

Specific operations and statistical properties (e.g. distributions) for compositional data were introduced by Aitchison (1986) and were further developed by Egozcue and Pawlowsky-Glahn (2006). The basic operations on the simplex that are relevant for the sensitivity analysis are summarized below. For a comprehensive description of these and other properties, the reader is referred to Aitchison (1982).

- Vector addition of composition $\mathbf{x} \in S^D$ and composition $\mathbf{x}' \in S^D$ (also called perturbation) (Aitchison, 1986):

$$\mathbf{x} \oplus \mathbf{x}' = \left(\frac{x_1 \cdot x'_1}{\sum_{i=1}^{D} x_i \cdot x'_i}, \frac{x_2 \cdot x'_2}{\sum_{i=1}^{D} x_i \cdot x'_i}, ..., \frac{x_D \cdot x'_D}{\sum_{i=1}^{D} x_i \cdot x'_i} \right) \qquad (15.2)$$

  For a detailed discussion on the visualization, the role and the interpretation of addition in the simplex, we refer to Aitchison and Ng (2005) and von Eynatten et al. (2002).

- Scalar multiplication of a composition $\mathbf{x} \in S^D$ by a scalar $\omega \in \mathbb{R}$ (also called power transformation) (Aitchison, 1986):

$$\omega \odot \mathbf{x} = \left(\frac{x_1^\omega}{\sum_{i=1}^D x_i^\omega}, \frac{x_2^\omega}{\sum_{i=1}^D x_i^\omega}, ..., \frac{x_D^\omega}{\sum_{i=1}^D x_i^\omega} \right) \tag{15.3}$$

- Aitchison distance between composition $\mathbf{x} \in S^D$ and composition $\mathbf{x}' \in S^D$ (Aitchison, 1983):

$$d_{\mathrm{A}}(\mathbf{x}, \mathbf{x}') = \sqrt{\frac{1}{2D} \sum_{i=1}^{D} \sum_{j=1}^{D} \left(\ln\left(\frac{x_i}{x_j}\right) - \ln\left(\frac{x_i'}{x_j'}\right) \right)^2} \tag{15.4}$$

  The Aitchison distance is a measure for the difference between two compositions $\mathbf{x}$ and $\mathbf{x}'$ (Aitchison, 1992). If one of the compositions corresponds to the baricenter (e.g. $\mathbf{x}' = \mathbf{p_0} = \left(\frac{\kappa_C}{D}, \frac{\kappa_C}{D}, ..., \frac{\kappa_C}{D}\right)$), then $d_{\mathrm{A}}(\mathbf{x}, \mathbf{p_0})$ is equal to the norm of $\mathbf{x}$, denoted as $\| \mathbf{x} \|_{\mathrm{A}}$.

Furthermore, it is worth mentioning that coordinates in the vector space can be transformed into a Cartesian coordinate system. A frequently used transformation is the isometric logratio (ILR) transformation, which preserves all metric properties (Egozcue et al., 2003). Although a coordinate transformation is not required within the presented SA method, it will be used for a better understanding of the operations in the simplex and as an alternative approach for sensitivity analysis in case of high-dimensional compositions ($D > 3$).

15.3.2. Soil texture in the simplex

The texture of a soil sample $\mathbf{x} = (C, Z, L)$ is defined by the distribution of the soil particle sizes C (clay, diameter < 2 μm), Z (sand, diameter > 2 mm) and L (silt, 2 μm $<$ diameter < 2 mm). Because the parts cannot vary independently (there are only two degrees of freedom), it is possible to visualize the soil texture, a 3D composition, in two dimensions by means of an equivalent representation in the texture triangle (Figure 15.1). This is an equilateral triangle with vertices at $\mathbf{p_1} = (100, 0, 0)$, $\mathbf{p_2} = (0, 100, 0)$ and $\mathbf{p_3} = (0, 0, 100)$. The three vertices are defined counter-clockwise and are connected through the segments $\mathbf{p_1 p_3}$, $\mathbf{p_3 p_2}$ and $\mathbf{p_2 p_1}$, scaled from 0 to 100.

In the texture triangle, three bisectors are defined as the straight lines through one of the vertices and the baricenter $\mathbf{p_0} = \left(\frac{100}{3}, \frac{100}{3}, \frac{100}{3}\right)$ (*cfr.* Figure 15.1). The bisector through vertex $\mathbf{p_1}$, $\mathbf{p_2}$ and $\mathbf{p_3}$ is referred to as B_1, B_2, and B_3, respectively, and has the property that the values of two parts of the composition are always equal on this line.

Scaling soil textures

Consider a soil texture $\mathbf{x_0} = (x_1, x_2, x_3)$ randomly sampled on the texture triangle. The texture can be (re)scaled with a factor ω following the operation $\omega \odot \mathbf{x_0}$. For a number of discrete ω values between -10 and 10, the result of this operation is shown in Figure 15.2. All rescaled textures lie on a compositional straight line going through the baricenter and through two of the vertices. The compositional line is expressed by an equation of the form $\mathbf{x}(\omega) = \mathbf{p_0} \oplus (\omega \odot \mathbf{x_0})$ in which the sampled texture determines the direction of the line. The higher the value of ω, the further the scaled texture is removed from the sampled texture. Negative values of ω cause the scaled textures to fall into the other side of the texture triangle when divided by bisector B_1.

Adding soil textures

A set of other textures, generated as $\omega \odot \mathbf{x}'$, can be added to the sampled texture $\mathbf{x_0}$ by means of the operation $\mathbf{x} = \mathbf{x_0} \oplus (\omega \odot \mathbf{x}')$ in which ω takes a value between 1 and 100. This results in a compositional straight line in which $\mathbf{x_0}$ is the starting point and $\mathbf{x}'$ is the direction. Here, we consider the directions of the bisectors B_1, B_2 and B_3 such that $\mathbf{x}'$ and $\omega \odot \mathbf{x}'$ lie on one of these bisectors. The result of adding rescaled versions of $\mathbf{x}'$ to the sampled texture $\mathbf{x_0}$ is shown in Figure 15.3. In this case, the compositional straight lines also define straight lines on the ternary diagram and go through the sampled texture and one of the vertices. In case the rescaled textures $\omega \odot \mathbf{x}'$ do not lie on the bisectors, but are obtained after rescaling $\mathbf{x}' \oplus \mathbf{x_0}$, the compositional straight line connects the sampled texture with two of the vertices (Figure 15.3). The resulting equation is then given by $\mathbf{x}(\omega) = \mathbf{x_0} \oplus (\omega \odot (\mathbf{x}' \oplus \mathbf{x_0}))$.

15.4. Local sensitivity analysis on compositional data

The aim of the local sensitivity analysis is to measure the effect of perturbing a specific composition $\mathbf{x}$, *i.e.* inducing small relative changes to the composition, on the model output y. The sensitivity of y with respect to $\mathbf{x}$ is expressed as a sensitivity function that is defined as the derivative of y with respect to $\mathbf{x}$ and is evaluated at one particular value of $\mathbf{x}$ by using the finite difference approximation. Therefore, small changes in $\mathbf{x}$ need to be imposed that, considering the closed character of $\mathbf{x}$, imply a change in each of its parts x_i $(i = 1, 2, ..., D)$ while maintaining $\sum_{i=1}^{D} x_i = \kappa_C$. The model output y is said to be sensitive to model input $\mathbf{x}$ if small changes in $\mathbf{x}$ produce large changes in y. On the contrary, y is called insensitive to $\mathbf{x}$ if small changes in $\mathbf{x}$ have almost no effect on y.

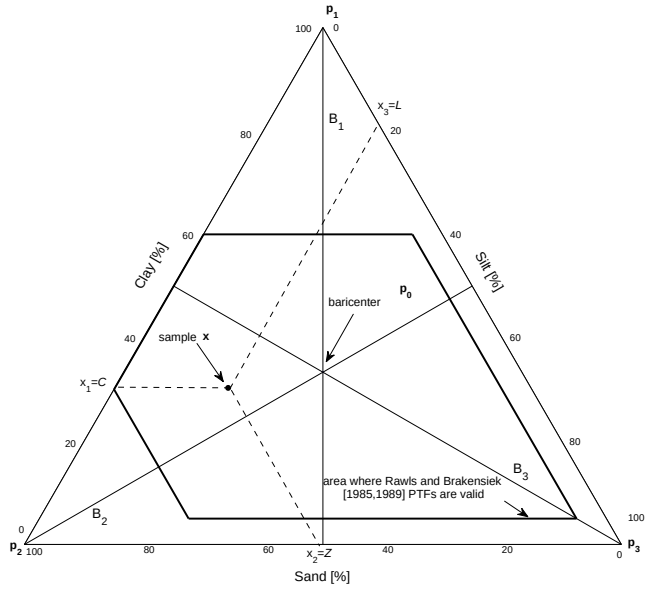


Figure 15.1: Representation of a sample $\mathbf{x} = (C, Z, L)$ in the texture triangle with indication of the bisectors B_1, B_2 and B_3, the vertices $\mathbf{p_1}$, $\mathbf{p_2}$, $\mathbf{p_3}$ and the boundary conditions on the PTFs of Rawls and Brakensiek (1985, 1989).

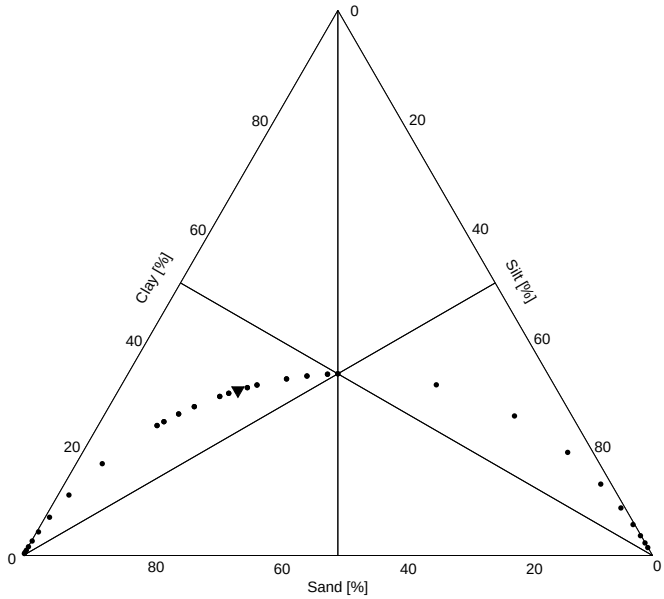


Figure 15.2: Rescaling a sampled soil texture $\mathbf{x_0}$ (indicated with a triangle) with a factor ω, varying between -10 and 10, following the operation $\omega \odot \mathbf{x_0}$.

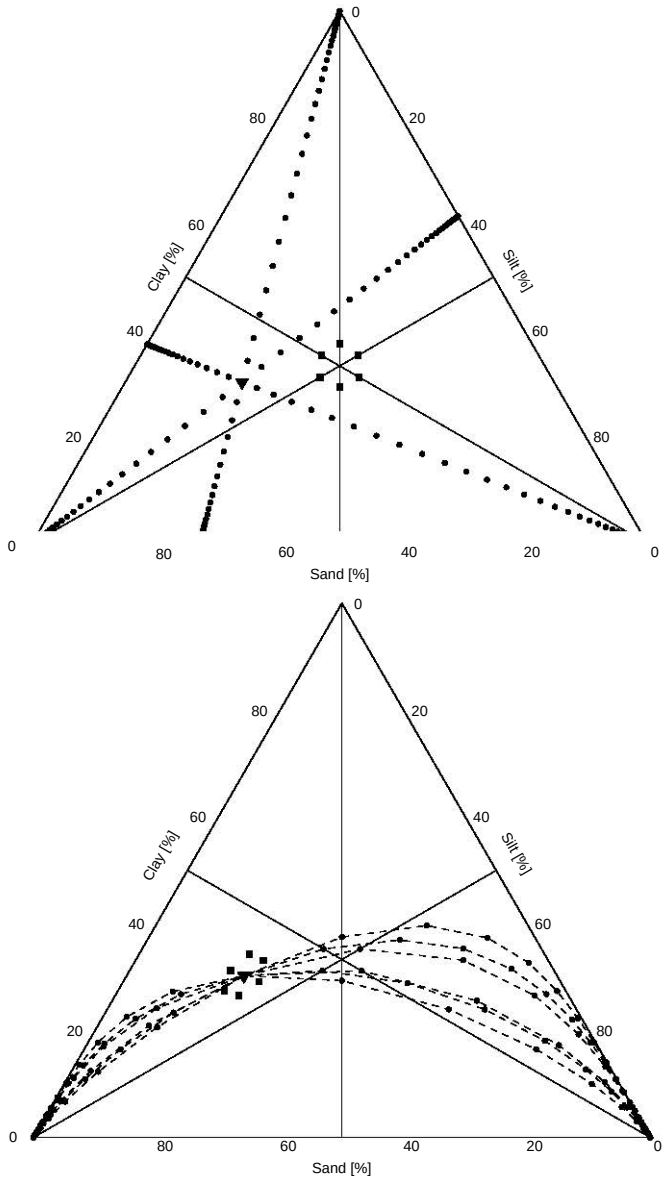


Figure 15.3: Adding to a sampled soil texture $\mathbf{x_0}$ (indicated with a triangle) (a) a set of other textures situated on the bisectors of the texture triangle and obtained after scaling texture $\mathbf{x}'$ (indicated with squares), and (b) a set of other textures situated in the neigbourhood of the sampled texture with the same direction of $\mathbf{x}'$ and obtained after scaling $\mathbf{x_0}$ (indicated with squares).

15.4.1. Perturbing in the 2D Euclidean space

The methodology presented in this section for perturbing compositional data is built by analogy with a perturbation in a two-dimensional Euclidean space. Suppose we want to simultaneously perturb two inputs x_1^* and x_2^* with a factor ξ, four possible outcomes are evident and are given by the Cartesian coordinates $(x_1^* \cdot (1+\xi), x_2^* \cdot (1+\xi))$, $(x_1^* \cdot (1-\xi), x_2^* \cdot (1-\xi))$, $(x_1^* \cdot (1+\xi), x_2^* \cdot (1-\xi))$, $(x_1^* \cdot (1-\xi), x_2^* \cdot (1+\xi))$. These are the intersections of the bisectors of the Cartesian coordinate system and the circle with centre $\mathbf{x} = (x_1^*, x_2^*)$ and radius $d = \sqrt{(\xi x_1^*)^2 + (\xi x_2^*)^2}$. The circle defines all possible perturbations and is therefore further referred to as the perturbation circle. Since it is impossible to evaluate an infinite number of perturbations, only a limited set of perturbed points, e.g. the four points on the bisectors, can be considered.

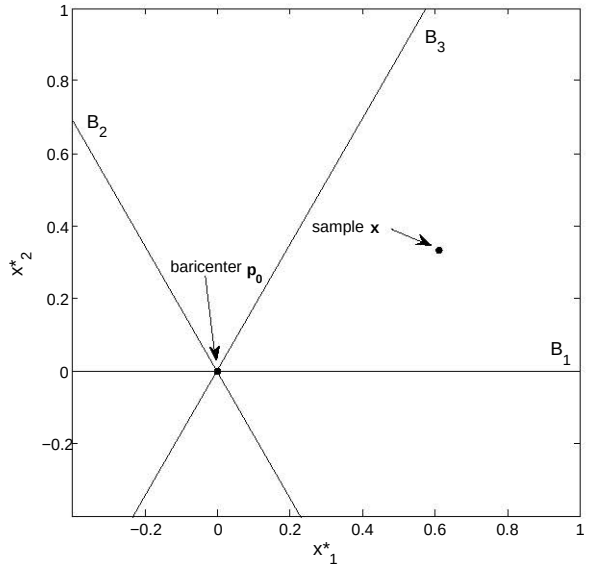
This idea is adopted for the perturbation of a 3D composition $\mathbf{x} = (x_1, x_2, x_3)$. Consider a random sample $\mathbf{x}$ from S^3 which is defined by its triangular coordinates in a ternary diagram (*cfr.* Figure 15.1). Through ILR transformation, the composition can be represented in the 2D Euclidean space by means of the Cartesian coordinates (Figure 15.4 (a)):

$$(x_1^*, x_2^*) = \left(\frac{1}{\sqrt{6}} \ln \left(\frac{x_1^2}{x_2 \cdot x_3} \right), \frac{1}{\sqrt{2}} \ln \left(\frac{x_2}{x_3} \right) \right) \tag{15.5}$$

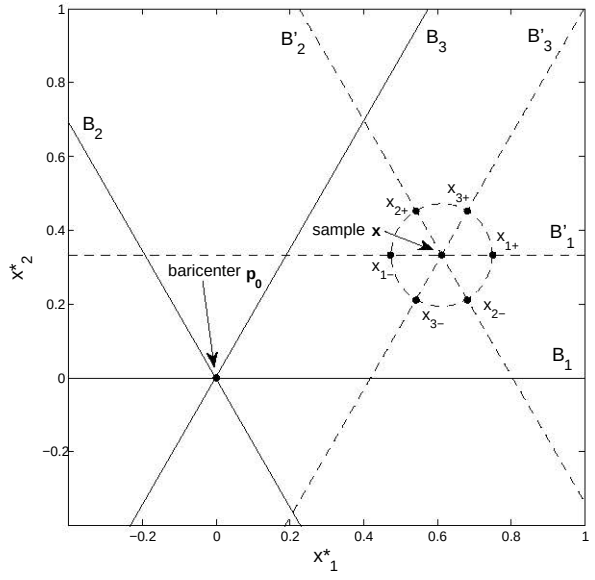
Likewise, any geometric shape on the ternary diagram can be transformed. Figure 15.4 (a) shows that after ILR transformation, the bisectors B_1, B_2, and B_3 preserve their angles of 60 ° (*cfr.* Section 15.3.2) and the baricenter $\mathbf{p_0}$ forms the origin of the Cartesian system in which the bisectors intersect. The perturbation of sample $\mathbf{x}$ with factor ξ can now be performed in the Euclidean space, following the methodology described above. First of all, the perturbation circle with centre $\mathbf{x} = (x_1^*, x_2^*)$ and radius $d = \sqrt{(\xi x_1^*)^2 + (\xi x_2^*)^2}$ is constructed (Figure 15.4 (b)). As only a perturbation in the directions given by the bisectors (further called perturbation axes) is considered, the directions of the bisectors are transferred to composition $\mathbf{x}$ by means of a translation. The perturbed points are then defined by the intersections $\mathbf{x_{i+}}$ and $\mathbf{x_{i-}}$ between the translated bisectors B'_i ($i \in \{1, 2, .., D\}$) and the perturbation circle (Figure 15.4 (b)). Finally, the Cartesian coordinates of the perturbed compositions $\mathbf{x_{i+}}$ and $\mathbf{x_{i-}}$ are backtransformed to the simplex through an inverse ILR transformation (Egozcue et al., 2003) (Figure 15.5, see further).

15.4.2. Perturbing in the simplex

Yet, in order to avoid the roundabout method of coordinate transformations, the operations in the Euclidean space are mimicked by operations in the simplex. This



(a)



(b)

Figure 15.4: (a) Representation of the baricenter $\mathbf{p_0}$, the sampled composition $\mathbf{x}$ and the bisectors B_1, B_2 and B_3 after ILR transformation and (b) illustration of perturbation in the 2D Euclidean space with indication of the perturbed compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ and the translated bisectors B'_1, B'_2 and B'_3.

results in following procedure to perturb a composition $\mathbf{x}$ with a factor ξ:

1. Perform one of the scalar multiplications $\mathbf{x}_{\pm} = (1 \pm \xi) \odot \mathbf{x}$ (Eq. (15.3)) in order to rescale the composition $\mathbf{x}$ with a factor ξ. The scaling factor ξ determines the magnitude of the perturbation, *i.e.* the higher the value of ξ, the more the perturbed composition will deviate from the sampled composition $\mathbf{x}$.

2. Calculate the Aitchison distance $d_{\mathrm{A}}(\mathbf{x}, \mathbf{x}_{\pm}) = d$ (Eq. (15.4)) in order to quantify the difference between the sampled composition $\mathbf{x}$ and the rescaled compositions $\mathbf{x}_{+}$ and $\mathbf{x}_{-}$. Note that for the same value of ξ, the value of d increases with increasing values of $\| \mathbf{x} \|_{\mathrm{A}}$.

3. Define a circle with centre $\mathbf{p_0}$ and radius d and determine the intersections between the circle and the perturbation axes, here the bisectors B_1, B_2, and B_3 (Figure 15.5). For each axis in direction $i \in \{1, 2, .., D\}$, this problem is solved by searching for the compositions $\mathbf{v_{i+}}$ and $\mathbf{v_{i-}}$ on the axis that satisfy the condition $\| \mathbf{v_i} \|_{\mathrm{A}} = d$ (the procedure is described below). The resulting compositions are further called directional vectors because they are necessary to transfer the direction of the perturbation axes to the sampled composition $\mathbf{x}$.

4. Add the directional vectors $\mathbf{v_{i+}}$ and $\mathbf{v_{i-}}$ ($i \in \{1, 2, .., D\}$) to the composition $\mathbf{x}$ (Eq. (15.2)). This results in three pairs of new compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ that lie on the perturbation circle (Figure 15.5). Since the performed operation preserves the distance in the simplex, the perturbation circle around $\mathbf{x}$ has radius d. Although its circular shape is distorted in the simplex (the further from the baricenter, the more distortion; examples of circles are shown in Figure 15.5), the definition of a circle remains valid.

In summary, when perturbing composition $\mathbf{x}$ in the simplex S^D with a fixed factor ξ following the methodology described above, we obtain three pairs of new compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ with $i \in \{1, 2, .., D\}$ that are a subset of all possible perturbations, defined by the circle with centre $\mathbf{x}$ and radius $d = d_a(\mathbf{x}, \mathbf{x_{i\pm}})$. Note that also $M \in \mathbb{N}$ other perturbation axes could have been chosen by selecting M points on a circle around the baricenter, either at random or such that they form angles of $\frac{360}{2M}$ degrees, and by connecting each of the selected points with the baricenter. In case of M perturbation axes, steps 4 and 5 from the methodology would respectively result in M pairs of directional vectors $\{\mathbf{v_{i+}}, \mathbf{v_{i-}}\}$ and M pairs of perturbed compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$. In this study, we selected the bisectors as perturbation axes (hence, $M = 3$) such that the perturbed compositions (i) define angles of $60\,^{\circ}$ degrees on the perturbation circle, and (ii) lie on the translated bisectors B'_{i}, connecting $\mathbf{x}$ with vertex $\mathbf{p_i}$ (*cfr.* Figure 15.5). The compositional lines B'_{i} also illustrate the effect of increasing (or decreasing) the magnitude of perturbation in direction i, as the perturbed compositions always lie on this line but shift towards (or away from) vertex $\mathbf{p_i}$.

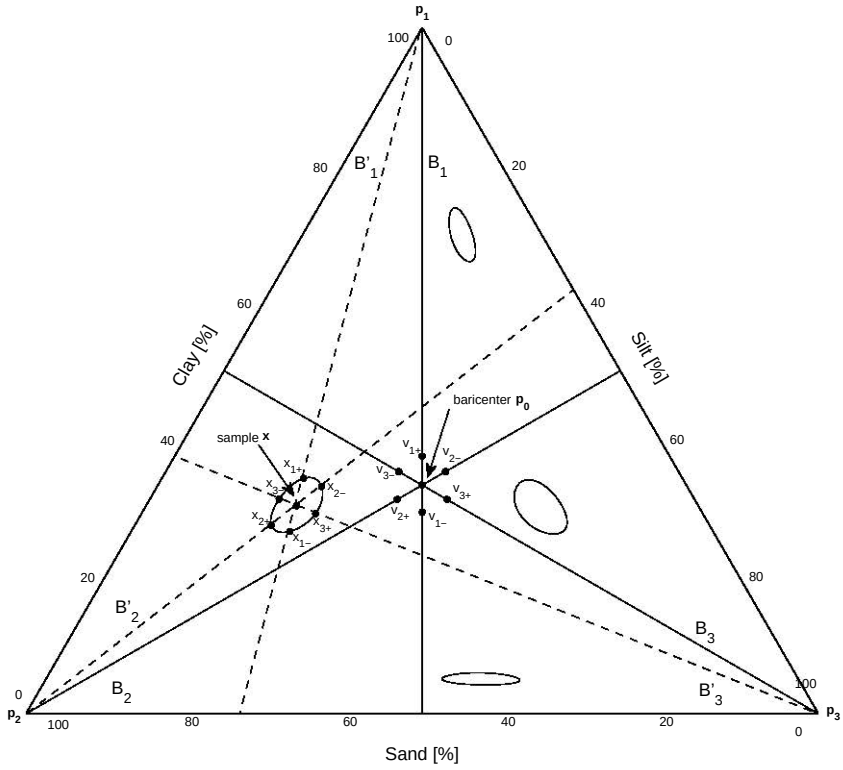


Figure 15.5: Perturbation in the simplex of a composition $\mathbf{x}$ sampled from the texture triangle with indication of the directional vectors $\{\mathbf{v_{i+}}, \mathbf{v_{i-}}\}$, the perturbed compositions $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$, the bisectors B_1, B_2 and B_3 and the translated bisectors B'_1, B'_2 and B'_3. Compositional circles at different locations in the texture triangle are shown for illustration.

Note that generalizing the methodology might raise some difficulties (Egozcue, 2012). For $M > 3$, calculating the directional vectors in the ternary diagram becomes complicated. In this case, it is advised to rely on a simplicial expression that computes the ILR coordinates of the directional vectors as the intersection points of the circle with radius d and center $\mathbf{p_0}$ and M perturbation axes, regularly distributed on the circle. Through inverse ILR transformation, the directional vectors are expressed in the simplex. For $D > 3$, e.g. when soil texture is described with more than three parts, the perturbation circle is not easily generalized to a sphere or hyper-sphere and the distribution of the perturbation axes might raise difficulties (Egozcue, 2012).

Procedure to determine the directional vectors

The directional vectors $\mathbf{v} = (v_1, v_2, v_3)$ are defined as the intersection between the bisectors B_1, B_2, B_3 and the circle with centre $\mathbf{p_0} = \left(\frac{100}{3}, \frac{100}{3}, \frac{100}{3}\right)$ and radius d. This definition implies the following:

1. The coordinates of $\mathbf{v}$ must respect the system of Cartesian equations that describe each of the bisectors. Since on the bisectors, two components are always equal, the problem boils down to finding two parts of $\mathbf{v}$ instead of three.

2. The norm of $\mathbf{v}$ should equal d, as $\mathbf{v}$ should be at a distance d from $\mathbf{p_0}$.

For the directional vectors on B_1, the problem is mathematically defined as:

$$\begin{cases} \|\mathbf{v}\|_{\mathrm{A}} = \sqrt{\dfrac{1}{6}\displaystyle\sum_{i=1}^{3}\sum_{j=1}^{3}\left(\ln\left(\dfrac{v_i}{v_j}\right)\right)^2} = d \\ v_2 = \dfrac{1-v_1}{2} \\ v_2 = v_3 \end{cases} \tag{15.6}$$

Solving this system of equations results in two solutions for v_1 and v_2:

$$\begin{cases} v_{1a} = \dfrac{1}{2e^{\sqrt{3/2}d}+1}, & v_{1b} = \dfrac{1}{2e^{-\sqrt{3/2}d}+1} \\ v_{2a} = \dfrac{e^{\sqrt{3/2}d}}{2e^{\sqrt{3/2}d}+1}, & v_{2b} = \dfrac{e^{-\sqrt{3/2}d}}{2e^{-\sqrt{3/2}d}+1} \end{cases} \tag{15.7}$$

To determine the directional vectors on B_2 and B_3, the same procedure is followed. Table 15.1 summarizes the coordinates of the resulting vectors. The minus sign indicates the point on the bisector that is the furthest away from its vertex, whereas the plus sign refers to the point on the bisector that is the closest to the vertex (*cfr.* Figure 15.5).

Table 15.1: Coordinates of the directional vectors.

	$\mathbf{v}_+$	$\mathbf{v}_-$
B_1	(v_{1a}, v_{2a}, v_{2a})	(v_{1b}, v_{2b}, v_{2b})
B_2	(v_{2a}, v_{1a}, v_{2a})	(v_{2b}, v_{1b}, v_{2b})
B_3	(v_{2a}, v_{2a}, v_{1a})	(v_{2b}, v_{2b}, v_{1b})

15.4.3. Calculating the sensitivity index

The methodology for perturbing a composition in the simplex allows for approximating the sensitivity function by means of the finite difference technique. The sensitivity function can be further summarized into a sensitivity index in order to express the sensitivity of y to small changes in $\mathbf{x}$ by means of a single value. In this section, it is described how the sensitivity index is evaluated at one particular value of $\mathbf{x}$.

In the simplex, a composition $\mathbf{x}$ is sampled uniformly at random from the sample space. Therefore, a Dirichlet distribution is defined, which is the multivariate generalization of the beta distribution and is parametrized by a vector $\boldsymbol{\alpha}$. The density function of the Dirichlet distribution is given by:

$$f_D(\mathbf{x};\boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{i=1}^{D}\alpha_i\right)}{\Pi_{i=1}^{D}\Gamma(\alpha_i)}\Pi_{i=1}^{D}x_i^{\alpha_i-1} \tag{15.8}$$

with $\mathbf{x} = (x_1, ..., x_D)$ a sample from S^D, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_D)$ the parameter vector and D the dimension of the sample space. In order to guarantee that the composition is sampled uniformly at random, the condition $\boldsymbol{\alpha} = 1$ should be fulfilled. The sampled composition $\mathbf{x}$ is thereupon perturbed in M different directions following the methodology described in Section 15.4.2.

After sampling and perturbing $\mathbf{x}$, the model output y_t is determined at time step t for both the sampled composition $\mathbf{x}$ and the perturbed compositions $\mathbf{x_{i+}}$ and $\mathbf{x_{i-}}$. For each direction i given by the perturbation axes, a forward and backward directional sensitivity function, respectively denoted as $\nabla_{\mathbf{v_{i+}}} y_t(\mathbf{x})$ and $\nabla_{\mathbf{v_{i-}}} y_t(\mathbf{x})$, are calculated using the finite difference technique (*cfr.* Section 4.1):

$$\begin{cases} \nabla_{\mathbf{v_{i+}}} y_t(\mathbf{x}) \approx \dfrac{y_t(\mathbf{x}\oplus\mathbf{v_{i+}}) - y_t(\mathbf{x})}{d_{\mathrm{A}}(\mathbf{x}\oplus\mathbf{v_{i+}},\mathbf{x})} = \dfrac{y_t(\mathbf{x_{i+}}) - y_t(\mathbf{x})}{d} \\ \nabla_{\mathbf{v_{i-}}} y_t(\mathbf{x}) \approx \dfrac{y_t(\mathbf{x}) - y_t(\mathbf{x}\oplus\mathbf{v_{i-}})}{d_{\mathrm{A}}(\mathbf{x}\oplus\mathbf{v_{i-}},\mathbf{x})} = \dfrac{y_t(\mathbf{x}) - y_t(\mathbf{x_{i-}})}{d} \end{cases} \tag{15.9}$$

with t the time step in the simulation period, $\mathbf{v_{i\pm}}$ the directional vectors for perturbation in direction i and d the Aitchison distance (*cfr.* Eq. (15.4)). Averaging both functions leads to a central, directional sensitivity function $\nabla_i y_t(\mathbf{x})$, which indicates the average change in y caused by opposite changes in $\mathbf{x}$ in the direction of perturbation axis i:

$$\nabla_i y_t(\mathbf{x}) \approx \frac{y_t(\mathbf{x}\oplus\mathbf{v_{i+}}) - y_t(\mathbf{x}\oplus\mathbf{v_{i-}})}{d_{\mathrm{A}}(\mathbf{x}\oplus\mathbf{v_{i+}},\mathbf{x}\oplus\mathbf{v_{i-}})} = \frac{y_t(\mathbf{x_{i+}}) - y_t(\mathbf{x_{i-}})}{2d}. \tag{15.10}$$

As the sensitivity function itself is not useful for sensitivity analysis (Saltelli et al., 2008b), it is summarized into an omnidirectional local sensitivity index S (by analogy with the sensitivity index proposed by Hill and Tiedeman (2007), *cfr.* Eq. (4.4)):

$$S(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M}\sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\nabla_i y_t(\mathbf{x})\right)^2} \tag{15.11}$$

with N the number of time steps in the model output and M the number of perturbation axes (here $M = 3$). The sensitivity index is hence a single value that reflects the average response of y when $\mathbf{x}$ is perturbed with a fixed factor ξ in M

Algorithm 7: Calculate the local sensitivity index S for a sampled composition $\mathbf{x}$

Data: - a composition $\mathbf{x}$ from the sample space S^D (Section 15.4.2);
- a perturbation factor ξ (Section 15.4.4);
- a set of M perturbation directions;

Result: sensitivity index S for y at $\mathbf{x}$

begin
  perform the scalar multiplication $(1 \pm \xi) \odot \mathbf{x} = \mathbf{x}_{\pm}$ (Section 15.4.2, step 1);
  calculate the Aitchison distance $d_{\mathrm{A}}(\mathbf{x}, \mathbf{x}_{\pm})$ (Section 15.4.2, step 2);
  for *each perturbation direction* $i \in \{1, 2, .., M\}$ **do**
    determine the directional vectors $\{\mathbf{v_{i+}}, \mathbf{v_{i-}}\}$ (Section 15.4.2, step 3);
    add the directional vectors to $\mathbf{x}$ to obtain $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ (Section 15.4.2, step 4);
    apply the model to determine $y_t(\mathbf{x_{i+}})$ and $y_t(\mathbf{x_{i-}})$;
    approximate the central, directional sensitivity function $\nabla_i y_t(\mathbf{x})$ (Eq. (15.10))
  end
  calculate the omnidirectional local sensitivity index $S(\mathbf{x})$ (Eq. (15.11))
end

different directions, each covering two opposite changes in $\mathbf{x}$. It is calculated as the root mean squared difference in the model output resulting from two oppositely perturbed compositions, averaged over the different perturbation directions. As such, the sensitivity index can be easily updated when $\nabla_i y_t(\mathbf{x})$ is calculated for additional perturbation directions, *i.e.* $M > 3$. In case the model output y is time-independent, *i.e.* $N = 1$, then S reduces to the mean absolute difference in the model output resulting from two oppositely perturbed compositions. An overview of the methodology to calculate the sensitivity index is given in Algorithm 7.

Note that for $D > 3$, a large number of points on the perturbation (hyper-)sphere will be required to compute the sensitivity index. In this case, computation of the sensitivity function can be simplified by estimating the directional derivatives based on the ILR coordinates of the sampled composition:

$$\nabla_{\mathbf{v_i}} y_t(\mathbf{x}) = \left(\frac{\partial y_t}{\partial x_1^*}, \frac{\partial y_t}{\partial x_2^*}, ..., \frac{\partial y_t}{\partial x_{D-1}^*} \right) \cdot \left(v_{1,i}^*, v_{2,i}^*, ..., v_{D-1,i}^* \right)^\top \tag{15.12}$$

As such, any directional derivative can be computed if for each of the $D - 1$ orthogonal axis on the ILR coordinate space, the gradient $\partial y_t / \partial x^*$ has been determined (Egozcue, 2012).

15.4.4. Optimizing the perturbation factor

Because the choice of the perturbation factor ξ determines the quality of the sensitivity function (as discussed in Section 4.1.1), an optimization procedure is included in the presented SA framework. The basic idea is to make both the numerical errors (if ξ is too small) and the errors due to model nonlinearities (if ξ is too large) as small as possible by minimizing the difference in model sensitivity when inducing opposite changes in the model input, *i.e.* the difference between the sensitivity functions $\nabla_{\mathbf{v_{i+}}} y_t(\mathbf{x})$ and $\nabla_{\mathbf{v_{i-}}} y_t(\mathbf{x})$ should be as small as possible when perturbing in direction i. Although several measures can be used to quantify this difference (*cfr.* Section 4.1.1), the sum of squared difference between their absolute values is selected in this study and is denoted as $C_i(\mathbf{x})$:

$$C_i(\mathbf{x}) = \frac{1}{N} \sum_{t=1}^{N} \left(\left| \nabla_{\mathbf{v_{i+}}} y_t(\mathbf{x}) \right| - \left| \nabla_{\mathbf{v_{i-}}} y_t(\mathbf{x}) \right| \right)^2. \tag{15.13}$$

By taking the absolute value of the sensitivity functions, we allow that opposite changes in $\mathbf{x}$ result in non-opposite, but similar, model responses. Since the sensitivity analysis explores M directions on the perturbation circle, M values of $C_i(\mathbf{x})$ are obtained on which the minimization procedure needs to be carried out. In order to solve this optimization problem, the maximum value over all $C_i(\mathbf{x})$ (with $i \in \{1, 2, .., M\}$), further called $C_{\max}(\mathbf{x})$, is used as objective function. For the SA problem in this study, it is chosen to limit ξ to a set of fixed values, such that $\xi \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ although other intermediate values could have been chosen as well. An overview of the methodology to optimize the perturbation factor for a sampled composition is given in Algorithm 8.

15.5. Results and discussion

The experimental set-up consists of two main steps: in the first step, 100 textures are sampled from the texture triangle to determine the optimal factor ξ for perturbing soil texture. In the second step, 5000 textures are sampled from the texture triangle and are used as input to the hydrological model to evaluate the response of the simulated soil moisture when texture is perturbed with the optimal perturbation factor.

15.5.1. Identification of the optimal perturbation factor

Hundred compositions are sampled from the texture triangle according to a Dirichlet distribution (with $\boldsymbol{\alpha} = 1$). For each sampled texture, the perturbation factor ξ is evaluated for the values in $\{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ by applying the

Algorithm 8: Optimize the perturbation factor for a sampled composition

Data: - a composition $\mathbf{x}$ from the sample space S^D (Section 15.4.2);
- a set of L perturbation factors ξ (Section 15.4.4);
- a set of M perturbation directions;

Result: optimal perturbation factor ξ at $\mathbf{x}$

begin
  apply the model to determine $y_t(\mathbf{x})$;
  for *each* $\xi \in \{\xi_1, \xi_2, .., \xi_L\}$ **do**
    perform the scalar multiplication $(1 \pm \xi) \odot \mathbf{x} = \mathbf{x}_\pm$ (Section 15.4.2, step 1);
    calculate the Aitchison distance $d_\mathrm{A}(\mathbf{x}, \mathbf{x}_\pm)$ (Section 15.4.2, step 2);
    for *each perturbation direction* $i \in \{1, 2, .., M\}$ **do**
      determine the directional vectors $\{\mathbf{v_{i+}}, \mathbf{v_{i-}}\}$ (Section 15.4.2, step 3);
      add the directional vectors to $\mathbf{x}$ to obtain $\{\mathbf{x_{i+}}, \mathbf{x_{i-}}\}$ (Section 15.4.2, step 4);
      apply the model to determine $y_t(\mathbf{x_{i+}})$ and $y_t(\mathbf{x_{i-}})$;
      approximate the sensitivity functions $\nabla_{\mathbf{v_{i+}}} y_t(\mathbf{x})$ and $\nabla_{\mathbf{v_{i-}}} y_t(\mathbf{x})$ (Eq. (15.9));
      calculate $C_i(\mathbf{x})$ (Eq. 15.13);
    end
    determine $C_{\max}(\mathbf{x})$ as $\max_{i=1}^{M} C_i(\mathbf{x})$;
  end
  select the value of ξ for which $C_{\max}(\mathbf{x})$ is minimal;
end

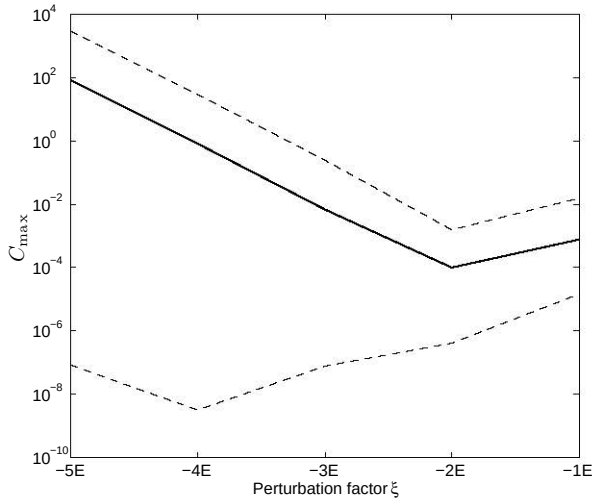
methodology as described in Algorithm 8. For each ξ, we obtain 100 values of $C_{\max}$ of which the mean, the minimum and the maximum are shown in Figure 15.6(a). The mean value of $C_{\max}$ clearly shows a minimum for $\xi = 10^{-2}$, which suggests that this value is optimal for perturbing a broad range of textures. Although the minimum value of $C_{\max}$ is lowest for $\xi = 10^{-4}$, this perturbation factor can be discarded as being optimal since the spread over the different values of $C_{\max}$ is very large for this value of ξ. Or stated differently: when $\xi = 10^{-4}$ would be selected as optimal, it would result in a very low value of $C_{\max}$ for only a limited number of textures whereas the majority of the samples would be characterized by a larger value of $C_{\max}$. These findings are in correspondence with the frequency distribution of the optimal ξ-values (Figure 15.6(b)), which shows that for the major part (about 65%) of the samples, $C_{\max}$ is minimal when they are perturbed with 10^{-2}. For 25% and 10% of the samples, the optimal value of ξ is respectively smaller and larger than 10^{-2}. The samples from the group with an optimal ξ smaller than 10^{-2} show a relatively heterogeneous distribution on the texture triangle (Figure 15.7) with the highest concentration around textures with a clay content above 40% or a sand content between 20% and 50%. The samples from the group with an optimal

ξ larger than 10^{-2} are mainly located around textures with a sand and clay content of 30%.

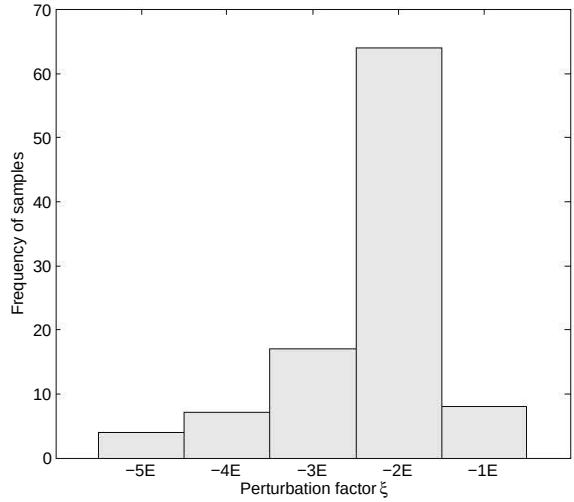
For practical purposes, it is chosen to use a fixed value of 10^{-2} for the perturbation factor, as it would unnecessarily increase the complexity of the sensitivity analysis when making ξ dependent on the sampled texture. Consequently, deviations from the optimal value are mainly located within the texture classes *clay loam* and *loam* (*cfr.* Figure 15.7).

15.5.2. Identification of sensitive regions in the texture triangle

Identifying sensitive regions in the texture triangle with respect to the estimation of soil parameters or with respect to the prediction of soil moisture is useful for the model user as it allows for reducing the uncertainty in the predicted variable. Since standard available soil information is often limited to a soil map of the study area and a number of sparsely distributed soil texture measurements, the measurement is assumed to be representative for the corresponding soil map unit. Consequently, the model user attributes the same particle size distribution to all locations falling into that soil map unit, whereas the soil texture at a location different from the measurement point, but within the same soil map unit, may (largely) deviate from the sampled texture. Although it is assumed that the spatial variability within a homogeneous soil map unit covers only a minor part of the total variability in texture that is enclosed within the definition of the corresponding soil type, the discrepancy between the scale of measurement and the scale of model application might introduce large uncertainties in the model output. If large uncertainties in either the estimated SHPs or the simulated soil moisture are not acceptable (depending on the objective of the study), the uncertainty about the potential bias in the measured soil texture due to spatial variability should be further reduced through additional data collection. If the pattern in sensitivity is identified, the following rule of thumb to prioritize additional data collection can be applied: "If the sampled texture, which is assumed to be representative for a certain soil map unit, is located within a region of high sensitivity in the texture triangle, then additional texture samples within the area corresponding to this soil map unit, as delineated on the soil map, should be taken". By accounting for the spatial variability, a more accurate estimate of the representative (most probable) texture for the given soil map unit can be formulated and can be used to reduce the uncertainty in the model output. On the contrary, if the sampled texture is located within a region of low sensitivity in the texture triangle, the discrepancy between the scale of measurement and the scale of model application will have a low impact on the predicted variable, and taking additional samples may therefore be discarded, unless a high spatial variability in soil texture exists.



(a)



(b)

Figure 15.6: The mean (full line), minimum and maximum (dotted lines) values of $C_{\max}$ as a function of the perturbation factor ξ (a) and the frequency distribution of the optimal perturbation factor (b), for 100 sampled compositions.

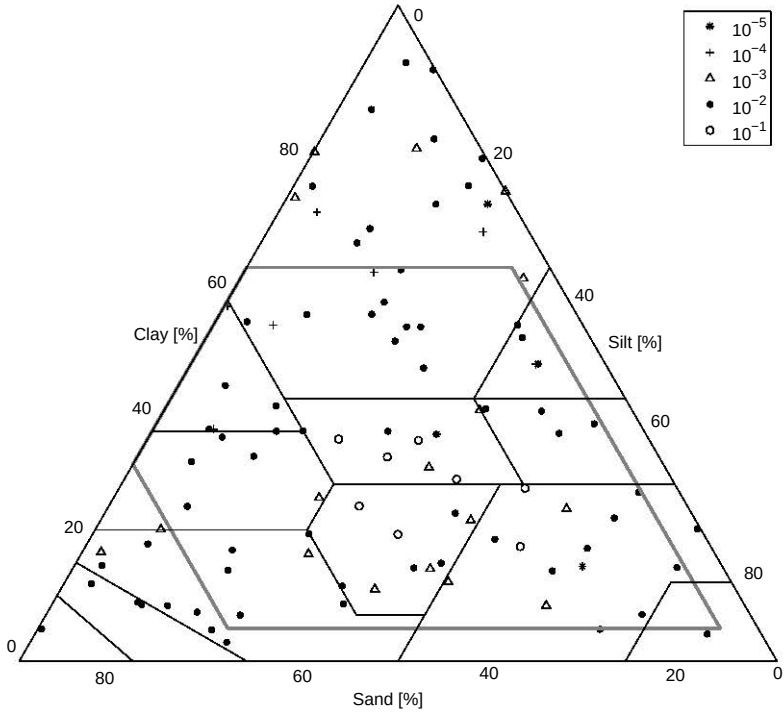


Figure 15.7: Distribution of the optimal perturbation factor ξ (100 samples) in the texture triangle with delineation of the USDA soil class (the corresponding class labels can be found in Figure 7.2).

Sensitivity of soil hydraulic parameters

Five thousand compositions $\mathbf{T} = (C, Z, L)$ are sampled from the texture triangle according to a Dirichlet distribution (with $\boldsymbol{\alpha} = 1$) and are perturbed with $\xi = 10^{-2}$. For both the sampled and perturbed compositions, the corresponding SHPs are estimated with the PTFs of Rawls and Brakensiek (1985), on which the sensitivity function $\nabla_i \mathrm{SHP}(\mathbf{T})$ with $i \in \{1, 2, 3\}$ and the sensitivity indices S_{SHP} are calculated by applying the methodology as described in Algorithm 7. A contour plot of S_{SHP} (Figure 15.8 (a)-(e)) reveals that the sensitivity pattern highly depends on the parameter under consideration, although the patterns in S_{θ_s} and S_{ψ_c} show a remarkable resemblance. For these parameters, the hot spot of high sensitivity is located around textures with a clay content of 60-80% and a sand content of 20-40%. The sensitivities S_{θ_r} and S_λ show a pattern that is highly dominated by the clay content: an increase in the clay content causes an increase (decrease) in the sensitivity if the clay content is lower (higher) than 30%. Although the clay content of the hot spot matches for S_{θ_r} and S_λ, the corresponding sand content is different: around 0% for the former and around 70% for the latter. On the contrary, the pattern in S_{K_s} is dominated by the sand content: the higher the sand content, the higher the sensitivity. The order of magnitude of the sensitivity index

should be interpreted with respect to the corresponding SHP. Therefore, the mean predicted SHP over the entire texture triangle is given as a reference in Figure 15.8. Also note that results outside the validity zone of the PTFs should be interpreted with care (*cfr.* Section 15.2 and Figure 15.1). This zone is indicated on the contour plots in Figure 15.8. In summary, the potential uncertainty in the predicted SHPs due to the discrepancy in scale between measurement and model application highly varies across the texture triangle and among the different SHPs, making it very difficult to formulate general guidelines to reduce the uncertainty in the predicted SHPs.

Sensitivity of soil moisture

After determining the sensitivity of the estimated SHPs towards the soil texture, the SHPs are used as input to the hydrological model TOPLATS in order to simulate the daily soil moisture content θ_t during the year 2006 at the simulation location under Belgian weather conditions (*cfr.* Section 15.2). The 5000 sampled textures and their perturbed textures are successively attributed to the simulation location on which the corresponding sensitivity functions $\nabla_i \theta_t(\mathbf{T})$ with $i \in \{1, 2, 3\}$ and sensitivity index S_θ are calculated as described in Algorithm 7. Figure 15.8 (f) shows a contour plot of S_θ as a function of $\mathbf{T} = (C, Z, L)$ and reveals a rather simple sensitivity pattern. For textures with a clay content lower than 35% or higher than 70%, the sensitivity is strongly determined by the clay content. In case $C < 35\%$, the textural sensitivity increases with increasing values of the clay content, whereas in case $C > 70\%$ the textural sensitivity decreases with increasing values of the clay content. For soils with a clay content between 35% and 70%, S_θ is also highly influenced by the percentage of sand in the soil. The hot spot of high sensitivity is located around textures with a clay and sand content of 55% and 45%, respectively. This means that for these measured textures, the potential uncertainty in θ that is associated with the scaling issue will be the highest, but can, however, be reduced through additional data collection.

Evaluation of the USDA class as sensitivity region

The objective of this section is to investigate whether a soil map of a region with indication of the USDA soil classes can be used as a rudimentary tool to set up the texture sampling strategy prior to data collection. As such, the discrepancy between the scale of measurement and the scale of model application within that region is optimally managed with respect to the uncertainty in the model prediction. For the prediction in a region that goes together with an USDA soil class that is attributed a high sensitivity towards soil texture, it is important to reduce the uncertainty in the textural variability within that region. Obviously, sufficient samples should be taken to accurately estimate the representative texture. By using the representative texture, a lower model prediction uncertainty is achieved. Otherwise, if the soil map indicates that the prediction will take place in a region where the soil class shows a low sensitivity towards texture, then resources can be saved and data collection can be limited to a single soil sample. However, this

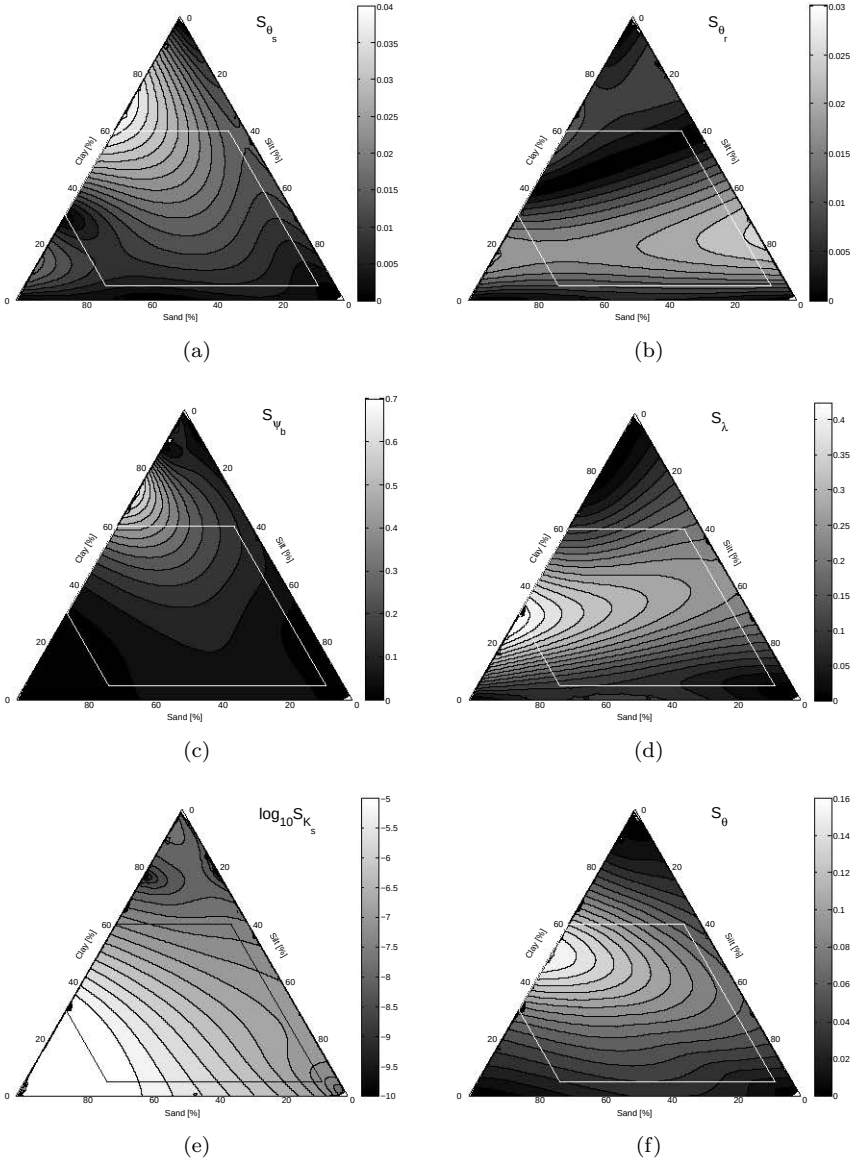


Figure 15.8: Contour plot of the sensitivity index across the texture triangle for the estimated soil hydraulic parameters (a)-(e) θ_s (mean is 0.18 $\mathrm{m}^3 \cdot \mathrm{m}^{-3}$), θ_r (mean is 0.03 $\mathrm{m}^3 \cdot \mathrm{m}^{-3}$), ψ_c (0.16 m), λ (mean is 0.49), and $\log_{10} K_s$ (mean is $-6.11\,\mathrm{m} \cdot \mathrm{s}^{-1}$) and for the simulated soil moisture content θ (f). Results outside the validity zone of the PTFs (indicated by a white line) and near the borders of the triangle should be interpreted with care.

strategy is only valid under the assumption that the textural variability within that region is low.

In order to associate regions of high and low sensitivity of the SHP estimation with the commonly used USDA soil classification, S_{SHP} is averaged over the samples falling into the same USDA soil class, further denoted as $\overline{S}_{\mathrm{SHP}}$ (Table 15.2). For θ_s and ψ_c, the soil class corresponding to the highest $\overline{S}_{\mathrm{SHP}}$ is *clay*, whereas for θ_r, λ and K_s, this is respectively *silt loam*, *sandy clay loam* and *sandy loam*, which are also the classes that contain the hot spot of high sensitivity for their respective SHP (Figure 15.8). However, the hot spot only covers a part of the entire soil class, such that advising a higher sampling density to formulate a representative texture for the corresponding soil map units will not always be relevant. Based on these results, we may argue that the USDA classification is only useful as a preliminary indication for the sensitivity in SHP prediction because the variation in sensitivity within an USDA class is often high. As a consequence, the USDA soil map is suboptimal when used as a tool to optimize the sampling strategy with respect to the potential uncertainty in the estimated SHPs that is associated with the scaling issue. These findings call for a refinement of the USDA soil classes as they seem to be too rough to accurately describe the soil texture. In addition, the definition of soil texture should evolve towards a representation with more than three grain classes.

Table 15.2: Average sensitivity index within the USDA soil classes, for the different soil hydraulic parameters. The class with the highest average sensitivity is indicated in boldface, whereas the class with the lowest average sensitivity is underlined.

Soil Class	$\overline{S}_{\theta_s}$	$\overline{S}_{\theta_r}$	$\overline{S}_{\psi_b}$	$\overline{S}_{\lambda}$	$\overline{S}_{K_s}$
Sand	5.84E-04	5.02E-04	5.02E-04	7.21E-03	1.77E-06
Loamy Sand	1.33E-03	1.44E-03	2.24E-03	2.04E-02	4.84E-06
Sandy Loam	4.72E-03	8.12E-03	2.55E-02	1.09E-01	**1.07E-05**
Sandy Clay Loam	4.10E-03	5.33E-03	2.85E-02	1.47E-01	5.02E-06
Sandy Clay	4.34E-03	8.02E-04	3.35E-02	5.72E-02	3.77E-07
Loam	4.41E-03	6.41E-03	3.37E-02	7.31E-02	7.89E-07
Clay Loam	6.25E-03	4.35E-03	5.07E-02	1.00E-01	2.38E-07
Silt Loam	6.90E-03	**1.52E-02**	5.95E-02	1.02E-01	5.97E-07
Silt	6.48E-04	2.11E-03	6.91E-03	6.26E-03	2.52E-08
Silty Clay Loam	3.25E-03	4.79E-03	2.18E-02	6.40E-02	4.46E-08
Silty Clay	3.22E-03	1.83E-03	2.62E-02	5.39E-02	1.91E-08
Clay	**2.39E-02**	5.80E-03	**2.61E-01**	**1.49E-01**	9.45E-08

By analogy, S_θ is averaged over each USDA soil class and the resulting $\overline{S}_\theta$ is shown together with its standard deviation in Figure 15.9. Based on the total range in S_θ, four sensitivity classes are defined: low sensitivity ($0 \leq \overline{S}_\theta < 0.04$), medium sensitivity ($0.04 \leq \overline{S}_\theta < 0.08$), high sensitivity ($0.08 \leq \overline{S}_\theta < 0.12$) and very high sensitivity ($0.12 \leq \overline{S}_\theta$). The soil class *sandy clay* is attributed the highest $\overline{S}_\theta$ and falls into the high sensitivity class, which is obvious as this class contains the hot

spot of high sensitivity (*cfr.* Figure 15.8). However, the variation in sensitivity within that soil class is very large and ranges from medium to very high. Also other soil classes enclose more than one sensitivity class, e.g. *clay* and *silt*, which would require to re(de)fine those soil classes with respect to their sensitivity. On the contrary, some USDA classes completely fall within a single sensitivity class. The classes *loamy sand* and *sand* therefore correctly represent a low sensitivity whereas the classes *loam*, *silty clay loam* and *silty clay* represent a medium sensitivity. This supports the earlier findings to use the USDA soil classification (and hence the USDA soil map) only as a preliminary indication of the model output sensitivity towards textural changes. The dominant sensitivity class that is associated with the USDA class is then an indication of the sampling density needed to formulate a representative texture for the given USDA class. For example the clayey soil classes (e.g. *sandy clay*, *clay loam* and *clay*) will require a higher sampling density.

15.5.3. Identification of the hydrological model behaviour

The scatterplot in Figure 15.11 shows how the model response S_θ is related to the average annual soil moisture content $\theta_{avg} = \frac{1}{N}\sum_{t=1}^{N}\theta_t$ simulated with TOPLATS, from which it is clear that very high model sensitivities are only expected if the simulated soil moisture has a value between 0.2 and 0.4, with a maximum around 0.3. On the contrary, low sensitivities occur for both very low ($\theta_{avg} < 0.2$) and very high soil moisture contents ($\theta_{avg} > 0.45$). This suggests that the more extreme (dry or wet) the soil moisture content becomes, the less uncertainty in the simulation result is involved when there is a discrepancy between the scale of texture measurement and the scale of model application. Similarly, scatterplots between S_θ and the SHPs (Figure 15.11) reveal that the sensitivity can only be very high if the SHPs take specific values: θ_s, θ_r, ψ_c, λ and K_s should be within the range $[0.42, 0.49]$, $[0.1, 0.12]$, $[0.05, 0.6]$, $[0.4, 1]$ and $[5 \cdot 10^{-8}, 5 \cdot 10^{-6}]$, respectively. The parameter values for which a maximum in S_θ is recorded, are combined to construct the SMRC that involves the highest uncertainty in θ. The so-called 'high sensitivity' SMRC shows a rather linear behaviour (Figure 15.10) that is characteristic for fine-textured soils with a low effective porosity, *i.e.* $\theta_s - \theta_r$. For the sake of simplicity, it can be said that this SMRC corresponds to low values of θ_s, ψ_c and λ, and a high value of θ_r. On the contrary, a low sensitivity of the simulated soil moisture is not exclusively related to specific values of the SHPs, since for a broad range of SHP values the corresponding S_θ falls into the low sensitivity class. Nevertheless, it is observed that the more the SHP values deviate from the specified range that gives rise to a very high sensitivity, S_θ shifts towards the low sensitivity class. This means that the sensitivity of the simulated soil moisture is certainly low in case θ_s, ψ_c and λ have a high value and θ_r has a low value. The so-called 'low sensitivity' SMRC that results from this soil parameter combination is characteristic for coarse-textured soils with a high effective porosity and shows a

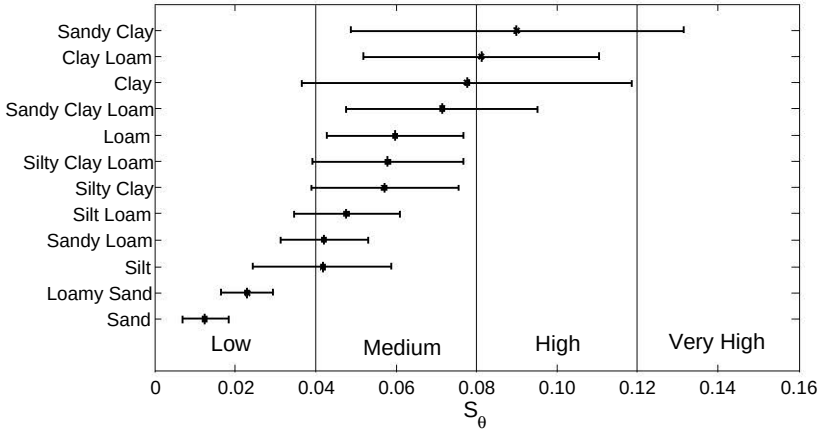


Figure 15.9: Average sensitivity index $\overline{S}_\theta$ for the 12 USDA soil classes with indication of the standard deviation and the sensitivity classes low ($0 \leq \overline{S}_\theta < 0.04$), medium ($0.04 \leq \overline{S}_\theta < 0.08$), high ($0.08 \leq \overline{S}_\theta < 0.12$) and very high ($0.12 \leq \overline{S}_\theta$).

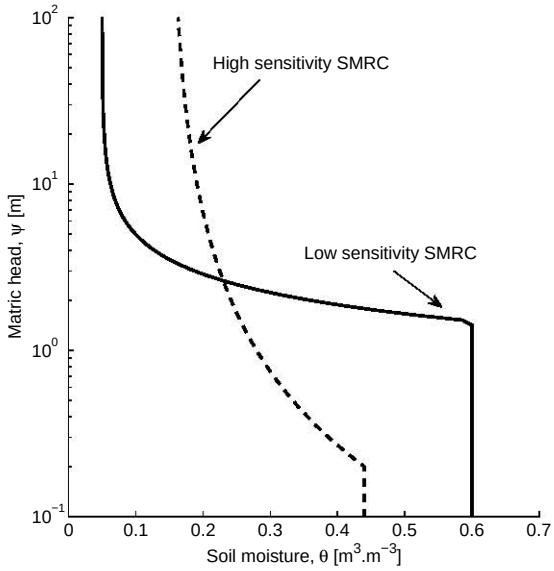


Figure 15.10: Soil moisture retention curve (SMRC) for which the sensitivity of the simulated soil moisture to textural changes is the highest and the lowest, respectively.

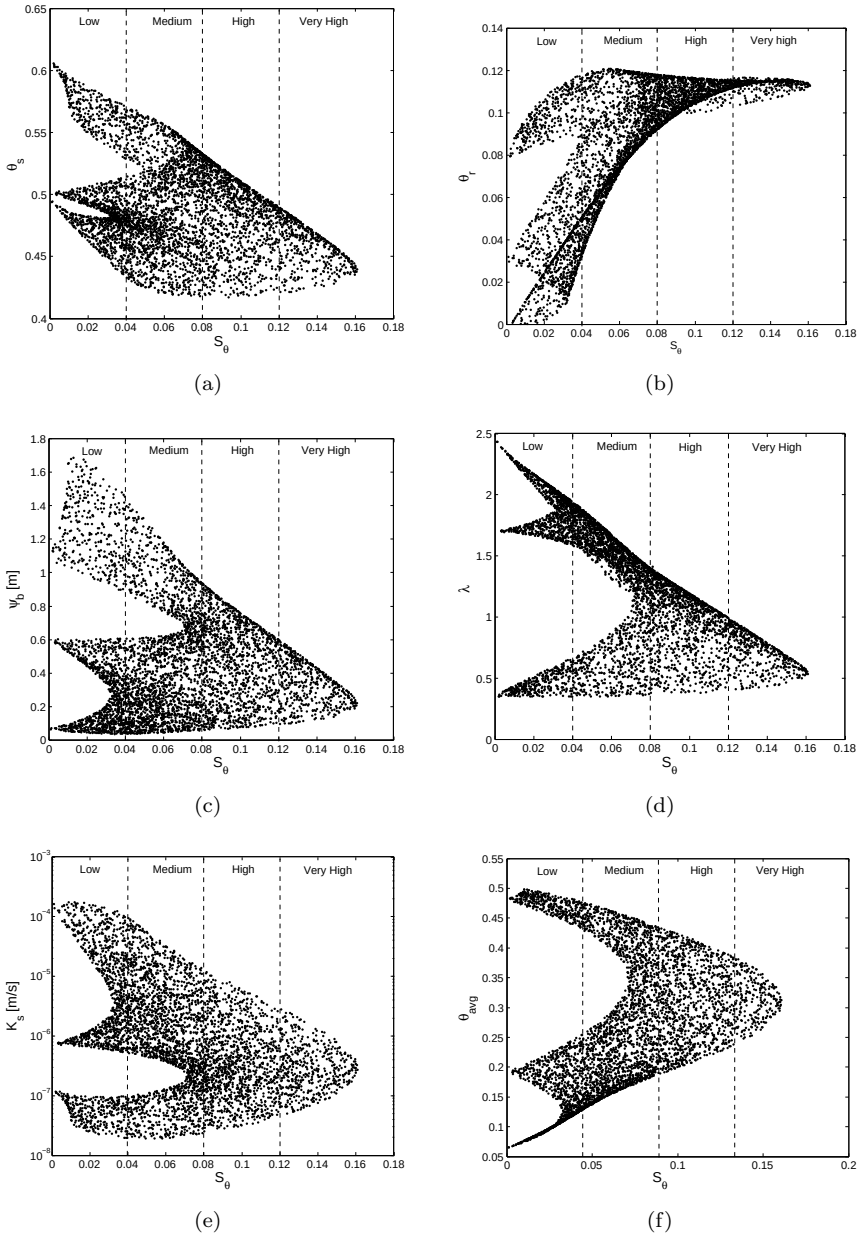


Figure 15.11: Scatterplot between S_θ and the soil hydraulic parameters θ_s, θ_r, ψ_c, λ, K_s (a)-(e) and the average simulated soil moisture content θ_{avg} (f) (5000 sampled textures).

highly nonlinear behaviour.

15.6. Conclusion

Considering the omnipresence of compositional data in the geosciences (and thus also in hydrological modelling), we developed a method to perform a local sensitivity analysis on compositional model inputs. As the different parts of the input vary simultaneously, while preserving the closed character of the input, this method allows to abandon incorrect practice of OAT-SA. In the presented SA method, a sensitivity index is calculated based on the finite difference technique to approximate the directional derivatives of the model output with respect to the compositional model input. Local perturbations of the compositions were realized by operations in the simplex (for complex SA problems we suggest to implement the alternative approach using ILR coordinates) and we relied on the assumption that all possible perturbations are defined by a perturbation circle. Additionally, we supplemented the SA method with a procedure to optimize the perturbation factor in order to minimize numerical errors and errors due to model nonlinearities.

The SA method was subsequently applied to a hydrological model to assess the sensitivity of the simulated soil moisture content to changes in soil texture, for a high number of compositions in the texture triangle. In a first step, we found that the optimal factor to perturb soil texture is 10^{-2}. Although this value was found to be optimal in 65% of the cases, it was chosen to use a fixed value of ξ in order not to unnecessarily complicate the sensitivity analysis. However, one should be aware that in 10% of the cases this value is too low and might introduce numerical errors in the sensitivity analysis, and that in 25% of the cases this value is too high, which might result in errors due to the nonlinear behaviour of the hydrological model. Especially near the borders of the ternary diagram, deviation of the perturbation factor from its optimal value might affect the sensitivity analysis. A perturbation factor of 10^{-2} was used to perform a local SA on 5000 different textures, sampled according to a Dirichlet distribution from the texture triangle. The analysed models are the PTFs of Rawls and Brakensiek (1985) and the hydrological model TOPLATS of which the generated outputs are respectively the soil hydraulic parameters and the soil moisture content. Based on these model applications, the sensitivity index was calculated for both model outputs and was evaluated with respect to the position of the sampled texture in the texture triangle.

The results of the sensitivity analysis were found to be useful (i) to reduce the uncertainty in the modelled output when there is a discrepancy between the scale of measurement and the scale of model application and (ii) to gain more insight into the behaviour of the applied model, and more specifically on how it reacts on changes in the soil texture with respect to its position in the texture triangle. As such, we found that the simulated soil moisture is most sensitive to soil texture

when the measured clay content is about 55% and the sand content about 45%. This means that the potential uncertainty that is involved with the scaling issue will be the highest under these textural conditions. Therefore, when high uncertainties in the modelled output are not acceptable, it is advised to take one or more additional texture samples within the soil map unit that encloses the original sample such that a better estimate of the most probable texture can be formulated. Similarly, we identified zones of high sensitivity for the soil parameters, showing a high variability in their sensitivity pattern. We also investigated whether a soil map with indication of the USDA soil classes can be used as a tool to optimize the texture sampling strategy by reviewing the USDA soil classification with respect to the pattern in model output sensitivity. The results point out that USDA classes are only useful as a rudimentary indication for the sensitivity as they distinguish between high and low sensitivity, but comprise a large within-class-variability of the sensitivity. Especially the clayey soil classes *sandy clay*, *clay loam* and *clay* involve high to very high sensitivities, such that it is advised to apply a high(er) sampling density within these soil map units to calculate the representative texture. Furthermore, we were able to relate S_θ to the shape of the soil moisture retention curve and recorded the highest sensitivity when the values of θ_s, ψ_c and λ are low and the value of θ_r is high. The resulting curve is characteristic for fine-textured soils with a low effective porosity and shows a rather linear behaviour.

Note that all results are specific for the given experimental set-up and that extrapolation of the results towards other flow models and soil parameter estimation methods is not justified as long as the generality of the results is not pointed out. Moreover, the results should be interpreted in a broader perspective of model uncertainties because optimizing the sampling strategy based on the textural sensitivity is only relevant if model uncertainties related to the soil hydraulic parameter estimation, e.g. accuracy and reliability of the selected PTFs, have been assessed and minimized.

General conclusions and outlook

Contribution of the thesis

During the last decades, the complexity of hydrological models has considerably increased in order to better represent the hydrological processes. Since hydrologists have obtained models that are good representations of reality, they are being confronted with a new type of problem: uncertainty in the increasing number of model inputs and parameters that are needed to drive the complex models. Although the model structures themselves are (generally) highly reliable, the resulting model predictions are highly uncertain due to uncertainty in the numerous model inputs (assuming that uncertainty in the model structure is of minor importance although this is not necessarily the case). This prediction uncertainty should be integrated in the decision-making process by communicating the uncertainty to the decision makers such that they dispose of additional information concerning the weight that should be given to the model predictions. Especially when the hydrological model predictions are used in early-warning systems, e.g. for flooding, uncertainty cannot be disregarded from a societal point of view. Despite the growing demand for uncertainty analysis, it has not yet become a standard practice in hydrological modelling, mainly due to the lack of guidance and due to the high computational cost. As such, there is a growing need for more research on this topic.

The aim of the dissertation was to evaluate some uncertainty issues in hydrological modelling in order to gain more insight regarding the uncertainties that are most relevant and how these should be modelled in a correct way. We focused on (i) uncertainty related to land surface characterization, *i.e.* land cover and soil inputs and parameters and (ii) simulations in poorly gauged or ungauged basins. Under the latter condition, there is limited information about the land surface characteristics available, such that the land surface characterization is prone to uncertainty. Moreover, we mainly dealt with aleatoric uncertainty (due to natural variability) in the land surface characterization.

In a first instance, the thesis presents some methodologies for uncertainty and sensitivity analysis w.r.t. uncertain soil and land cover information, including for example a method to assess and propagate uncertainty in the soil hydraulic parameters due to within-soil-class variability (*cfr.* Chapter 14) and a method to assess and propagate uncertainty in the land cover classification (*cfr.* Chapter 12). Furthermore, we developed a sensitivity analysis method for compositional data with the aim to assess the sensitivity of the simulated soil moisture towards textural changes (*cfr.* Chapter 15). In addition, we also presented some alternative methods for uncertainty assessment to meet the need for less computationally

demanding methods. Examples are the estimation of a representative set of soil parameters based on the central point of the soil class (*cfr.* Chapter 14) and the approximation of Monte Carlo-based prediction uncertainty through uncertainty proxies (*cfr.* Chapter 12). In a second instance, the results of the uncertainty and sensitivity analyses allowed to formulate guidelines about how to set up strategies for data collection in order to efficiently reduce the prediction uncertainty. For example, guidelines concerning strategies to acquire remote sensing data for the classification of land cover (*cfr.* Chapter 11) or concerning the collection of soil texture measurements to estimate the soil hydraulic parameters (*cfr.* Chapter 15). It is also indicated in which situations modelling uncertainty is relevant by comparing the uncertainty in area-averaged and local model predictions (*cfr.* Chapter 10). Finally, the performed uncertainty and sensitivity analyses also provided more insight in the hydrological model behaviour (*cfr.* Chapters 10, 12, 14, 15).

Answer to the research questions

The research objectives of this thesis are: (i) give an overview of the basic concepts and techniques for uncertainty and sensitivity analysis, (ii) perform a calibration of the spatially distributed hydrological model for the study area under consideration, (iii) evaluate the impact of uncertainty in the land cover input on the hydrological model predictions, and (iv) evaluate the impact of uncertainty in the soil input on the hydrological model predictions. These objectives were achieved through several research questions, of which the answers were already obtained throughout the different chapters, but are briefly summarized below.

Which aspects are important when setting up a framework for uncertainty or sensitivity analysis?
A successful uncertainty analysis starts with a correct identification of the type of uncertainty because it determines the selection of an appropriate technique for uncertainty modelling and propagation. One should distinguish between epistemic and aleatoric uncertainty based on the origin (or the nature) of the uncertainty, respectively incomplete knowledge and natural variability. Aleatoric uncertainty is best represented using probability distributions, whereas epistemic uncertainty is more faithfully represented using probability boxes, possibility distributions or belief functions. For sensitivity analysis, the objective of the analysis is the crucial factor when selecting an appropriate technique. Special attention should be given to the correlation between the model inputs for which a sensitivity analysis is intended because the effect of simultaneously varying model inputs might not be the same as the superposition of their individual effects.

Does the selection of the calibration method have an influence on the calibration result?
Automatic calibration of the hydrological model pointed out that different calibration methods (here, shuffled complex evolution (SCE), dynamically dimensioned search (DDS) and weight-adaptive recursive parameter estimation (WARPE) method were evaluated) might result in a different set of calibrated parameter values. Although the model performance statistics (in terms of RMSE and Pearson correlation) indicated that the parameter values have been optimized, a good correlation between the simulated and observed discharge was not guaranteed. The calibration methods (SCE and DDS) that minimize an objective function were found unsuitable for calibrating the TOPLATS model. Only the Kalman filter-based method (WARPE) which minimizes the uncertainty in the estimated parameter values was able to produce simulations that closely resembled the observed discharge. It is therefore advised to carefully select an appropriate algorithm for the automatic calibration of a spatially distributed hydrological model and to evaluate the calibration result based on a combination of model performance statistics and a visual inspection of the simulated model output.

Is more detailed information about the land cover valuable for the prediction of water fluxes?
The effect of the level of land cover information (general land cover classes versus crop type classes) depends on the purpose of the hydrological model application. If modellers are interested in area-averaged predictions of the water fluxes, adding information about the specific crop type is not very useful as the effect on the predicted fluxes is negligible. On the contrary, if local predictions of the water fluxes are of interest, it is meaningful to increase the level of land cover information because an increased heterogeneity in the land cover map results in an increased heterogeneity in the hydrological model response and as such reduces the uncertainty in the water flux predictions.

How can classification uncertainty efficiently be assessed?
Uncertainty assessment is a valuable addition to the established accuracy assessment of the classification result as it allows for better interpretating the results or for improving the classified land cover map through uncertainty-weighted post-processing of the map. The prediction probability at the pixel level can be assessed in a computationally efficient way through Random Forests. This probabilistic classifier has few user-defined model parameters and has the ability to deal with high-dimensional data sets. Based on the produced membership probability vectors, uncertainty measures $U = 1 - p_{\max}$ or the entropy H can be calculated. Both measures produce similar uncertainty assessment results but show a different behaviour of the modelled probability: the linear behaviour of U is in contrast with the nonlinear behaviour of H. The entropy H is to be preferred as it exploits the entire probability vector whereas U only uses the maximum probability from the probability vector.

How can the dimensionality of remote sensing data be optimized w.r.t. the classification uncertainty?
High-dimensional remote sensing data sets are disadvantageous as they require a lot of resources, both financially in the acquisition step and computationally in the processing step. Therefore, it is interesting to search for an optimal configuration in the dimensionality of the data set in order to minimize the resource requirements while maintaining the quality of the derived product (here, a land cover map). If a reduction of the dimensionality in terms of acquisition is intended, multi-date models are to be preferred over multi-frequency models, and early-spring acquisitions are better avoided as they compromise the model uncertainty. In case of multi-date models, the frequency is of minor importance although for single-date models, the L-band is to be preferred over C-band. Alternatively, the data set can be reduced in terms of the derived features. Depending on the purpose of the classified land cover map, an accuracy-oriented or an efficiency-oriented reduction approach should be selected. The latter approach relies on a set of general features (identified as γ_{hv}, ϕ_{hhvv}, γ_{ll}) for land cover classification, while the former approach varies the set of most important features in function of the acquisition time. Identification of the most important features can be realized based on an evaluation of the variable importance, which is easily obtained through application of Random Forests.

Is propagation of the classification uncertainty relevant for the predictability of water fluxes and how can it be performed in a computationally efficient way?
When area-averaged predictions of the water fluxes are intended, assessing and propagating the classification uncertainty is not very useful, instead, reliable water flux predictions can be easily obtained by applying the hydrological model with the hard classified land cover map. On the contrary, when a spatial distribution of the model predictions is relevant for decision-making or is coupled to other spatially distributed models, it is advised to propagate the classification uncertainty because it gives rise to locally high prediction uncertainties. A Monte Carlo-based uncertainty propagation is to be preferred but if computationally infeasible, an alternative fast indicative uncertainty assessment can be used. With this method, the mean and standard deviation of the Monte Carlo-based results are effectively approximated by means of uncertainty proxies. Although the proxies are not able to reproduce the results of the Monte Carlo-based uncertainty assessment, they provide a more reliable estimate of the water fluxes than the simulations obtained with the hard classified land cover map, thus when disregarding the classification uncertainty.

How can the within-soil-class variability of the soil hydraulic parameters be modelled and propagated through the hydrological model?
Based on a sampling of a high number of textures on the texture triangle within the given soil class, a set of soil hydraulic parameters (SHPs) can be calculated by means of continuous pedotransfer functions (PTFs). The choice of the PTF is a crucial step since intrinsic differences in the PTFs remain present after propagating

the estimated SHPs through the hydrological model. If region specific PTFs are available they are to be preferred over non-region specific PTFs. The within-soil-class variability in the SHPs is then propagated by repeatedly applying the hydrological model with the different sets of SHPs. As such, the correlation between the soil parameters is preserved. Furthermore, a representative characterization of the uncertainty requires that higher-order moments of the probability distribution are taken into account as the non-normality might be strong, especially under very dry soil moisture conditions. When only a quick assessment of the mean system behaviour is intended, it is advised not to rely on literature-based values for the SHPs, but to use the SHPs of the central point (centre of gravity of the soil class) which can be estimated by means of continuous PTFs that are specific for the region of interest.

How can a sensitivity analysis on soil texture be performed?
Soil texture is a composition of the sand, silt and clay fraction and should be treated as such in the sensitivity analysis. This means that the particle size distributions should be varied simultaneously while assuring that they sum to one. Local perturbation of the soil texture can be realized by operations in the simplex (or by the alternative approach using the ILR coordinates) and results in a perturbation circle on which a set of perturbed textures is selected. Based on the selected set of perturbed textures, a sensitivity index can be calculated by approximating the directional derivatives of the model output to the compositional model input using the finite difference technique. The choice of the perturbation factor is an important step in the sensitivity analysis and should be chosen carefully. To perturb texture, we suggest to use a perturbation factor of 10^{-2} although for a limited number of cases this value might be too low (causing numerical errors) or too high (causing errors due to nonlinearity).

Which strategies for sampling soil texture can be applied to reduce the uncertainty in the soil moisture predictions?
If only a soil type map of the basin is available, information on the exact soil texture in terms of the sand, silt and clay content is lacking. For regions in the basin where the soil map indicates the presence of coarse-textured soils, the central point of the soil class can be used as a representative soil texture for that region. For regions with fine-textured soils, it is advised to derive a representative texture from one or more soil texture measurements (considering the high sensitivity of the soil moisture simulation to these type of soils). The more the sampled texture deviates from textures with a clay and sand content around 50%, the more additional samples will be needed to determine a representative soil texture. The representative texture can then be used in the hydrological model simulation. Note that this guideline was derived from simulations with TOPLATS in which the SHPs were estimated with the Rawls and Brakensiek (1985, 1989) PTFs. For any other model set-up, we recommended to carry out the sensitivity analysis as presented in Chapter 15 under the given conditions.

Recommendations for hydrological modelling

Throughout the different chapters in the manuscript, different sources of uncertainty were treated and analyzed. The results can be integrated and generalized in order to improve future hydrological modelling studies. Based on the answer to the research questions, following rules-of-thumb can be formulated:

1. Identify the purpose of the modelling study and indicate which level of detail for the water flux predictions is needed. This can save a lot of effort because not all sources of uncertainty are equally relevant under different conditions: **only assess uncertainty if relevant.** For example, if area-averaged predictions are intended the uncertainty in the land cover information is of minor importance. This means that in this case less detailed land cover data, e.g. readily available general land cover maps (as provided by international agencies or local government) might offer a good solution, and that assessing classification uncertainty is not useful.

2. When a land cover map (or another type of map) needs to be derived from remote sensing data (because there is no map available or does not have the required level of detail for local predictions), the acquisition of remote sensing data needs to be well-considered. A higher dimensionality in the input data does not necessarily lead towards a lower uncertainty in the classified land cover map. Disinformative data can mortgage the classification process such that removing the least important inputs from the classification process cause the classification uncertainty to decrease: **less is more, only collect informative data.** For land cover classification, especially the time of acquisition is a crucial step w.r.t. remote sensing data and needs to be carefully selected in function of the growing season.

3. Soil information is an important source of uncertainty (more important than land cover information) in physically-based hydrological models such that the collection of soil data deserves special attention. A soil type map of the region as single source of soil information induces a very high uncertainty in the model prediction because there is a high variation of the textural composition within soil classes and soil classes themselves are not the best way to classify soils w.r.t. their hydrological properties. To obtain reliable model predictions, it is necessary to **collect at least one soil texture sample within each soil map unit** in order to determine a representative particle size distribution. The sampling density is a function of the soil type: more clayey soils will require more samples. A set of representative soil properties can be derived from the representative soil texture, preferably by means of continuous pedotransfer functions that were developed for the specific region of interest as non-region specific pedotransfer functions have a dubious reliability and produce substantially different estimates.

4. The larger the modelling scale, the less uncertainty is involved: errors at the point scale compensate each other when averaged over the catchment such that the resulting error at the catchment scale is significantly lower. This effect is quite similar to the increasing complexity of natural processes at the finer scale, which is in contrast with the relative simplicity of large scale processes. As such, **uncertainty bounds should be interpreted with regard to the modelling scale.** The uncertainty at the point scale can be considered as an upper boundary of uncertainty at the catchment scale and represents the uncertainty as if the catchment was treated as a single hydrological unit. Increasing the heterogeneity in the description of the land surface increases the heterogeneity of the local model response, but also implies a heterogeneity in the input error such that local errors compensate each other when averaged over the catchment and ultimately reduce the error in the catchment scale model response.

Future perspectives

As mentioned earlier, uncertainty is an important aspect in hydrological modelling and should not be disregarded for a good interpretation of the model results. The results of this thesis and of previous research provide us with more insight regarding uncertainty modelling but are not sufficient to come up with a decent framework to identify the best method for uncertainty modelling and the best strategy for uncertainty reduction. More research is required to realize this.

In this dissertation, we only dealt with a few of the numerous sources of uncertainty that are present in distributed hydrological models. Future research should also deal with the remaining sources of uncertainty. Furthermore, we focused on aleatoric uncertainties, while epistemic uncertainty makes up a big part of the total uncertainty in hydrological modelling. A crucial question that remains unsolved is how to deal with epistemic uncertainties because these uncertainties cannot be modelled through probability theory and are subjective. This makes it difficult to correctly quantify and interpret epistemic uncertainty such that this type of uncertainty is often avoided and not taken into account in hydrological modelling. Future research should come up with solutions for this matter.

Throughout the different chapters, the sources of uncertainty were treated individually, while it would be interesting to know what happens to the model output if all input uncertainty is jointly propagated through the hydrological model. Would it result in a prediction uncertainty that is practically feasible for decision-making or would the uncertainty bounds explode and make the hydrological prediction worthless? Future research should aim at the evaluation of prediction uncertainty as the result of the simulatenous effect of the main sources of uncertainty.

Considering the large prediction uncertainty associated with a single source of uncertainty, it is very plausible that the total prediction uncertainty is very large. This raises the question whether spatially distributed physically-based models are the best solution to model the hydrological fluxes. In theory, this type of models involves a better representation of reality than the alternative lumped models, but the high number of input variables and parameters undermines the strong model structure. It should be investigated whether less complex models are not a better alternative than spatially distributed models for predictions in ungauged basins as it is easier to restrict the input uncertainty in the former type of models.

Bibliography

Aitchison, J. (1982). The statistical-analysis of compositional data. *Journal of the Royal Statistical Society Series B-Methodological*, 44(2):139–177.

Aitchison, J. (1983). Principal components analysis of compositional data. *Biometrika*, 70(1):57–65.

Aitchison, J. (1986). The Statistical Analysis of Compositional Data. In *Monographs on Statistics and Applied Probability*, page 416. Chapman and Hall Ltd, London (UK).

Aitchison, J. (1992). On criteria for measures of compositional difference. *Mathematical Geology*, 24(4):365–379.

Aitchison, J. and Egozcue, J. J. (2005). Compositional data analysis: Where are we and where should we be heading? *Mathematical Geology*, 37(7):829–850.

Aitchison, J. and Ng, K. W. (2005). The role of perturbation in compositional data analysis. *Statistical Modelling*, 5(2):173–185.

Ajami, N., Gupta, H., Wagener, T., and Sorooshian, S. (2004). Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology*, 298(1-4):112–135.

Alberga, V. (2007). A study of land cover classification using polarimetric SAR parameters. *International Journal of Remote Sensing*, 28(17):3851–3870.

Alberga, V., Satalino, G., and Staykova, D. K. (2008). Comparison of polarimetric SAR observables in terms of classification performance. *International Journal of Remote Sensing*, 29(14):4129–4150.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1988). Crop evapotranspiration - guidelines for computing crop water requirements. Technical report, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

Armstrong, R. N. and Martz, L. W. (2008). Effects of reduced land cover detail on hydrological model response. *Hydrological Processes*, 22(14):2395–2409.

Ascough, I. I. J. C., Green, T. R., Ma, L., and Ahjua, L. R. (2005). Key criteria and selection of sensitivity analysis methods applied to natural resource models. In Zerger, A. and Argent, R. M., editor, *MODSIM 2005: International Congress on Modelling and Simulation: Advances and applications for management and decision making*, pages 2463–2469.

Ayyub, B. M. (2001). *Elicitation of expert opinions for uncertainty and risks*. CRC Press, New York.

Barth, G. R., Hill, M. C., Illangasekare, T. H., and Rajaram, H. (2001). Predictive modeling of flow and transport in a two-dimensional intermediate-scale, heterogeneous porous medium. *Water Resources Research*, 37(10):2503–2512.

Baudrit, C. and Dubois, D. (2006). Practical representations of incomplete probabilistic knowledge. *Computational Statistics & Data Analysis*, 51(1):86–108.

Baudrit, C., Guyonnet, D., and Dubois, D. (2007). Joint propagation of variability and imprecision in assessing the risk of groundwater contamination. *Journal of Contaminant Hydrology*, 93(1-4):72–84.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of changes. *Philosophical Transactions*, 53:370–418.

Benke, K. K., Lowell, K. E., and Hamilton, A. J. (2008). Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. *Mathematical and Computer Modelling*, 47(11-12):1134–1149.

Bernier, J. (1987). Elements of Bayesian analysis of uncertainty in reliability and risk models. In Duckstein, L. and Plate, E. J., editors, *Engineering reliability and risk in water resources*, volume 124 of *NATO ASI Series E: Applied Science*, pages 405–422. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modeling. *Advances in Water Resources*, 16(1):41–51.

Beven, K. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1):1–12.

Beven, K. (2006a). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1-2, SI):18–36.

Beven, K. (2006b). On undermining the science? *Hydrological Processes*, 20(14):3141–3146.

Beven, K. and Binley, A. (1992). The future of distributed models - model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298.

Beven, K. and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4):11–29.

Beven, K. J. and Kirkby, M. J. (1979). A physically based variable contributing area model of basin hydrology. *Hydrologic Science Bulletin*, 24(1):43–69.

Blaes, X., Vanhalle, L., and Defourny, P. (2005). Efficiency of crop identification based on optical and SAR image time series. *Remote Sensing of Environment*, 96(3-4):352–365.

Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modeling - a review. *Hydrological Processes*, 9(3-4):251–290.

Boateng, S. and Cawlfield, J. D. (1999). Two-dimensional sensitivity analysis of contaminant transport in the unsaturated zone. *Ground Water*, 37(2):185–193.

Bormann, H. (2007). Analysis of the suitability of the German soil texture classification for the regional scale application of physical based hydrological model. *Advances in Geoscience*, 11:7–131.

Bormann, H. (2008). Sensitivity of a soil-vegetation-atmosphere-transfer scheme to input data resolution and data classification. *Journal of Hydrology*, 351(1-2):154–169.

Bouma, J. (1989). Using soil survey data for quantitative land evaluation. *Advances in Soil Science*, 9:177–213.

Boyle, D. P., Gupta, H. V., and Sorooshian, S. (2000). Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*, 36(12):3663–3674.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., editors (1984). *Classification and Regression Trees*. Chapman and Hall, New York, 2nd edition.

Breshears, D. D., Kirchner, T. B., and Whicker, F. W. (1992). Contaminant transport through agroecosystems - assessing relative importance of environmental, physiological, and management factors. *Ecological Applications*, 2(3):285–297.

Breuer, L., Eckhardt, K., and Frede, H. G. (2003). Plant parameter values for models in temperate climates. *Ecological Modelling*, 169(2-3):237–293.

Breuer, L., Huisman, J. A., and Frede, H. G. (2006). Monte Carlo assessment of uncertainty in the simulated hydrological response to land use change. *Environmental Modeling & Assessment*, 11(3):209–218.

Brooks, R. H. and Corey, A. (1964). Hydraulic properties of porous media. Hydrology Paper No 3, Colorado State University.

Brown, J. D. (2004). Knowledge, uncertainty and physical geography: towards the development of methodologies for questioning belief. *Transactions of the Institute of British Geographers*, 29(3):367–381.

Brown, K. M., Foody, G. M., and Atkinson, P. M. (2009). Estimating per-pixel thematic uncertainty in remote sensing classifications. *International Journal of Remote Sensing*, 30(1):209–229.

Brutsaert, W. (1973). Review of Green's functions for linear open channels. *Journal of Engineering Mechanics Division ASCE*, 99:1247–1257.

Buchanan, J. T., Henig, E. J., and Henig, M. I. (1998). Objectivity and subjectivity in the decision making process. *Annals of Operations Research*, 80:333–345.

Bylander, T. (2002). Estimating generalization error on two-class datasets using out-of-bag estimates. *Machine Learning*, 48:287–297.

Campbell, G. S. (1977). *An introduction to environmental biophysics*. Springer-Verlag, New York, USA.

Campolongo, F., Cariboni, J., and Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10):1509–1518.

Camps-Valls, G., Mooij, J., and Schoelkopf, B. (2010). Remote sensing feature selection by kernel dependence measures. *IEEE Geoscience and Remote Sensing Letters*, 7(3):587–591.

Canters, F., De Genst, W., and Dufourmont, H. (2002). Assessing effects of input uncertainty in structural landscape classification. *International Journal of Geographical Information Science*, 16(2):129–149.

Chen, F., Mitchell, K., Schaake, J., Xue, Y. K., Pan, H. L., Koren, V., Duan, Q. Y., Ek, M., and Betts, A. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research-Atmospheres*, 101(D3):7251–7268.

Chow, V. T., Maidment, D. R., and Mays, L. W. (1988). *Applied Hydrology*. McGraw-Hill, USA.

Christensen, E. and Dall, J. (2002). EMISAR: a dual-frequency, polarimetric airborne SAR. In *Proceedings IGARSS 2002: IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing, Remote Sensing: Integrating our view of the planet*, volume 3, pages 1711–1713.

Christensen, E. L., Skou, N., Dall, J., Woelders, K., Jorgensen, J. H., Granholm, J., and Madsen, S. N. (1998). EMISAR: An absolutely calibrated polarimetric L- and C-band SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 36(6):1852–1865.

Christiaens, K. and Feyen, J. (2002). Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrological Processes*, 16(2):373–391.

Cloke, H. L., Pappenberger, F., and Renaud, J. P. (2008). Multi-Method Global Sensitivity Analysis (MMGSA) for modelling floodplain hydrological processes. *Hydrological Processes*, 22(11):1660–1674.

Cloude, S. R. and Pottier, E. (1997). An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1):68–78.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46.

Cornelis, W. M., Ronsyn, J., Van Meirvenne, M., and Hartmann, R. (2001). Evaluation of pedotransfer functions for predicting the soil moisture retention curve. *Soil Science Society of America Journal*, 65(3):638–648.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, 20(6):682–690.

Crawford, N. H. and Linsley, R. K. (1966). Digital simulation in hydrology, Stanford Watershed Model IV. Technical report, Department of Civil Engineering, Stanford University, USA.

Croke, B., Merritt, W., and Jakeman, A. (2004). A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments. *Journal of Hydrology*, 291(1-2):115–131.

Crosetto, M., Ruiz, J., and Crippa, B. (2001a). Uncertainty propagation in models driven by remotely sensed data. *Remote Sensing of Environment*, 76(3):373–385.

Crosetto, M., Ruiz, J. A. M., and Crippa, B. (2001b). Uncertainty propagation in models driven by remotely sensed data. *Remote Sensing of Environment*, 76(3):373–385.

Crosetto, M., Tarantola, S., and Saltelli, A. (2000). Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture Ecosystems & Environment*, 81(1):71–79.

Crow, W. T., Drusch, M., and Wood, E. F. (2001). An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8):1622–1631.

Cryer, S. A. and Havens, P. L. (1999). Regional sensitivity analysis using a fractional factorial method for the USDA model GLEAMS. *Environmental Modelling & Software*, 14(6):613–624.

Cullen, A. C. and Frey, H. C. (1999). *Probabilistic techniques in exposure assessment.* Plenum Press, New York, USA.

Dawdy, D. R. and O'Donnell, T. (1965). Mathematical models of catchment behavior. *Journal of Hydraulic Engineering - ASCE*, 91:123–137.

De Fries, R. S. and Los, S. O. (1999). Implications of land-cover misclassification for parameter estimates in global land-surface models: An example from the simple biosphere model (SiB2). *Photogrammetric Engineering and Remote Sensing*, 65(9):1083–1088.

De Pauw, D. J. W. and Vanrolleghem, P. A. (2006). Practical aspects of sensitivity function approximation for dynamic models. *Mathematical and Computer Modelling of Dynamical Systems*, 12(5):395–414.

Deardorff, J. W. (1978). Efficient prediction of ground surface-temperature and moisture, with inclusion of a layer of vegetation. *Journal of Geophysical Research-Oceans and Atmospheres*, 83(NC4):1889–1903.

Del Frate, F., Licciardi, G., and Duca, R. (2009). Autoassociative neural networks for features reduction of hyperspectral data. In *Proceedings of 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 62–65.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multi-valued mapping. *Annals of Mathematical Statistics*, 38(2):325–339.

Deng, H., Ye, M., Schaap, M. G., and Khaleel, R. (2009). Quantification of uncertainty in pedotransfer function-based parameter estimation for unsaturated flow modeling. *Water Resources Research*, 45:W04409.

Dettinger, M. D. and Wilson, J. L. (1981). 1st-order analysis of uncertainty in numerical-models of groundwater-flow: 1 mathematical development. *Water Resources Research*, 17(1):149–161.

Dickinson, R. E. (1984). Modeling evapotranspiration for three-dimensional climate models. In Hanson, J. E. and Takahashi, T., editors, *Climate Processes and Climate Sensitivity*, volume 29 of *Geophysical Monograph Series*, pages 58–72. AGU, Washington, D.C.

Dierking, W. and Skriver, H. (2002). Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):618–636.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society Series B-Methodological*, 57(1):45–97.

Duan, Q. (2003). Global optimization for watershed model calibration. In Duan, Q., editor, *Calibration of watershed models*, volume 6 of *Water Science and Application*, pages 89–104. AGU, Washington.

Duan, Q., Sorooshian, S., and Gupta, V. K. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 24(4):1015–1031.

Duan, Q. Y., Sorooshian, S., and Gupta, V. K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158(3-4):265–284.

Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis*, 51(1):47–69.

Dubois, D., Foulloy, L., and Mauris, G. (2004). Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10:273–297.

Dubois, D., Nguyen, H. T., and Prade, H. (2000). Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps. In Dubois, D. and Prade, H., editors, *Fundamentals of fuzzy sets*, pages 343–438. Kluwer Academic Publishers, Boston.

Eagleson, P. S. (1978). Climate, soil, and vegetation. 3. Simplified model of soil-moisture movement in liquid-phase. *Water Resources Research*, 14(5):722–730.

Eckhardt, K., Breuer, L., and Frede, H. G. (2003). Parameter uncertainty and the significance of simulated land use change effects. *Journal of Hydrology*, 273(1-4):164–176.

Egozcue, J. J. (2012). Interactive comment on "Local sensitivity analysis for compositional data with application to soil texture in hydrologic modelling" by L. Loosvelt et al. *Hydrology and Earth System Sciences Discussion*, 9:C4913–C4926.

Egozcue, J. J. and Pawlowsky-Glahn, V. (2006). Simplical geometry for compositional data. In *Compositional Data Analysis in the Geoscience: From Theory to Practice*, pages 145–159. Geological Society, London (UK).

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barcelo-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3):279–300.

Ek, M. and Mahrt, L. (1991). OSU 1D PBL Model: User's Guide, version 1.0.4. Technical report, Department of Atmospheric Sciences, Oregon State University, Corvallis.

Everitt, B. S. (1992). *The Analysis of Contingency Tables*. Chapman and Hall, London, 2nd edition.

Famiglietti, J. S. and Wood, E. F. (1994). Multiscale modeling of spatially variable water and energy balance processes. *Water Resources Research*, 30(11):3061–3078.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

Ferrazzoli, P. (2002). SAR for agriculture: Advances, problems and prospects. In Wilson, A, editor, *Proceedings of the Third International Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, volume 475 of *ESA SP Publications*, pages 47–56.

Ferreira, V. A., Weesies, G. A., Yoder, D. C., Foster, G. R., and Renard, K. G. (1995). The site and condition specific nature of sensitivity analysis. *Journal of Soil and Water Conservation*, 50(5):493–497.

Ferro-Famil, L., Pottier, E., and Lee, J.-S. (2001). Unsupervised classification of multifrequency and fully polarimetric SAR images based on the H/A/Alpha-Wishart Classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11):2332–2342.

Ferson, S. and Ginzburg, L. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2-3):133–144.

Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D. M., and Sentz, K. (2003). Construction of probability boxes and Demster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratories, Albuquerque, California.

Finke, P. A. and Wösten, J. H. M. (1996). Effects of uncertainty in major input variables on simulated functional soil behaviour. *Hydrological Processes*, 10(5):661–669.

Finn, J. T. (1993). Use of the average mutual information index in evaluating classification error and consistency. *International Journal of Geographical Information Systems*, 7(4):349–366.

Foglia, L., Hill, M. C., Mehl, S. W., and Burlando, P. (2009). Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function. *Water Resources Research*, 45:W06427.

Foody, G. M. (1995). Cross-entropy for the evaluation of the accuracy of a fuzzy land-cover classification with fuzzy ground data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(5):2–12.

Foody, G. M. (2006). What is the difference between two maps? A remote senser's view. *Journal of Geographical Systems*, 8(2):119–130.

Foody, G. M. and Atkinson, P. M. (2002). *Uncertainty in remote sensing and GIS*. Wiley, Chichester.

Frate, F. D., Shiavon, G., Solimini, D., Borgeaud, M., Hoekman, D. H., and Vissers, M. A. M. (2003). Crop classification using multiconfiguration C-band SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7):1611–1619.

Frey, H. C. and Patil, S. R. (2002). Identification and review of sensitivity analysis methods. *Risk Analysis*, 22(3):553–578.

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232.

Fritz, S. and See, L. (2005). Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, 19(7):787–807.

Fritz, S. and See, L. (2008). Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, 14(5):1057–1075.

Funtowicz, S. O. and Ravetz, J. R. (1990). *Uncertainty and quality in science for policy*. Kluwer, Dordrecht, the Netherlands.

Gardner, W. R. (1958). Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Science*, 85:228–232.

Gevrey, M., Dimopoulos, I., and Lek, S. (2006). Two-way interaction of input variables in the sensitivity analysis of neural network models. *Ecological Modelling*, 195(1-2):43–50.

Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4):294–300.

Goncalves, L. M. S., Fonte, C. C., Julio, E. N. B. S., and Caetano, M. (2009). A method to incorporate uncertainty in the classification of remote sensing images. *International Journal of Remote Sensing*, 30(20):5489–5503.

Goodchild, M. F., Guoqing, S., and Shiren, Y. (1992). Development and test of an error model for categorical data. *International Journal of Geographical Information Systems*, 6(2):87–104.

Grayson, R. B., Moore, I. D., and McMahon, T. A. (1992). Physically based hydrologic modeling, 2. Is the concept realistic? *Water Resources Research*, 28(10):2659–2666.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34(4):751–763.

Gutmann, E. D. and Small, E. E. (2005). The effect of soil hydraulic properties vs. soil texture in land surface models. *Geophysical Research Letters*, 32(2):L02402.

Gwo, J. P., Toran, L. E., Morris, M. D., and Wilson, G. V. (1996). Subsurface stormflow modeling with sensitivity analysis using a Latin-hypercube sampling technique. *Ground Water*, 34(5):811–818.

Hack, H., Bleiholder, H., Buhr, L., Meier, U., Schnock-Fricke, U., Weber, E., and Witzenberger, A. (1992). Einheitliche codierung der phnologischen entwicklungsstadien mono- und dikotyler pflanzen erweiterte bbch-skala, allgemein -. *Nachrichtenbl. Deut. Pflanzenschutzd.*, 44:265–270.

Haralick, R. M., Shanmuga, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems Man and Cybernetics*, SMC3(6):610–621.

Harremoes, P. and Madsen, H. (1999). Fiction and reality in the modelling world - Balance between simplicity and complexity, calibration and identifiability, verification and falsification. *Water Science and Technology*, 39(9):1–8.

Hastings, W. K. (1970). Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–&.

Helton, J. C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive-waste disposal. *Reliability Engineering & System Safety*, 42(2-3):327–367.

Helton, J. C. and Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69.

Heuvelink, G. B. M. and Burrough, P. A. (2002). Developments in statistical approaches to spatial uncertainty and its propagation. *International Journal of Geographical Information Science*, 16(2):111–113.

Heuvelmans, G., Muys, B., and Feyen, J. (2004). Evaluation of hydrological model parameter transferability for simulating the impact of land use on catchment hydrology. *Physics and Chemistry of the Earth*, 29(11-12):739–747.

Hill, M. C. and Tiedeman, C. R. (2007). *Effective groundwater model calibration: With analysis of data, sensitivities, predictions and uncertainty.* John Wiley, New York.

Hoekman, D. H. and Vissers, M. A. M. (2003). A new polarimetric classification approach evaluated for agricultural crops. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12):2881–2889.

Hoeting, J., Madigan, D., Raftery, A., and Volinsky, C. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, 14(4):382–401.

Hofer, E. (1999). Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. *Computer Physics Communications*, 117(1-2):21–34.

Hoffmann-Riem, H. and Wynne, B. (2002). In risk assessment, one has to admit ignorance - Explaining there are things we can't know could improve public confidence in science. *Nature*, 416(6877):123.

Hogue, T. S., Sorooshian, S., Gupta, H., Holz, A., and Braatz, D. (2000). A multistep automatic calibration scheme for river forecasting models. *Journal of Hydrometeorology*, 1(6):524–542.

Hopmans, J. W., Simunek, J., Romano, N., and Durner, W. (2002). Simultaneous determination of water transmission and retention properties. Inverse methods. In Dane, J. H. and Topp, G. C., editors, *Methods of Soil Analysis*, pages 963–1008. Soil Science Society of America Book Series No. 5.

Hossain, F. and Anagnostou, E. (2005). Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture. *Advances in Water Resources*, 28(12):1336–1350.

Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3):651–674.

Hwang, J. T., Dougherty, E. P., Rabitz, S., and Rabitz, H. (1978). Greens function method of sensitivity analysis in chemical-kinetics. *Journal of Chemical Physics*, 69(11):5180–5191.

Ibrahim, M. A., Arora, M. K., and Ghosh, S. K. (2005). Estimating and accommodating uncertainty through the soft classification of remote sensing data. *International Journal of Remote Sensing*, 26(14):2995–3007.

Iman, R. L. and Helton, J. C. (1988). An investigation of uncertainty and sensitivity analysis techniques for computer-models. *Risk Analysis*, 8(1):71–90.

Jacquemin, B. and Noilhan, J. (1990). Sensitivity study and validation of a land surface parameterization using the Hapex-Mobilhy data set. *Boundary-Layer Meteorology*, 52(1-2):93–134.

Jacques, J., Lavergne, C., and Devictor, N. (2006). Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering & System Safety*, 91(10-11):1126–1134.

Jacquin, A. P. (2010). Possibilistic uncertainty analysis of a conceptual model of snowmelt runoff. *Hydrology and Earth System Sciences*, 14(8):1681–1695.

Jajarmizadeh, M., Harun, S., and Salarpour, M. (2012). A review on theoretical consideration and types of models in hydrology. *Journal of Environmental Science and Technology*, 5(5):249–261.

Jamison, K. and Lodwick, W. (2002). The construction of consistent possibility and necessity measures. *Fuzzy Sets and Systems*, 132(1):1–10.

Jansen, M. J. W. (1998). Prediction error through modelling concepts and uncertainty from basic data. *Nutrient Cycling in Agroecosystems*, 50(1-3):247–253.

Jarvis, P. G., James, G. B., and Landsberg, J. J. (1976). Coniferous forest. In Monteith, J. L., editor, *Vegetation and the atmosphere*, volume 2, pages 171–240. Academic Press, London.

Jenson, S. K. and Domingue, J. O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing*, 54(11):1593–1600.

Jones, H. G. (1992). *Plants and Microclimate: A Quantitative Approach to Environmental Plant Physiology*, volume 56 of *Irrigation and drainage paper*. Cambridge University Press, Cambridge, 2nd edition.

Kaiser, K. and Guggenberger, G. (2003). Mineral surfaces and soil organic matter. *European Journal of Soil Science*, 54(2):219–236.

Kasiviswanathan, K. S. and Sudheer, K. P. (2013). Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. *Stochastic Environmental Research and Risk Assessment*, 27(1):137–146.

Kleijnen, J. P. C., Vanham, G., and Rotmans, J. (1992). Techniques for sensitivity analysis of simulation-models - A case-study of the CO2 greenhouse-effect. *Simulation*, 58(6):410–417.

Klir, G. J. and Folger, T. A. (1988). *Fuzzy sets, uncertainty and information*. Prentice Hall, New Jersey, USA.

Kotoda, K. (1986). Estimation of river basin evapotranspiration. In *Environmental Research Center Papers No. 8*, volume 8, page 66. Environmental Research Center, University of Tsukuba, Ibaraki, Japan.

Kuczera, G. (1988). On the validity of 1st-order prediction limits for conceptual hydrologic-models. *Journal of Hydrology*, 103(3-4):229–247.

Kuczera, G. and Parent, E. (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*, 211(1-4):69–85.

Kyriakidis, P. C. and Dungan, J. L. (2001). A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological model predictions. *Environmental and Ecological Statistics*, 8(4):311–330.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174.

Lee, J.-S., Grunes, M. R., Ainsworth, T. L., Du, L.-J., Schuler, D. L., and Cloude, S. R. (1999). Unsupervised classification using polarimetric decomposition and the Complex Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2249–2258.

Lee, J.-S., Grunes, M. R., and Pottier, E. (2001). Quantitative comparison of classification capability: fully polarimetric versus dual and single-polarization SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11):2343–2351.

Lenhart, T., Eckhardt, K., Fohrer, N., and Frede, H. (2002). Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth*, 27(9-10):645–654.

Li, K. Y., Coe, M. T., Ramankutty, N., and De Jong, R. (2007). Modeling the hydrological impact of land-use change in West Africa. *Journal of Hydrology*, 337(3-4):258–268.

Li, W. and Zhang, C. (2011). A Markov Chain geostatistical framework for land-cover classification with uncertainty assessment based on expert-interpreted pixels from remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8):2983–2992.

Liang, X. and Guo, J. Z. (2003). Intercomparison of land-surface parameterization schemes: Sensitivity of surface energy and water fluxes to model parameters. *Journal of Hydrology*, 279(1-4):182–209.

Liaw, A. and Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3):18–22.

Lin, H., Chen, J. S., Pei, Z. Y., Zhang, S. L., and Hu, X. Z. (2009). Monitoring sugercane growth using ENVISAT ASAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8):2572–2580.

Liu, Y. and Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, 43(7):W07401.

Livne, E. and Svoray, T. (2011). Components of uncertainty in primary production model: the study of DEM, classification and location error. *International Journal of Geographical Information Science*, 25(3, SI):473–488.

Loosvelt, L., Vernieuwe, H., Pauwels, V. R. N., De Baets, B., and Verhoest, N. E. C. (2013). Local sensitivity analysis for compositional data with application to soil texture in hydrologic modelling. *Hydrology and Earth System Sciences*, 17(2):461–478.

Ma, L., Ascough, J., Ahuja, L., Shaffer, M., Hanson, J., and Rojas, K. (2000). Root zone water quality model sensitivity analysis using Monte Carlo simulation. *Transactions of the ASAE*, 43(4):883–895.

Madsen, H., Wilson, G., and Ammentrop, H. (2002). Comparison of different automated strategies for calibration of rainfall-runoff models. *Journal of Hydrology*, 261(1-4):48–59.

Madsen, S. N., Christensen, E. L., Skou, N., and Dall, J. (1991). The Danish SAR system: design and initial tests. *IEEE Transactions on Geoscience and Remote Sensing*, 29(3):417–476.

Mahrt, L. and Pan, H. (1984). A 2-layer model of soil hydrology. *Boundary-layer Meteorology*, 29(1):1–20.

Martin-Clouaire, R., Cazemier, D., and Lagacherie, P. (2000). Representing and processing uncertain soil information for mapping soil hydrological properties. *Computers and Electronics in Agriculture*, 29(1-2):41–57.

Maskey, S. (2004). *Modelling uncertainty in flood forecasting systems.* Taylor & Francis, London, UK.

Maskey, S., Guinot, V., and Price, R. (2004). Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. *Advances in Water Resources*, 27(9):889–898.

Massmann, C. and Holzmann, H. (2012). Analysis of the behavior of a rainfall-runoff model using three global sensitivity analysis methods evaluated at different temporal scales. *Journal of Hydrology*, 475:97–110.

Matheussen, B., Kirschbaum, R., Goodman, I., O'Donnell, G., and Lettenmaier, D. (2000). Effects of land cover change on streamflow in the interior Columbia River Basin (USA and Canada). *Hydrological Processes*, 14(5):867–885.

Mayer, L. M., Schick, L. L., Hardy, K. R., Wagal, R., and McCarthy, J. (2004). Organic matter in small mesopores in sediments and soils. *Geochimica et Cosmochimica Acta*, 68(19):3863–3872.

McBratney, A. B., Minasny, B., Cattle, S. R., and Vervoort, R. W. (2002). From pedotransfer functions to soil inference systems. *Geoderma*, 109(1-2):125–152.

McIver, D. K. and Friedl, M. A. (2001). Estimating pixel-scale land cover classification confidence using nonparametric machine learning methods. *IEEE Transactions on Geoscience and Remote Sensing*, 39(9):1959–1968.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

McLaughlin, D. (2002). An integrated approach to hydrologic data assimilation: interpolation, smooting and filtering. *Advances in Water Resources*, 25(8-12):1275–1286.

McNairn, H., Shang, J., Jiao, X., and Champagne, C. (2009). The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification. *IEEE Transactions on Geoscience and Remote Sensing*, 47(12):3981–3992.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153157.

McRae, G. J., Tilden, J. W., and Seinfeld, J. H. (1982). Global sensitivity analysis - A computational implementation of the Fourier amplitude sensitivity test (FAST). *Computers & Chemical Engineering*, 6(1):15–25.

Meier, U., editor (2001). *Growth stages of mono- and dicotyledonous plants*, volume 3. BBCH Monograph, German Federal Biological Research Centre for Agriculture and Forestry.

Meyer, P., Rockhold, M., and Gee, G. (1997). Uncertainty analysis of infiltration and subsurface flow and transport for SDMP sites. *NUREG/CR-6565; U.S. Nuclear Regulatory Commission.*

Miller, S. N., Guertin, D. P., and Goodrich, D. C. (2007). Hydrologic modeling uncertainty resulting from land cover misclassification. *Journal of the American Water Resources Association*, 43(4):1065–1075.

Milly, P. C. D. (1986). An event-based simulation-model of moisture and energy fluxes at a bare soil surface. *Water Resources Research*, 22(12):1680–1692.

Milly, P. C. D. (1991). A refinement of the combination equations for evaporation. *Surveys in Geophysics*, 12(1-3):145–154.

Minasny, B. and McBratney, A. B. (2002). Uncertainty analysis for pedotransfer functions. *European Journal of Soil Science*, 53(3):417–429.

Mittal, G. and Singh, D. (2008). Critical analysis of polarimetric PALSAR data for landcover classification. In *International Conference on Recent Advances in Microwave Theory and Applications, Proceedings*, pages 434–437.

Montanari, A. (2007). What do we mean by "uncertainty"? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes*, 21(6):841–845.

Montanari, A., Shoemaker, C. A., and van de Giesen, N. (2009). Introduction to special section on uncertainty assessment in surface and subsurface hydrology: An overview of issues and challenges. *Water Resources Research*, 45:W00B00.

Moore, R. J. and Clarke, R. T. (1981). A distribution function-approach to rainfall runoff modeling. *Water Resources Research*, 17(5):1367–1382.

Muleta, M. and Nicklow, J. (2005). Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology*, 306(1-4):127–145.

Mulvany, T. J. (1850). On the use of self registering rain and flood gauges. *Proceedings of the Institution of Civil Engineers*, 4(2):1–8.

Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: 1. A discussion of principles. *Journal of Hydrology*, 10:282–290.

Nelder, J. A. and Mead, R. (1965). A simplex-method for function minimization. *Computer Journal*, 7(4):308–313.

Nemes, A., Timlin, D. J., Pachepsky, Y. A., and Rawls, W. J. (2009). Evaluation of the Rawls et al. (1982) pedotransfer functions for their applicability at the US national scale. *Soil Science Society of America Journal*, 73(5):1638–1645.

Neuman, S. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment*, 17(5):291–305.

Nguyen, H. T. (1978). Note on extension principle for fuzzy sets. *Journal of Mathematical Analysis and Applications*, 64(2):369–380.

Ni-Meister, W. (2008). Recent advances on soil moisture data assimilation. *Physical Geography*, 29(1):19–37.

Nicodemus, K. K. and Malley, J. D. (2009). Predictor correlation impacts machine learning algorithms: implications for genomic studies. *Bioinformatics*, 25(15):1884–1890.

Nicodemus, K. K., Malley, J. D., Strobl, C., and Ziegler, A. (2010). The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinformatics*, 11(110).

Nicodemus, K. K. and Shugart, Y. Y. (2007). Impact of linkage disequilibrium and effect size on the ability of machine learning methods to detect epistasis in case-control studies. *Genetic Epidemiology*, 31(6):611.

Oleson, K. W., Driese, K. L., Maslanik, J. A., Emery, W. J., and Reiners, W. A. (1997). The sensitivity of a land surface parameterization scheme to the choice of remotely sensed land-cover datasets. *Monthly Weather Review*, 125(7):1537–1555.

Ott, B. and Uhlenbrook, S. (2004). Quantifying the impact of land-use changes at the event and seasonal time scale using a process-oriented catchment model. *Hydrology and Earth System Sciences*, 8(1):62–78.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1):217–222.

Pan, F., Zhu, J., Ye, M., Pachepsky, Y. A., and Wu, Y.-S. (2011). Sensitivity analysis of unsaturated flow and contaminant transport with correlated parameters. *Journal of Hydrology*, 397(3-4):238–249.

Pappenberger, F. and Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5):W05302.

Pappenberger, F., Frodsham, K., Beven, K., Romanowicz, R., and Matgen, P. (2007). Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrology and Earth System Sciences*, 11(2):739–752.

Pappenberger, F., Harvey, H., Beven, K., Hall, J., and Meadowcroft, I. (2006a). Decision tree for choosing an uncertainty analysis methodology: a wiki experiment. *Hydrological Processes*, 20(17):3793–3798.

Pappenberger, F., Iorgulescu, I., and Beven, K. J. (2006b). Sensitivity analysis based on regional splits and regression trees (SARS-RT). *Environmental Modelling & Software*, 21(7):976–990.

Park, N.-W. and Chi, K.-H. (2006). Land-cover classification using multitemporal/polarization C-band SAR data. In *2006 IEEE International Geoscience and Remote Sensing Symposium, vols 1-8*, IEEE International Symposium on Geoscience and Remote Sensing (IGARSS), pages 188–191.

Park, N. W. and Chi, K. H. (2008). Integration of multitemporal/polarization C-band SAR data sets for land-cover classification. *International Journal of Remote Sensing*, 29(16):4667–4688.

Passerat de Silans, A., Bruckler, L., Thony, J. L., and Vauclin, M. (1989). Numerical modeling of coupled heat and water flows during drying in a stratified bare soil-comparison with field observations. *Journal of Hydrology*, 105:109–138.

Pauwels, V. and Wood, E. (2000). The importance of classification differences and spatial resolution of land cover data in the uncertainty in model results over boreal ecosystems. *Journal of Hydrometeorology*, 1(3):255–266.

Pauwels, V. R. N. (2008). A multistart weight-adaptive recursive parameter estimation method. *Water Resources Research*, 44(4):W04416.

Pauwels, V. R. N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N. E. C., and Mattia, F. (2009). Optimization of soil hydraulic model parameters using synthetic aperture radar data: An integrated multidisciplinary approach. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2):455–467.

Pauwels, V. R. N., Verhoest, N. E. C., De Lannoy, G. J. M., Guissard, V., Lucau, C., and Defourny, P. (2007). Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. *Water Resources Research*, 43(4):W04421.

Pauwels, V. R. N. and Wood, E. F. (1999). A soil-vegetation-atmosphere transfer scheme for the modeling of water and energy balance processes in high latitudes 1. Model improvements. *Journal of Geophysical Research*, 104(D22):27811–27822.

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., and Wheater, H. S. (2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global Nest Journal*, 13(3):193–214.

Peck, R. and Devore, J. L. (1996). *Statistics - The exploration and analysis of data*. Brooks/Cole Publishing Company, London.

Peters, J., De Baets, B., Van doninck, J., Calvete, C., Lucientes, J., De Clercq, E., Ducheyne, E., and Verhoest, N. E. C. (2011). Absence reduction in entomological surveillance data to improve niche-based distribution models for Culicoides imicola. *Preventive Veterinary Medicine*, 100(1):15–28.

Peters, J., De Baets, B., Verhoest, N. E. C., Samson, R., Degroeve, S., De Becker, P., and Huybrechts, W. (2007). Random forests as a tool for ecohydrological distribution modelling. *Ecological Modelling*, 207(2-4):304–318.

Peters, J., Verhoest, N. E. C., Samson, R., Boeckx, P., and De Baets, B. (2008). Wetland vegetation distribution modelling for the identification of constraining environmental variables. *Landscape Ecology*, 23(9):1049–1065.

Peters, J., Verhoest, N. E. C., Samson, R., Van Meirvenne, M., Cockx, L., Vekerdy, Z., and De Baets, B. (2009). Uncertainty propagation in vegetation distribution models based on ensemble learning. *Ecological Modelling*, 220(6):791–804.

Peters-Lidard, C. D., Zion, M. S., and Wood, E. F. (1997). A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes. *Journal of Geophysical Research*, 102(D4):4303–4324.

Plate, E. J. and Duckstein, L. (1987). Reliability in hydraulic design. In Duckstein, L. and Plate, E. J., editors, *Engineering reliability and risk in water resources*, volume 124 of *NATO ASI Series E: Applied Science*, pages 27–60. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Ratto, M., Young, P. C., Romanowicz, R., Pappenberger, F., Saltelli, A., and Pagano, A. (2007). Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. *Hydrology and Earth System Sciences*, 11(4):1249–1266.

Rawls, W. J. and Brakensiek, D. L. (1985). Prediction of soil water properties for hydrologic modelling. In Jones, E. and Ward, T. J., editors, *Watershed Management in the Eighties*, Proceedings of Symposium, 30 Apr. - 1 May 1985, pages 293–299. ASCE, Denver, New York.

Rawls, W. J. and Brakensiek, D. L. (1989). Estimation of soil water retention and hydraulic properties. In Morel-Seytoux, H., editor, *Unsaturated flow in hydrologic modelling - Theory and practice*, pages 275–300. Kluwer Academic Publishing, Dordrecht.

Rawls, W. J., Brakensiek, D. L., and Saxton, K. E. (1982). Estimation of soil water properties. *Transactions of the Americal Society of Agricultural Engineers*, 25(5):1316–1320.

Reddy, R. K. and Haldar, A. (1992). A random-fuzzy reliability analysis of engineering systems. In Ayyub, B. M., Gupta, M. M., and Kanal, L. N., editors, *Analysis and management of uncertainty: Theory and applications*, pages 319–329. North-Holland, New York, USA.

Refsgaard, J., Henriksen, H., Harrar, W., Scholten, H., and Kassahun, A. (2005). Quality assurance in model based water management - review of existing practice and outline of new approaches. *Environmental Modelling & Software*, 20(10):1201–1215.

Refsgaard, J. C., van der Sluijs, J. P., Hojberg, A. L., and Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software*, 22(11):1543–1556.

Reichle, R. H. (2008). Data assimilation methods in the earth sciences. *Advances in Water Resources*, 31:1411–1418.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, 130(1):103–114.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 46:W05521.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics*, 1:318–333.

Rojas, R., Feyen, L., and Dassargues, A. (2008). Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging. *Water Resources Research*, 44(12):W12418.

Romanowicz, A. A., Vanclooster, M., Rounsevell, M., and La Junesse, I. (2005). Sensitivity of the SWAT model to the soil and land use data parametrisation: a case study in the Thyle catchment, Belgium. *Ecological Modelling*, 187(1):27–39.

Saich, P. and Borgeaud, M. (2000). Interpreting ERS SAR signatures of agricultural crops in Flevoland, 1993-1996. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):651–657.

Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517.

Saltelli, A., Chan, K., and Scott, M., editors (2008a). *Sensitivity Analysis*. John Wiley & Sons, New York.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisona, M., and Tarantola, S. (2008b). *Global Sensitivity Analysis: The Primer*. John Wiley, Chichester, UK.

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., Commission, E., and Ispra, J. R. C. (2006). Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering & System Safety*, 91(10-11):1109–1125.

Saltelli, A., Tarantola, A., Campolongo, F., and Ratto, M. (2004). *Sensitivity analysis in practice - A guide to assessing scientific models*. John Wiley & Sons, New York, USA.

Saltelli, A., Tarantola, S., and Campolongo, F. (2000). Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395.

Saltelli, A., Tarantola, S., and Chan, K. P. S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56.

Saxton, K. E. and Rawls, W. J. (2006). Soil water characteristics estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, 70(5):1569–1578.

Schaap, M. G., Leij, F. L., and van Genuchten, M. T. (1998). Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, 62:847–855.

Scheerlinck, K., Vernieuwe, H., and De Baets, B. (2012). Zadeh's extension principle for continuous functions of non-interactive variables: A parallel optimization approach. *IEEE Transactions on Fuzzy Systems*, 20(1):96–108.

Scheerlinck, K., Vernieuwe, H., Verhoest, N. E. C., and De Baets, B. (2013). Practical computing with interactive fuzzy variables. *Applied Soft Computing*. submitted.

Schmadel, N. M., Neilson, B. T., and Stevens, D. K. (2010). Approaches to estimate uncertainty in longitudinal channel water balances. *Journal of Hydrology*, 394(3-4):357–369.

Scholten, H., Kassahun, A., Refsgaard, J. C., Kargas, T., Gavardinas, C., and Beulens, A. J. M. (2007). A methodology to support multidisciplinary model-based water management. *Environmental Modelling & Software*, 22(5, SI):743–759.

Shafer, G. A. (1976). *Mathematical theory of evidence*. Princeton University Press, Princeton, New Jersey.

Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal*, 19:549–558.

Shaw, E. M. (1983). *Hydrology in Practice*. Chapman and Hall, London.

Sherman, L. K. (1932). Streamflow from rainfall by the unit graph method. *Engineering News Record*, 108:501–505.

Shi, J. C., Dozier, J., and Rott, H. (1994). Snow mapping in alpine regions with synthetic-aperture radar. *IEEE Transactions on Geoscience and Remote Sensing*, 32(1):152–158.

Silvan-Cardenas, J. L. and Wang, L. (2008). Sub-pixel confusion-uncertainty matrix for assessing soft classifications. *Remote Sensing of Environment*, 112(3):1081–1095.

Simard, M., Grandi, G. D., Saatchi, S., and Mayaux, P. (2002). Mapping tropical coastal vegetation using JERS-1 and ERS-1 radar data with a decision tree classier. *International Journal of Remote Sensing*, 23(7):1461 1474.

Singh, R. (1998). Land use cover changes, extreme events and ecohydrological responses in the Himalayan region. *Hydrological Processes*, 12(13-14):2043–2055.

Singh, V. and Woolhiser, D. (2002). Mathematical modeling of watershed hydrology. *Journal of Hydrologic Engineering*, 7(4):270–292.

Singh, V. P., editor (1995). *Computer models of watershed hydrology*. Water Resources Publications, Littleton, USA.

Singh, V. P. and Frevert, D., editors (2006). *Watershed models*. CRC Press, Boca Raton, USA.

Sivapalan, M., Beven, K., and Wood, E. F. (1987). On hydrologic similarity, 2. A scaled model for runoff prediction. *Water Resources Research*, 23(12):2266–2278.

Skriver, H. (2012). Crop classification by multi-temporal C- and L-band single and dual polarization, and fully polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6):2138–2149. 2012.

Skriver, H., Mattia, F., Satalino, G., Balenzano, A., Pauwels, V. R. N., Verhoest, N. E. C., and Davidson, M. (2011). Crop classification using short-revisit multitemporal SAR data. *Journal of Selected Topics in Earth Observations and Remote Sensing*, 4(2):423–431.

Skriver, H., Svendsen, M. T., and Thomsen, A. G. (1999). Multitemporal C- and L-band polarimetric signatures of crops. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2413–2429.

Sleutel, S., De Neve, S., Singier, B., and Hofman, G. (2006). Evaluation of pedo-transfer functions. *Soil Use and Management*, 22:188–196.

Smets, P. (1997). Imperfect information: imprecision-uncertainty. In Motro, A. and Smets, P., editors, *Uncertainty management in information systems. From needs to solutions*, pages 225–254. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Smets, P. (1999). *B.1.2. Theories of uncertainty*. Institute of Physics Publishing, Bristol, UK.

Soares, J. V., Rennó, C. D., Formaggio, A. R., da Costa Freitas Yanasse, C., and Frery, A. C. (1997). An investigation of the selection of texture features for crop discrimination using SAR imagery. *Remote Sensing of Environment*, 59(5):234–247.

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computation*, 1(4):407–414.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280.

Soet, M. and Stricker, J. N. M. (2003). Functional behaviour of pedotransfer functions in soil water flow simulation. *Hydrological Processes*, 17(8):1659–1670.

Sorooshian, S., Duan, Q. Y., and Gupta, V. K. (1993). Calibration of rainfall-runoff models - application of global optimization to the Sacramento Soil-Moisture Accounting model. *Water Resources Research*, 29(4):1185–1194.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15:72–101.

Stankiewicz, K. A. (2006). The efficiency of crop recognition on ENVISAT ASAR images in two growing seasons. *IEEE Transactions on Geoscience and Remote Sensing*, 44(4):806–814.

Strobl, C., Boulesteix, A., Zeilies, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(25):doi 10.1186/1471–2105–8–25.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, 9(307):doi 10.1186/1471–2105–9–307.

Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K. (2007). Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences*, 11(2):793–817.

Tekalign, T. and Hammes, P. S. (2005). Growth and productivity of potato as influenced by cultivar and reproductive growth II. Growth analysis, tuber yield and quality. *Scientia Horticulturae*, 105(1):29–44.

Tian, Y., Dickinson, R. E., Zhou, L., Zeng, X., Dai, Y., Myneni, R. B., Knyazikhin, Y., Zhang, X., Friedl, M., Yu, I. I., Wu, W., and Shaikh, M. (2004). Comparison of seasonal and spatial variations of leaf area index and fraction of absorbed photosynthetically active radiation from Moderate Resolution Imaging Spectroradiometer (MODIS) and Common Land Model. *Journal of Geophysical Research-Atmospheres*, 109(D1):D01103.

Tietje, O. and Hennings, V. (1996). Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. *Geoderma*, 69(1-2):71–84.

Tietje, O. and Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, 57(4):1088–1095.

Tolson, B. A. and Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1):W01413.

Troch, P. A., Smith, J. A., Wood, E. F., and De Troch, F. P. (1994). Hydrologic controls of large floods in a small basin - central Appalachian case-study. *Journal of Hydrology*, 156(1-4):285–309.

Tso, B. and Mather, P. M. (1999). Crop discrimination using multi-temporal SAR imagery. *International Journal of Remote Sensing*, 20(12):2443–2460.

Twarakavi, N. K. C., Sakai, M., and Simunek, J. (2009). An objective analysis of the dynamic nature of field capacity. *Water Resources Research*, 45:W10410.

Twarakavi, N. K. C., Simunek, J., and Schaap, M. G. (2010). Can texture-based classification optimally classify soils with respect to soil hydraulics? *Water Resources Research*, 46:W01501.

Ulaby, F. T., Moore, R. K., and Fung, A. K. (1986). *Microwave Remote Sensing. Active and Passive*, volume 3. Artech House.

Unwin, D. (1995). Geographical information systems and the problem of "error and uncertainty". *Progress in Human Geography*, 19(4):549–558.

Vachaud, G. and Chen, T. (2002). Sensitivity of computed values of water balance and nitrate leaching to within soil class variability of transport parameters. *Journal of Hydrology*, 264(1-4):87–100.

van Asselt, M. B. A. and Rotmans, J. (1996). Uncertainty in perspective. *Global Environmental Change-Human and Policy Dimensions*, 6(2):121–157.

van der Sluijs, J. P., Risbey, J. S., Kloprogge, P., Ravetz, J. R., Funtowicz, S. O. Quintana, S. C., Pereira, A. G., De Marchi, B., Petersen, A. C., Janssen, P. H. M., Hoppe, R., and Huijs, S. W. F. (2003). RIVM/MNP guidance for uncertainty assessment and communication: Detailed guidance. Technical report, University of Utrecht, Utrecht, Netherlands.

van Genuchten, M. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5).

Van Shaar, J. R., Haddeland, I., and Lettenmaier, D. P. (2002). Effects of land-cover changes on the hydrological response of interior Columbia River basin forested catchments. *Hydrological Processes*, 16(13):2499–2520.

Vazquez, R. F., Beven, K., and Feyen, J. (2009). GLUE based assessment on the overall predictions of a MIKE SHE application. *Water Resources Management*, 23(7):1325–1349.

Vereecken, H., Maes, J., and Feyen, J. (1990). Estimating unsaturated hydraulic conductivty from easily measured soil properties. *Soil Science*, 141(1):1–12.

Vereecken, H., Maes, J., Feyen, J., and Darius, P. (1989). Estimating the soil-moisture retention characteristics from texture, bulk-density, and carbon content. *Soil Science*, 148(6):389–403.

Verhoest, N. E. C., De Baets, B., Mattia, F., Satalino, G., Lucau, C., and Defourny, P. (2007). A possibilistic approach to soil moisture retrieval from ERS synthetic aperture radar backscattering under soil roughness uncertainty. *Water Resources Research*, 43(7):W07435.

Vernieuwe, H., Verhoest, N. E. C., Lievens, H., and De Baets, B. (2011). Possibilistic soil roughness identification for uncertainty reduction on SAR-retrieved soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2):628–638.

von Eynatten, H., Pawlowsky-Glahn, V., and Egozcue, J. J. (2002). Understanding perturbation on the simplex: A simple method to better visualize and interpret compositional data in ternary diagrams. *Mathematical Geology*, 34(3):249–257.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S. (2003). A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*, 39(8):doi 10.1029/2002WR001642.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, 44:W00B09.

Wagener, T. and Gupta, H. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, 19(6):378–387.

Wagner, B., Tarnawski, V. R., Hennings, V., Muller, U., Wessolek, G., and Plagge, R. (2001). Evaluation of pedo-tranfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*, 102(3-4):275–297.

Walker, W. E., Harremoës, P., Rotmans, J., Van der Sluis, J. P., Van Asselt, M. B. A., Janssen, P., and Krayer von Kraus, M. P. (2003). Defining uncertainty - a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, 4(1):5–17.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London, UK.

Wang, D., Lin, H., Chen, J. S., Zhang, Y. Z., and Zeng, Q. W. (2010). Application of multi-temporal ENVISAT ASAR data to agricultural area mapping in the Pearl River Delta. *International Journal of Remote Sensing*, 31(6):1555–1572.

Waske, B. and Braun, M. (2009). Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5):450–457.

Waske, B., Schiefer, S., and Braun, M. (2006). Random feature selection for decision tree classification of multi-temporal SAR data. In *2006 IEEE International Geoscience and Remote Sensing Symposium, vols 1-8*, IEEE International Symposium on Geoscience and Remote Sensing (IGARSS), pages 168–171.

Webb, T. H. and Lilburne, L. R. (2005). Consequences of soil map unit uncertainty on environmental risk assessment. *Australian Journal of Soil Research*, 43(2):119–126.

Weber, A., Fohrer, N., and Moller, D. (2001). Long-term land use changes in a mesoscale watershed due to socio-economic factors - effects on landscape structures and functions. *Ecological Modelling*, 140(1-2):125–140.

Wegehenkel, M. (2002). Estimating of the impact of land use changes using the conceptual hydrological model THESEUS - a case study. *Physics and Chemistry of the Earth*, 27(9-10):631–640.

Wegehenkel, M., Heinrich, U., Uhlemann, S., Dunger, V., and Matschullat, J. (2006). The impact of different spatial land cover data sets on the outputs of a hydrological model - a modelling exercise in the Ucker catchment, North-East Germany. *Physics and Chemistry of the Earth*, 31(17):1075–1088.

Wegmüller, U. and Werner, C. (1997). Retrieval of vegetation parameters with SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1):18–24.

Wei, H., Nearing, M. A., and Stone, J. J. (2007). A comprehensive sensitivity analysis framework for model evaluation and improvement using a case study of the Rangeland Hydrology and Erosion Model. *Transactions of the ASABE*, 50(3):945–953.

Wheater, H. S. Jakeman, A. J., Beven, K., Beck, M. B., and McAleer, M. J. (1993). Progress and directions in rainfall-runoff modelling. In Jakeman, A. J., Beck, M. B., and McAleer, M. J., editors, *Modelling change in environmental systems*, pages 101–132. Whiley, Chichester.

Wösten, J. H. M., Lilly, A., Nemes, A., and Le Bas, C. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, 90(3-4):169–185.

Wösten, J. H. M., Pachepsky, Y. A., and Rawls, W. J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3-4):123–150.

Wu, S.-J., Yang, J.-C., and Tung, Y.-K. (2011). Risk analysis for flood-control structure under consideration of uncertainties in design flood. *Natural Hazards*, 58(1):117–140.

Xu, B. and Gong, P. (2007). Land-use/land-cover classification with multispectral and hyperspectral E0-1 data. *Photogrammetric Engineering and Remote Sensing*, 73(8):955–965.

Ye, M., Neuman, S., and Meyer, P. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research*, 40(5):W05113.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

Zadeh, L. A. (1975a). Concept of a linguistic variable and its application to approximate reasoning .1. *Information Sciences*, 8(3):199–249.

Zadeh, L. A. (1975b). Concept of a linguistic variable and its application to approximate reasoning .2. *Information Sciences*, 8(4):301–357.

Zadeh, L. A. (1975c). Concept of a linguistic variable and its application to approximate reasoning .3. *Information Sciences*, 9(1):43–80.

Zadeh, L. A. (1999). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 100(S):9–34.

Zadeh, L. A. (2005). Toward a generalized theory of uncertainty (GTU) - an outline. *Information Sciences*, 172(1-2):1–40.

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, 14(6):415–421.

Zeng, X. Q. and Yeung, D. S. (2003). A quantified sensitivity measure for multilayer perceptron to input perturbation. *Neural Computation*, 15(1):183–212.

Zhang, J. and Goodchild, M. F. (2002). *Uncertainty in Geographical Information*. Taylor and Francis, London.

Zhang, Y. and Schilling, K. (2006). Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A field observation and analysis. *Journal of Hydrology*, 319(1-4):328–338.

Zhang, Y., Wu, L., Neggaz, N., Wang, S., and Wei, G. (2009). Remote-sensing image classification based on an improved probabilistic neural network. *Sensors*, 9(9):7516–7539.

Zhou, Y., McLauglin, D., and Entekhabi, D. (2006). Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Monthly Weather Review*, 134(8):2128–2142.

Zhu, J., Pohlmann, K. F., Chapman, J. B., Russell, C. E., Carroll, R. W. H., and Shafer, D. S. (2010). Sensitivity of solute advective travel time to porosities of hydrogeologic units. *Ground Water*, 48(3):442–447.

Zimmerman, H. J. (1996). *Fuzzy set theory and its applications*. Kluwer Academic Publisher, Dordrecht, the Netherlands.

Zimmermann, H. J. (1997). A fresh perspective on uncertainty modelling: Uncertainty vs. modelling. In Ayyub, B. M. and Gupta, M. M., editors, *Uncertainty analysis in engineering and sciences*, pages 353–364. Kluwer Academic Publishers, Massachussets, USA.

Zou, T., Yang, W., Dai, D., and Sun, H. (2010). Polarimetric SAR image classification using multifeatures combination and extremely randomized clustering forests. *Eurasip Journal on Advances in Signal Processing 2010*, page doi 10.1155/2010/465612.

Curriculum Vitae

Personalia

Naam	Lien Loosvelt
Geslacht	vrouwelijk
Nationaliteit	Belg
Geboortedatum	21/05/1984
Adres	Munkelostraat 30 8750 Wingene, België
Telefoon	+32/(0)472599120
E-mail adres	Lien.Loosvelt@UGent.be lienloosvelt@hotmail.com

Opleiding

2002-2007:	Bio-ingenieur Land- en Bosbeheer, Universiteit Gent *Scriptie*: *Schaal afhankelijke invloed van stenige bedekking op infiltratie, runoff en sediment verlies in ariede gebieden van Chili.* Promotoren: Prof. dr. ir. D. Gabriëls dr. ir. W. Cornelis
2009-2012:	Milieucoördinator Type A, IVPV Gent *Scriptie*: *Gebruik van het model PLEASE bij de identificatie van risicogebieden voor fosfor uitspoeling in Vlaanderen.* Promotor: Prof. dr. ir. S. De Neve

Loopbaanoverzicht-werkervaring

09/2007-11/2007:	Kwaliteitscontrole, Libeltex Meulebeke, Belgium
02/2008-02/2014:	Assistent faculteit Bio-ingenieurswetenschappen, Universiteit Gent Vakgroep: Vakgroep Bos- en Waterbeheer Labo: Laboratorium voor Hydrologie en Waterbeheer Promotor: Prof. dr. ir. Niko Verhoest

Conferenties

2010	European Geosciences Union General Assembly 2010, 2-7 mei, Wenen, Oostenrijk, mondelinge + poster presentatie
2012	EUROSOIL 4th International Congress of the European Confederation of Soil Science Societies (ECSSS), 2-6 juli, Bari, Italië, mondelinge presentatie
2012	American Geophysical Union's 45th Fall Meeting, 3-7 december, San Fransisco, Verenigde Staten, mondelinge + poster presentatie

Wetenschappelijke output

Gepubliceerd

L. Loosvelt, V.R.N. Pauwels, W.M. Cornelis, G.J.M. De Lannoy and N.E.C. Verhoest (2011). Impact of soil hydraulic parameter uncertainty on soil moisture modelling. *Water Resources Research*, 47: W03505

L. Loosvelt, J. Peters, H. Skriver, H. Lievens, F.M.B. Van Coillie, B. De Baets and N.E.C. Verhoest (2012). Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *International Journal of Applied Earth Observation and Geoinformation International Journal of Applied Earth Observation and Geoinformation*, 19: 173-184.

L. Loosvelt, J. Peters, H. Skriver, B. De Baets and N.E.C. Verhoest (2012). Impact of reducing polarimetric SAR input on the uncertainty of crop classifications based on the Random Forests algorithm. *IEEE Transactions on*

Geoscience and Remote Sensing, 50(10): 4185-4200.

L. Loosvelt, H. Vernieuwe, V.R.N. Pauwels, B. De Baets and N.E.C. Verhoest (2013). Local sensitivity analysis for compositional data with application to soil texture in hydrologic modelling. *Hydrology and Earth System Sciences*, 17: 461-478.

Ingediend

L. Loosvelt, V.R.N. Pauwels, and N.E.C. Verhoest (2013). On the significance of crop type information for the simulation of catchment hydrology. *Hydrological Processes*, ingediend op 28 augustus 2013.

L. Loosvelt, V.R.N. Pauwels, B. De Baets and N.E.C. Verhoest (2013). Assessing hydrologic prediction uncertainty resulting from soft land cover classification. *Journal of Hydrology*, ingediend op 26 september 2013.