Ghent University
Faculty of Sciences
Department of Physics and Astronomy

The application of the ALARO-0 model for regional climate modeling in Belgium:

extreme precipitation and unfavorable conditions for the dispersion of air pollutants under present and future climate conditions

Rozemien De Troch

Ghent University
Faculty of Sciences
Department of Physics and Astronomy

Supervisors

Prof. Dr. Piet Termonia	Royal Meteorological Institute of Belgium & Universiteit Gent
Dr. Rafiq Hamdi	Royal Meteorological Institute of Belgium

Members of the jury

Prof. Dr. Maarten Baes	Universiteit Gent
Dr. Radmila Brožková	Czech Hydrometeorological Institute
Prof. Dr. Herwig Dejonghe	Universiteit Gent
Prof. Dr. Dominique Fonteyn	Universiteit Gent
Prof. Dr. Erik Kjellström	Swedish Meteorological and Hydrological Institute
Prof. Dr. Jan Nyssen	Universiteit Gent
Prof. Dr. Jean-Pascal van Ypersele	Université catholique de Louvain
Prof. Dr. Patrick Willems	Katholieke Universiteit Leuven

Ghent University
Faculty of Sciences

Department of Physics and Astronomy
Krijgslaan 281 (S9), B-9000 Ghent, Belgium

Tel.: +32 9 264.47.98
Fax.: +32 9 264.49.89

Dissertation presented in partial fulfillment
of the requirements for the degree of Doctor of Science:
Physics
Academic year 2015-2016

Acknowledgements

I remember it was about five years ago, at the end of my postgraduate studies in Weather and Climate Modeling at the University of Ghent, that Piet told me he had in mind to submit a BELSPO project proposal for a PhD on climate research. In the first place I was a bit surprised about his suggestion. After my geographer studies, it had never been really my ambition to do a PhD. Even less doing research being surrounded by physicists, mathematicians, or engineers. Indeed, I was a bit scared that my background as a geographer would not be sufficient. If this would be the reason not to do a PhD, then that's precisely the reason to do it, Piet said. A few months later, in January 2012, I started my first day in the research department of the RMI, working on a PhD project that aimed to investigate the impact of climate change in Belgium based on regional climate scenarios.

This anecdote exactly demonstrates how Piet always managed to motivate me to explore new aspects of my research. Today, five years later, I am writing the acknowledgements of my PhD thesis. I can fairly say that I would not have been able to come at this stage without the help of my supervisor Piet. I would like to thank you for the many opportunities and confidence you gave me over the past years. The numerous scientific conferences and meetings that I could attend allowed me to share my work and to discuss with other people in my fields of interest. Your critical insights, and constructive feedbacks helped me to grow as a researcher. Even with your very busy agenda, you were always very enthusiastic to make time for the 'real science' when we discussed on new results or on how to proceed with my research.

Another person to who I am very thankful is my co-supervisor Rafiq. No matter what question I had, you were always ready and helpful to answer them with a lot of patient. Together we were the 'regional climate modeling group' at the department. I was often impressed by your dedication to urban climate, with the Urban Heat Island effect in Brussels having no secrets for you. A few months later Olivier joined our little climate team. The many discussions we had at work and at conferences always motivated me and gave me many new ideas on how to proceed with my research. It was always a pleasure to be overwhelmed with your enthusiasm and good scientific criticism for climate science. Besides that I will never forget the numerous "wedden voor 10 euro dat het nu lukt" that we had during the experimental set-up of the EURO-CORDEX runs, (y)our wheelchair experience during the Regional Climate Modeling Workshop in Lund and your subtle addiction to Kinder chocolates that you enthusiastically shared with your colleagues. One more year later, Julie made our climate group complete. Thanks to you I got to know more on soil moisture-atmosphere feedbacks, and in between our discussions we could share some fun 'geographer news'.

Furtermore, I thank my other colleagues from the ALADIN team: Alex for the R-help, Lesley for resolving all kinds of technical and computer related issues, Hans for the

Table of Contents

List of Figures

List of Tables

List of abbreviations

A

ACCEPTED	Assessment of Changing Conditions, Environmental Policies, Time-activities, Exposure and Disease
ACRANEB	Action de Recherche Petite Echelle Grande Echelle (ARPEGE) Calcul Radiatif Avec NEBulosité scheme for radiation
ALADIN	Aire Limitée Adaptation Dynamique développement INternational
ALARO-0	ALADIN and AROME (Application de la Recherche à l'Opérationnel à Meso-Echelle) combined model, first baseline version released in 1998
AOGCM	Atmosphere-Ocean coupled General Circulation Model
AR4	Fourth Assessment Report of the Intergovernmental Panel on Climate Change
AR5	Fifth Assessment Report of the Intergovernmental Panel on Climate Change
ARF	Areal Reduction Factor
ARPEGE	Action de Recherche Petite Echelle Grande Echelle
AWS	Automatic Weather Station

B

BBE	Big-Brother Experiment
BIME	Brussels Institute for Management of the Environment

C

CC	Clausius-Clapeyron
CH_4	Methane [μg m^{-3}]
CMIP3	Coupled Model Intercomparison Project phase 3
CMIP5	Coupled Model Intercomparison Project phase 5
CNRM	Centre National de Recherches Météorologiques

CO_2	Carbon dioxide
CORDEX	COordinated Regional climate Downscaling EXperiment
CORDEX.be	COmbining the Regional Downscaling EXpertise in BElgium: CORDEX and beyond
CPM	Convection-Permitting Model
CRM	Convection-Resolving Model
CRU	Climate Research Unit
CTM	Chemical Transport Model

E

ECMWF	European Centre for Medium-Range Weather Forecasts
ECUME	Exchange Coefficients from the Unified Multicampaigns Estimates
ENIAC	Electronic Numerical Integrator and Computer
ERA-40	40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis
EVA	Extreme Value Analysis

F

FMR	Fouquart-Morcrette Radiation scheme

G

GCM	General Circulation Model or Global Climate Model
GEV	Generalized Extreme Value distribution
GHG	GreenHouse Gas
GLAMEPS	Grand Limited Area Model Ensemble Prediction System
GPD	Generalized Pareto Distribution

H

H_2O	Water vapor
HadGEM3-RA	Hadley centre Global Environmental Model version 3
HIRLAM	HIgh Resolution Limited Area Model

I

IC   Initial Conditions
IDF   Intensity-Duration-Frequency relationship
IFS   Integrated Forecast System
IPCC   Intergovernmental Panel on Climate Change
ISBA   Interactions between Soil, Biosphere and Atmosphere

K

K-S test   Kolmogorov-Smirnov test

L

LAM   Limited Area Model
LBC   Lateral Boundary Condition
LT   Local Time
LSN   Large-Scale Nudging

M

MAE   Mean Absolute Error

N

NARCCAP   North American Regional Climate Change Assessment Program
NER   Net Exchange Rate
NO_x Nitrogen oxide [μg m^{-3}]
NO_2 Nitrogen dioxide [μg m^{-3}]
NWP   Numerical Weather Prediction

O

O_3 Ozone [μg m^{-3}]

P

PAV	Potential Added Value
PDF	Probability Density Function
PFRACT	Fractional contribution of extreme precipitation to total precipitation
PM	Particulate Matter [μg m^{-3}]
POT	Peak-Over-Threshold
PRUDENCE	Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects
PSS	Perkins Skill Score

Q

Q-Q	Quantile-Quantile

R

RC LACE	Regional Cooperation for Limited Area modeling for Central Europe
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RMI	Royal Meteorological Institute of Belgium
RMSE	Root Mean Square Error
RRTM	Rapid Radiative Transfer Model

S

SO$_2$	Sulfur dioxide [μg m^{-3}]
SRES	Special Report on Emission Scenarios
SURFEX	SURFace EXternalisée

T

TCR	Transient Climate Response
TEB	Town Energy Balance
TKE	Turbulent Kinetic Energy

U

UHI   Urban Heat Island

W

WD   number of Wet Days
WRF   Weather Research and Forecasting model
WTF   Wet-Time Fraction

1-9

3MT   Modular Multiscale Microphysics and Transport scheme

List of symbols

a	-	Slope of linear regression
b	-	Intercept of linear regression
c_p	J mol^{-1} K^{-1}	Specific heat capacity at constant pressure
d	Hour	Precipitation duration
D	-	Kolmogorov-Smirnov test statistic
e_s	Pa	Saturation pressure of water vapor
g	m s^{-2}	Gravity
G	W m^{-2}	Global solar radiation
G_T	-	Growth rate as function of the return period T
H_0	-	Null hypothesis
i_T	-	Return level estimate given by the Intensity-Duration-Frequency relation
l	m	Transport length
L_v	J kg^{-1}	Latent heat of water vapor
n	-	Sample size
N	okta	Cloudiness
p	Pa	Atmospheric pressure
p_s	Pa	Surface pressure
p_0	Pa	Standard reference pressure
q	g kg^{-1}	Specific humidity
q_{sat}	g kg^{-1}	Saturation specific humidity
R	J mol^{-1} K^{-1}	Universal gas constant of air
R_d	J kg^{-1} K^{-1}	Gas constant of dry air
R_v	J kg^{-1} K^{-1}	Gas constant of water vapor
R^2	-	Determination coefficient
T	Year	Return period
T_a	K	Atmospheric air temperature
T_v	K	Virtual temperature
u	m s^{-1}	Zonal wind speed
$\overline{u}$	m s^{-1}	Mean horizontal wind speed
v	m s^{-1}	Meridional wind speed
w	m s^{-1}	Wind speed at 10 m
x_m	-	Return level estimate of the Generalized Pareto Distribution, giving the amount of extreme precipitation corresponding to a given number of observations m

x_T	-	Return level estimate of the Generalized Extreme Value distribution, giving the amount of extreme precipitation corresponding to a return period T
Z	m	Geopotential height
α	-	Significance level
γ	-	Shape parameter of the Generalized Extreme Value distribution
Δt	s	Time step
Δx	km	Horizontal grid spacing
ε	-	Ratio of the gas constant for dry air to that of water vapor
η	-	Scaling factor in the power law formulation of the Generalized Extreme Value parameters
θ	K	Potential temperature
μ	-	Location parameter of the Generalized Extreme Value distribution
ν	s^{-1}	Brunt-Väisälä frequency
ξ	-	Shape parameter of the Generalized Pareto Distribution
σ	-	Scale parameter of the Generalized Pareto Distribution and the Generalized Extreme Value distribution
σ_c	-	Convective area fraction
ϕ_s	m	Geopotential at the surface

Nederlandse samenvatting
–Summary in Dutch–

Klimaatmodellen, en in het bijzonder regionale klimaatmodellen vinden hun oorsprong in numerieke weersvoorspellingsmodellen. Belangrijke ontwikkelingen in weersvoorspellingsmodellen leiden midden jaren 1950 tot de eerste toepassingen van globale klimaatmodellen. De huidige globale klimaatmodellen hebben een ruimtelijke resolutie van ongeveer $100-200$ km, en zijn een essentieel hulpmiddel om grootschalige mechanismen en fenomenen in ons klimaatsysteem, zoals de algemene atmosferisch circulatie, te bepalen en beter te begrijpen. De lage ruimtelijke resolutie van de globale modellen schiet echter tekort om vele belangrijke regionale en lokale aspecten in ons klimaatsysteem, en diens onderliggende kleinschalige fysische processen, zoals bijvoorbeeld extreme neerslag, in rekening te brengen. Om een oplossing te bieden voor dit schaalverschil tussen enerzijds de grootschalige informatie afkomstig van globale klimaatmodellen en anderzijds de kleinschalige en lokale informatie die uiterst relevant is voor impactstudies, werd de zogenaamde neerschalingstechniek ingevoerd.

Een veel gebruikte neerschalingstechniek, is de dynamische neerschaling, en in het bijzonder de nesting procedure. Hierbij worden grootschalige meteorologische velden van globale klimaatmodellen of van analyses van observaties gebruikt als initiële meteorologische randvoorwaarden voor een hoge-resolutie Limited Area Model (een model over een beperkt gebied zoals Europa of België) of regionaal klimaatmodel. De voorbije decennia kenden de regionale klimaatmodellen een opmerkelijke vooruitgang in hun ontwikkeling, gekenmerkt door een betere beschrijving van landschaps- en oppervlakte eigenschappen, en kleinschalige fysische effecten, evenals een verhoging in de ruimtelijke resolutie (tot minder dan 10 km). Vandaag de dag zijn regionale klimaatmodellen een veelgebruikt hulpmiddel voor regionale klimaatmodellering, waarbij meerjarige simulaties worden uitgevoerd om belangrijke regionale en lokale klimaatprocessen, zoals bijvoorbeeld extreme gebeurtenissen, te bestuderen.

Deze thesis beoogt in detail te onderzoeken in welke mate het operationele weermodel van het Koninklijk Meteorologisch Instituut (KMI) van België (i.e. ALARO-0) gebruikt kan worden voor regionale klimaatmodellering in België van (i) extreme neerslag en van (ii) meteorologische condities die ongunstig zijn voor de verspreiding van luchtvervuiling. Neerslag is één van de meest belangrijke klimaatvariabelen. Bovendien spelen de onderliggende neerslagprocessen een cruciale rol in de atmosferische toestand en het regionale en globale klimaat. Een correcte beschrijving van de neerslagprocessen in de klimaatmodellen is dus van uiterst belang. Tekortkomingen in de parametrisaties voor neerslag, en in het bijzonder voor diepe convectie, maken het echter voor klimaatmodellen doorgaans moeilijk om de ruimtelijke en temporele variaties, evenals de frequentie en intensiteit van neerslag correct te modelleren. De nieuwe fysische parametrisaties voor diepe convectie en wolken in het ALARO-0 model werden specifiek ontwikkeld in de context

van numerieke weersvoorspellingen voor de mesoschalen tot de schalen waarop convectie plaatsvindt (zogenaamde schalen in de "grijze" zone). Er werd inderdaad aangetoond dat deze fysische parametrisaties, of het "Modular Multiscale Microphysics and Transport (3MT)" schema, gekenmerkt worden door hun meerschaligheid, hetgeen consistente en realistische weersvoorspellingen impliceert voor ruimtelijke resoluties gaande van 10 km tot hoge resoluties van 4 km.

Sinds 2010 wordt het ALARO-0 model in het onderzoeksdepartement van het KMI gebruikt voor regionale klimaatmodellering. Een gedetailleerde regionale klimaat modelleringsstudie voor België werd sindsdien echter nog niet uitgevoerd. Het eerste doel van dit thesisonderzoek bestaat erin om in een klimaat context een gedetailleerde beschrijving en validatie uit te voeren van het model en van de nieuwe fysische parametrisaties dewelke het model kenmerken. Voor neerslag, en in het bijzonder extreme neerslag, wordt de validatie van de neerschalingsresultaten van het ALARO-0 model uitgevoerd voor een reeks van ruimtelijke en temporele resoluties, hetgeen ons tevens toelaat om als tweede doel te onderzoeken in welke mate het model waardevolle fijnschalige temporele en ruimtelijke details toevoegt aan het lage resolutie klimaatmodel dat de globale randvoorwaarden biedt. Algemeen omvatten modelresultaten van een neerschaling verschillende soorten onzekerheden dewelke gelinkt zijn aan (i) modelformulering, (ii) onzekerheden in de antropogene factoren voor klimaatforcering, en (iii) natuurlijke klimaatvariabiliteit. Een ensemble van meerdere modelsimulaties, dat toelaat om deze onzekerheden en spreiding op de modelresultaten te kwantificeren, vereist voldoende rekencapaciteit, hetgeen voor een kleine instelling zoals het KMI niet mogelijk is. Bijgevolg stellen we als laatste en derde doel in deze thesis om op een kwalitatieve manier de onzekerheden van de regionale neerschalingsresultaten in te schatten. Deze kwalitatieve inschatting gebeurt door onze modelresultaten onder toekomstige klimaatsomstandigheden in een ruimere context te plaatsen, en ze te vergelijken met andere regionale klimaatsprojecties, dewelke idealiter gebaseerd zijn op hetzelfde forceringsscenario.

Om deze doelstellingen te bereiken worden twee belangrijke onderzoeksstappen ondernomen. In eerste instantie wordt het Belgisch ALARO-0 weermodel gevalideerd voor klimatologische tijdschalen, waarbij zogenaamde "perfecte randvoorwaarden" afkomstig van globale reanalyses gebruikt worden aan de randen van het geneste ALARO-0 domein. Vervolgens wordt het ALARO-0 model gebruikt voor een klimaatprojectie van het A1B scenario beschreven door het "Intergovernmental Panel on Climate Change (IPCC)", waarbij een globale simulatie van klimaatverandering dynamisch neergeschaald wordt, gebruikmakend van het ALARO-0 model. De ALARO-0 simulaties in deze thesis worden uitgevoerd met een verticale resolutie van 46 modelniveaus, en een ruimtelijke resolutie tot 4 km, hetgeen overeenstemt met de kleinste atmosferische microschaal. Bovendien is deze horizontale en verticale resolutie veel hoger dan de resoluties van $100-200$ km en 12 km die gehanteerd worden in de "state-of-the-art" globale en regionale klimaatmodellen die gebruikt worden in internationale initiatieven zoals het "Model Intercomparison Project Phase 5 (CMIP5)" en het EURO-CORDEX project.

Een vergelijking van de verschillende opties in de parametrisaties voor straling- en oppervlakte toont aan dat de combinatie van het stralingsschema ACRANEB en het oppervlakteschema ISBA een aanvaardbare configuratie is voor de simulatie van neerslag met het ALARO-0 model.

Algemeen tonen de resultaten van de validatie van extreme neerslag op dagelijkse en subdagelijkse schaal aan dat het ALARO-0 model in staat is om op een consistente en correcte manier de relevante eigenschappen voor neerslag in achting te nemen, en dit voor at-

mosferische en overeenkomstige temporele schalen gaande van de micro- tot de mesoschaal. De validatie van dagelijkse neerslag in België toont aan dat de nieuwe 3MT parametrisatie, en in het bijzonder de meerschalige eigenschap van de parametrisatie, verantwoordelijk is voor een correcte simulatie van extreme zomerneerslag voor verschillende horizontale resoluties, gaande van 40 km tot 4 km resolutie. Vervolgens wordt onderzocht in welke mate het ALARO-0 model in staat is om verschillende kenmerkende aspecten van subdagelijkse neerslag voor verschillende temporele en ruimtelijke resoluties, te simuleren. Uit de resultaten van deze validatie blijkt de meerschaligheid van het ALARO-0 model voor de simulatie van dagelijkse zomerneerslag niet van toepassing te zijn voor de simulatie van subdagelijkse neerslag. Ten opzichte van de lage resolutie simulaties, duiden bovendien de modelresultaten van de simulatie op een hoge resolutie van 4 km op een significante toegevoegde waarde in de beschrijving van de dagelijkse neerslagcyclus, extreme neerslaghoeveelheden, en belangrijke schalingseigenschappen.

Deze positieve resultaten voor de validatie laten ons toe om in een volgende stap het model te gebruiken voor de berekening van een klimaatprojectie. De toekomstige veranderingen in extreme neerslag en meteorologische condities die ongunstig zijn voor wintersmog episodes, als gevolg van verhoogde broeikasgasconcentraties beschreven door het A1B scenario van het IPCC, worden in het tweede en laatste deel van deze thesis onderzocht. De validatie van de controle simulatie brengt significante biases aan het licht, die in hoofdzaak toegeschreven kunnen worden aan modelfouten afkomstig van het globale model CNRM-CM3 dat gebruikt wordt voor de neerschaling met het ALARO-0 model. De toekomstige veranderingen worden onderzocht volgens de gevoeligheid van het model voor veranderingen in de klimaatforcering, waarbij het verschil tussen de toekomstige scenario en controle simulatie gekwantificeerd wordt.

Wat betreft de veranderingen in extreme winterneerslag, kunnen we in de context van de model biases en de modelresultaten van andere projecties, met een zekere betrouwbaarheid een toekomstige verhoging in de uurlijkse neerslaghoeveelheden verwachten. Voor de zomer daarentegen zijn de negatieve veranderingen in extreme en uurlijkse neerslag meer onzeker. De afname is niet significant en kleiner dan de respectievelijke biases, en bovendien niet in overeenstemming met de gevonden resultaten in andere voorgaande studies voor West- en Centraal-Europa. Deze tegenstrijdigheden in de resultaten voor extreme zomerneerslag kunnen toegewezen worden aan de transitiezone waarin België zich bevindt, en de sterke afhankelijkheid van de parametrisaties, en in het bijzonder de parametrisaties voor diepe convectie, die een belangrijke bron van onzekerheid vormen in de projectie van extreme zomerneerslag.

Om het effect van klimaatverandering op wintersmog episodes in Brussel te bestuderen, worden twee verschillende stabiliteitsindices geanalyseerd. Beide indices, de transport index en de Pasquill klassen, zijn gebaseerd op meteorologische condities die de verspreiding van luchtvervuiling bepalen. Deze methodologie, waarbij de gevoeligheid van ongunstige omstandigheden voor de dispersie van wintersmog polluenten in een toekomstig klimaat onderzocht wordt met behulp van veranderingen in de frequentie van stabiliteitsindices, is nieuw en voor zover geweten niet eerder gebruikt in voorgaande studies. De biases aanwezig in de transport index voor de controle simulatie worden in rekening gebracht door rechtstreeks op de transport lengtes een bias correctie uit te voeren. Na de bias correctie tonen de resultaten een aanvaardbare afwijking in de frequentie van de gesimuleerde transport lengtes ten opzichte van de geobserveerde. Dezelfde bias correctie wordt toegepast op de transport lengtes van de scenario simulatie, hetgeen toelaat de onzekerheid op de toekomstige veranderingen in de frequentie van lage transport lengtes te kwantificeren. Zowel de

resultaten voor de veranderingen in de frequentie van lage transport lengtes evenals van de stabiele Pasquill klassen suggereren een tendens naar meer stabiele meteorologische condities en dus een mogelijke achteruitgang van de luchtkwaliteit tijdens wintersmog episodes in Brussel.

Uit bovenstaande resultaten concluderen we dat het ALARO-0 weermodel van het KMI gebruikt kan worden voor regionale klimaatmodellering in België, en in het bijzonder voor de toepassing van extreme neerslag en de meteorologische condities die ongunstig zijn voor de verspreiding van luchtvervuiling tijdens wintersmog episodes. Deze algemene conclusie laat ons tenslotte toe te stellen dat het grote potentieel van de hoge horizontale en verticale resolutie van de neergeschaalde modelresultaten relevante klimaatinformatie levert dat gebruikt kan worden voor verdere impactstudies van bijvoorbeeld het stedelijk hitte-eilandeffect, extreme neerslag, en de meteorologische condities die ongunstig zijn voor de verspreiding van luchtvervuiling.

English summary

Climate models, and in particular regional climate models (RCMs) find their origin in Numerical Weather Prediction (NWP) models. Important developments in NWP models have in the mid 1950's led to the first applications of Global Climate Models (GCMs). The current GCMs have a spatial resolution of around $100-200$ km, and are an essential tool to determine and correctly understand the large-scale mechanisms and phenomena in our climate system, such as the general atmospheric circulation. However, the coarse spatial resolution of the GCMs falls short to take into account many key regional and local aspects and underlying subgrid scale processes, such as for example extreme precipitation. To account for this scale difference between the large-scale information from the GCMs on the one hand, and the small-scale and local information which is extremely relevant for impact studies on the other, the downscaling technique has been introduced.

A frequently used downscaling technique, is the dynamical downscaling, and more specifically the nesting approach. In this approach large-scale meteorological fields from either a GCM or from analyses of observations are used to provide the initial and time-dependent meteorological Lateral Boundary Conditions (LBCs) for the high-resolution Limited Area Model (LAM) or RCM. Over the last decades, RCMs have undergone enormous improvements in their development, characterized by important advancements in their representation of landscape and surface features, in their description of subgrid-scale physical effects, and in their spatial resolution (down to 10 km). Nowadays, the RCMs have become a popular tool for regional climate modeling, in which multiyear simulations are carried out to study important regional and local climate processes, such as extreme events.

The ultimate aim of this thesis is to study in detail to what extent the operational NWP model of the Royal Meteorological Institute of Belgium (RMI) (i.e. ALARO-0) can be used for regional climate modeling in Belgium of (i) extreme precipitation, and of (ii) the unfavorable meteorological conditions for the dispersion of air pollution. Precipitation is one of the most important climate variables. Furthermore, the underlying precipitation processes play a crucial role in the state of the atmosphere and the regional and global climate. Hence, a correct description of the precipitation processes in the climate models is crucial. However, deficiencies in the parameterizations for precipitation, and in particular for deep convection, prevent the climate models to correctly simulate spatial and temporal variations, as well as the frequency and intensity of precipitation. The new physics parameterizations for deep convection and clouds in the ALARO-0 model were specifically designed in the context of NWP, and aimed to be used for the mesoscale to the convection-permitting scales (i.e. so-called "gray-zone" scales). The multiscale aspect of the physics parameterizations, called Modular Multiscale Microphysics and Transport (3MT), has indeed been demonstrated through consistent and realistic weather forecasts at spatial resolutions ranging from 10 km up to high resolutions of 4 km.

In the research department of the RMI, the ALARO-0 model is since 2010 used for regional climate simulations. Although, a detailed regional climate modeling study for Bel-

gium is since then not carried out. The first aim of this thesis is therefore to carry out a detailed description and validation of the model and its new physics parameterizations in a climate context. For precipitation, and more specifically extreme precipitation, the validation of the downscaling results is executed for a wide range of spatial and temporal resolutions. As a second goal, this allows us to explore to what extent the model adds valuable fine-scale temporal and spatial details to its driving coarse-resolution global model. The downscaled model results generally consist of multiple sources of uncertainties which are related to (i) model formulation, (ii) uncertainties in the anthropogenic climate forcing factors, and (iii) natural variability. A multi-model ensemble which allows to quantify these uncertainties and the spread of the model results, requires sufficient computing capacity, which is impossible for a small institute such as the RMI. Hence, the last and third goal of this thesis is to assess the uncertainties of the regional downscaling results in a qualitative manner. This qualitative assessment is done through comparison of our downscaled future climate model results w.r.t. other RCM climate projections, which ideally use the same scenario of natural and anthropogenic forcing.

The above objectives are addressed by two main research steps. In a first step the Belgian ALARO-0 NWP model is validated for climatological time scales, by driving the model with "perfect boundary conditions" coming from global reanalyses. In a following step, the model is applied for a climate projection under the A1B scenario as described by the Intergovernmental Panel on Climate Change (IPCC), in which a global climate change simulation is dynamically downscaled using the ALARO-0 model. In this thesis, the ALARO-0 simulations with a vertical resolution of 46 model levels, are carried out up to high spatial resolutions of 4 km, corresponding to the finest atmospheric (micro)scales. Furthermore, this horizontal and vertical resolution is much higher than the state-of-the-art GCM and RCM resolutions of roughly $100-200$ km and 12 km as used in international initiatives such as the Model Intercomparison Project Phase 5 (CMIP5) and the EURO-CORDEX project.

A comparison of the different options in the radiation and surface parameterizations demonstrates that the combination of the radiation scheme ACRANEB and the surface scheme ISBA is an acceptable configuration for the simulation of precipitation with the ALARO-0 model.

Overall, the results of the validation of extreme daily and subdaily precipitation show that the ALARO-0 model is able to consistently and correctly reproduce the relevant precipitation characteristics, and this for a wide range of atmospheric and corresponding temporal scales varying from the micro- to the mesoscale. The validation of daily precipitation in Belgium demonstrates that the new 3MT parameterization, and its multiscale character, are responsible for a correct simulation of extreme summer precipitation at multiple horizontal resolutions, ranging from 40 km to 4 km resolution. Subsequently it is investigated to what extent the ALARO-0 model is able to simulate several subdaily precipitation characteristics at different temporal as well as spatial resolutions. The results of this validation suggest that the multiscale character of the ALARO-0 model as apparent in the simulation of the daily precipitation climatology, is not valid for the simulation of subdaily precipitation. Compared to the low-resolution simulations, the high 4-km model results demonstrate a significant added value in the description of the daily precipitation cycle, very high precipitation amounts, and important scaling properties.

These positive results from the validation allow us to apply the model in a next step for the calculation of a climate projection. The future changes in extreme precipitation and the meteorological conditions which are unfavorable for winter smog episodes, as a consequence of increased greenhouse gas (GHG) concentrations described by the A1B scenario

of the IPCC, are investigated in the second and last part of this thesis. The validation of the control simulation reveal significant biases, which can be attributed to model errors that are present in the driving GCM CNRM-CM3. The future changes are explored through a sensitivity of the model for changes in the climate forcing, in which the differences between the future scenario and the control simulation are quantified.

When it comes to the extreme winter precipitation, we can expect, taking into account the model biases and the projection results from other modeling studies, to some level of confidence a future increase in the hourly precipitation amounts. However, for summer precipitation the negative changes in extreme and hourly precipitation are more uncertain. The negative changes are not significant and smaller than their respectively biases, and the negative response is in disagreement with the modeling results for western- and central Europe from previous studies. These disagreements can be attributed to the transition zone in which Belgium is located, and the strong dependency of the parameterizations, and in particular the parameterizations for deep convection, which are an important source of uncertainty in the projection of extreme summer precipitation.

To study the climate change impact on winter smog episodes in Brussels, two different stability indices are analyzed. Both indices, the transport index and the Pasquill stability classes, are based on the meteorological conditions determining the dispersion of air pollution. This methodology, in which the sensitivity of unfavorable conditions to the dispersion of winter smog pollutants under future climate conditions is explored through changes in the frequency of stability indices, is new and has to our knowledge never been used in previous studies. The biases which are present in the transport index obtained from the control simulation are taken into account by a bias correction directly applied on the transport lengths themselves. After the bias correction, the results show an acceptable deviation in the frequency of modeled transport lengths with respect to the observed ones. The same bias correction is applied on the transport lengths of the scenario simulation. This allows to quantify the uncertainty in the future changes of the frequencies of low transport length values. Both the results of the changes in the frequency of low transport length values as well as in the frequency of stable Pasquill classes suggest a tendency towards more stable meteorological conditions, and hence a possible deterioration of the air quality during winter smog events in Brussels.

Based upon these results we conclude that the ALARO-0 model of the RMI can be used for regional climate modeling in Belgium, and in particular for the application of extreme precipitation and the meteorological conditions which are unfavorable for the dispersion of air pollution during winter smog episodes. This general conclusion allows us to state that the great potential of the high horizontal and vertical resolution of the downscaled model results provide relevant climate information that can be used as a forcing for impact studies on for example the Urban Heat Island effect (UHI), extreme precipitation, and the meteorological conditions which are unfavorable for the dispersion of air pollution.

1
GENERAL INTRODUCTION

The ultimate aim of this thesis is to investigate to what extent the Belgian ALARO-0 Numerical Weather Prediction (NWP) model can be applied for regional climate modeling of (i) (extreme) precipitation and of (ii) the unfavorable meteorological conditions for the dispersion of air pollution in Belgium. Regional Climate Models (RCMs) in general, and the ALARO-0 model in particular, find their origins in NWP. This research is motivated by the fact that a detailed study of the application of the ALARO-0 NWP model for regional climate modeling, has since its use for regional climate simulations in 2010, never been carried out.

As a general introduction to this research theme, the origins of regional climate modeling in NWP, the principle of the downscaling approach, and some relevant conceptual issues to the downscaling approach are discussed below. This chapter ends with formulating the key research goals that will be addressed in this thesis.

1.1   Origins of regional climate modeling

1.1.1   Legacy of Numerical Weather Prediction (NWP)

Climate models, and more specifically regional climate models, find their origin in NWP. The beginning of modern NWP started with John von Neumann, who was one of the leading mathematicians of the 20th century. During the 1950's, von Neumann played a key role in the development of the computer industry. His interests in thermodynamics made him realize that a means of solving numerically the complex equations behind it, would greatly accelerate the understanding in these topics. With a team of engineers, programmers, mathematicians and meteorologists, von Neumann started the "Electronic Computer Project". He indeed recognized weather forecasting as an ideal problem for an automatic computer. At that time the meteorological group was led by Jule Charney. Charney's pioneering analyses on the set of nonlinear primitive equations which describe the global atmospheric flow, resulted in a new set of equations where the gravity wave solutions were completely eliminated, i.e. the quasi-geostrophic system. By early 1950, the meteorology group had completed the necessary mathematical analysis, and under supervision of von Neumann the first successful NWP using an equivalent-barotropic quasi-geostrophic model was run on the only computer then available, the Electronic Numerical Integrator and Computer (ENIAC). The initial results from the four 24-hour forecasts that were made, were very encouraging. They clearly indicated that the large-scale features of the mid-tropospheric flow could be forecasted barotropically with a reasonable resemblance to reality (Lynch, 2008; Arakawa and Jung, 2011).

Around the same time, in 1955, Jacques van Isacker worked together with colleagues at the Royal Meteorological Institute of Belgium (RMI) on a barotropic model, where he delivered pioneering work for NWP in Belgium. The first operational forecasts for Western Europe with their barotropic model started in 1962, which took two hours to run a 48-h forecast (Persson, 2005).

These important advancements in NWP paved the way for the development of the early General Circulation Models or Global Climate Models (GCMs). The first long-range simulation of the general circulation of the atmosphere was performed by Phillips by the mid

1950's. He used a two-level quasi-geostrophic model with rudimentary physics. The integrations were done on a spatial grid of 16×17 points and for a period of about one month. Later, several major research groups developed more advanced GCMs as a fundamental tool for modeling the general circulation of the atmosphere and the understanding of the climate (Lynch, 2008).

1.1.2   The downscaling approach

The state-of-the art GCMs, with a typical horizontal resolution of roughly 1 to 2 degrees ($\approx$ 100-200 km), effectively allow for the simulation of large-scale climate features such as the general circulation of the atmosphere and the ocean, and sub-continental patterns of, for example, temperature and precipitation (Rummukainen, 2010; Flato et al., 2013). However, given the broad range of spatial and temporal scales at which atmospheric phenomena in the global climate system occur (Fig. 1.1), these coarse spatial resolutions fall short of many key regional and local aspects and underlying physical subgrid scale processes, such as extreme precipitation, which are of particular interest for impact researchers, stakeholders, and policy makers (Rummukainen, 2010; Prein et al., 2015). To resolve this scale discrepancy between global climate model output and the resolutions required for impact assessment, the downscaling technique has been introduced. It attempts to obtain regional or local detail from sparse observations or low resolution numerical simulations (Maraun et al., 2010; Rummukainen, 2010). The two main downscaling methods are known as statistical and dynamical downscaling. Statistical downscaling establishes robust statistical relationships between large(r) scale climate variables (e.g. the mean sea level pressure field) and observed local ones (such as temperature or precipitation) (Maraun et al., 2010; Rummukainen, 2010). There exist a great number of statistical downscaling methods (see e.g. Maraun et al. (2010) for a review on different methods).

Dynamical downscaling is based on the idea to produce high resolution over a region of interest by using a global or regional numerical model. Three approaches to dynamical downscaling can be distinguished. The first one consists of running global integrations at very high spatial resolutions. However, the large computational costs of this technique currently limits the runtime of such simulations to a few days or one month. The other technique uses a global model with a variable resolution. In this case, the regional grid spacing is reduced over the region of interest, while larger grid spacing is used elsewhere for computational efficiency (Rummukainen, 2010; Prein et al., 2015).

The third and most widely used dynamical downscaling approach, is the "nesting" technique. It originates from NWP and is based on the concept of one-way nesting, where large-scale meteorological fields from either a GCM or from analyses of observations are used to provide the initial and time-dependent meteorological Lateral Boundary Conditions (LBC) for the high-resolution so-called Limited Area Model (LAM) or RCM simulations (Fig. 1.2). The basic strategy underlying this one-way nesting approach is that the GCM can provide the response of the global circulation to large scale forcings, and that the LAM or RCM can account for the local or sub-GCM grid scale forcings (e.g. complex topographical features) in a physically-based way (Giorgi and Mearns, 1999; Denis et al., 2002). The nesting approach or limited-area modeling was first extensively used for NWP, where LAMs were mostly run for simulation times of a few days. Later, it has been demonstrated that

L_H	Lifetime	Stull (1988)	Pielke (1984)	Orlanski (1975)	Present	Atmospheric phenomena
	1 month		Synoptic Regional	Macro-α	Macro-α	General circulation, long waves
10 000 km		Macro				
				Macro-β	Macro-β	Synoptic cyclones
2000 km	1 week					
				Meso-α	Macro-γ	Fronts, hurricanes
200 km	1 day		Meso	Meso-β	Meso-β	Low-level jets, thunderstorm groups, mountain winds and waves, sea breeze, urban circulations
20 km		Meso				
	1 h			Meso-γ	Meso-γ	Thunderstorm, clear-air turbulence
2 km						
				Micro-α	Meso-δ	Cumulus, tornadoes, katabatic jumps
200 m	30 min	Micro				
			Micro	Micro-β	Micro-β	Plumes, wakes, waterspouts, dust devils
20 m	1 min					
					Micro-γ	
2 m	1 s	Micro δ		Micro-γ		Turbulence, sound waves
					Micro-δ	

Figure 1.1: Atmosperic scale definitions, where L_H corresponds to the horizontal length scale (adopted from Thunis and Bornstein, 1996).

LAMs could be run continuously for longer climatological time periods, such that multiyear nested model simulations were conducted for *regional climate modeling* in order to address in detail climate processes and add detail to large-scale data (Giorgi and Mearns, 1999; Rummukainen, 2010).

First applications of the nested LAMs in NWP, aimed for modeling the mesoscale processes, at grid resolutions higher than the synoptic and global models. The mesoscale, corresponding to the fundamental horizontal scale often referred as the first baroclinic Rossby radius of deformation, is thus a very important scale in NWP, at which many relevant atmospheric processes take place (Fig. 1.1). Nowadays, also atmospheric dynamical RCMs are used in a wide range of applications, varying from the reconstruction of regional-scale paleoclimatology to dynamical downscaling of climate change projections (Richard et al., 2002). Over the last decades, LAMs and RCMs have undergone enormous improvements in their development, characterized by important advancements in their spatial resolution (down to 10 km), in their representation of landscape and surface features,
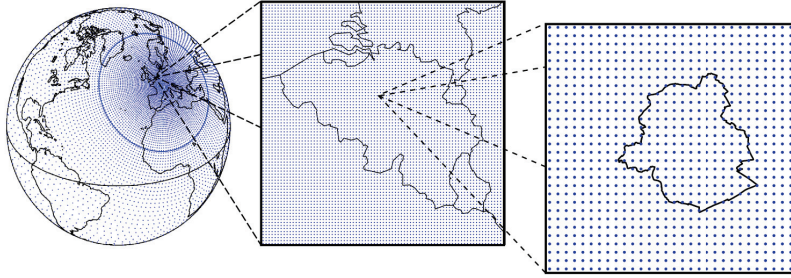
Figure 1.2: Nested modeling technique: Large-scale meteorological fields from Global Climate Models (GCMs) serve as initial Lateral Boundary Condition (LBC) for the high-resolution Limited Area Model (LAM) or Regional Climate Model (RCM) simulations.

and in their description of subgrid-scale physical effects (Rummukainen, 2010). Because of the ability of these high-resolution LAMs and RCMs to reproduce meaningful small-scale features over a limited region (Denis et al., 2002; Giorgi et al., 2004), they have become a popular tool in both the NWP and the climate community in particular for studying extreme events at the regional and local scales (e.g. Jones and Reid, 2001; Buonomo et al., 2007; Déqué and Somot, 2008; Dulière et al., 2011).

Nevertheless, some studies show that RCMs do not necessarily improve their driving GCM simulations or global reanalyses (e.g. Castro et al., 2005; Jacob et al., 2007; Sylla et al., 2010). The use of nested LAMs or RCMs as a climate downscaling technique indeed involves a number of technical as well as conceptual issues (Richard et al., 2002). The major technical issues related to the nesting technique are: (i) problems related to the LBCs driving the RCM, (ii) the divergence of the RCM fields from the large-scale atmospheric circulation provided by the driving global model, and (iii) the choice of the domain size and the spatial resolution of the RCM. Besides these technical issues (see the next Chapter 2 for a further discussion), also some conceptual issues or important scientific questions related to the nesting approach have been raised within the RCM community.

1.2 Conceptual issues

1.2.1 Parameterizations

Precipitation is one of the most important climate variables. Precipitation processes, including amongst others cloud microphysics, cumulus convection, planetary boundary layer processes, large-scale circulation, are key physical processes that strongly influence the atmosphere and the regional and global climate. Furthermore, future climate change is expected to have a great impact on society due to changes in precipitation patterns and variability. A correct representation of the precipitation processes in the models is thus of

a crucial importance. Nevertheless, it is still a big challenge for GCMs or RCMs to real-istically simulate the spatial and temporal variations, and both in frequency and intensity of precipitation. These errors in simulated precipitation fields are often associated with deficiencies in the *parameterizations* (Dai, 2006).

The finite model resolution of weather and climate models requires that all physical processes that occur on finer scales than those that are resolved by the model, need to be approximated. These approximations, known as parameterizations, are thus incorporated in the model in an attempt to relate the statistical effects of unresolved small-scale processes to the large scale processes that are explicitly resolved in the model (Arakawa, 2004; Rum-mukainen, 2010).

One critical subgrid process that has almost always been at the core of the efforts to nu-merically model the atmosphere, is the representation of deep convection through convection parameterization schemes. Deep convection plays a crucial role in the precipitation process in many regions of the world, and contributes to extreme events such as flash floods and landslides through heavy precipitation associated with mesoscale convective systems, squall lines, and tropical cyclones. The triggering of deep convection emerges from an interplay of processes acting at scales from the microscale to the synoptic scale (Fig. 1.1), making the parameterization of deep convection challenging. Furthermore, convection parameterization schemes interact with many other parameterization schemes (e.g. microphysics, radiation, and planetary boundary layer schemes), such that deficiencies in convection parameteriza-tion schemes indeed result in common errors in the precipitation field (Prein et al., 2015).

Since the beginning of the 21st century, advancements in high-performance computing, allowed further refinement of the spatial grids of RCMs well beyond 10 km resolution. At such small spatial scales (i.e. < 4 km), deep convection starts to be resolved explicitly, so that the error-prone convection parameterization schemes can be switched off in these so-called convection-permitting (CPMs) or convection-resolving models (CRMs) (Prein et al., 2015). As shown in Fig. 1.3, at the two ends of the climate models' horizontal resolution range (abscissa), GCMs and CRMs generally use a different kind of model physics in the representation of deep convection (ordinate, is a measure for the degree of parameterization, such as the reduction in the degree of freedom, increasing downwards) (Arakawa and Jung, 2011; Arakawa et al., 2011).

Both families of models have been developed from a different viewpoint w.r.t. the range of horizontal resolution on which they can be applied. Hence, in GCMs deep convection is usually highly parameterized (unresolved), while in CRMs these processes are explicitly simulated (resolved). Obviously, there have been many studies exploring the applicability of GCMs (CRMs) to higher (lower) resolutions as shown by the horizontal arrows in Fig. 1.3 (Arakawa and Jung, 2011; Arakawa et al., 2011).

The artificial separation in the spectrum of processes at different scales in a resolved and unresolved part by which the model physics is treated, can only be justified when the resulting error can be made arbitrarily small by using a higher resolution. This requires that, with an increase of the horizontal resolution, both the dynamics and physics of GCMs converge to those of the CRMs (as shown by the dashed curve in Fig. 1.3). If the GCM and CRM share the same dynamical core, it is expected that the convergence does take place as far as the model dynamics is concerned. However, for the conventional formulations of
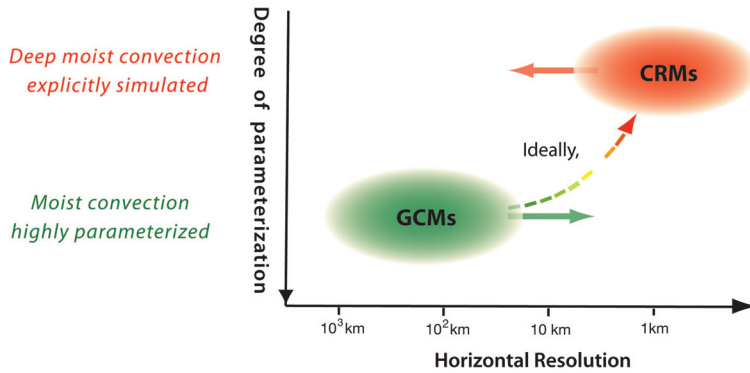
Figure 1.3: *Two families of atmospheric models with different model physics. The horizontal arrows represent attempts to broaden the applicability of GCMs (CRMs) to higher (lower) resolutions (adopted from Arakawa et al., 2011).*

model physics, and especially when deep convection is involved, the GCM model solution does not naturally converge to the solution of the CRM (or of the real atmosphere) as the resolution is refined (Arakawa and Jung, 2011; Arakawa et al., 2011). This issue, which has been referred by Arakawa (2004) as the so-called cumulus or deep convection parameterization problem, is schematically depicted in Fig. 1.4. The upper half of the loop in Fig. 1.4, represents the effects of the resolved processes on the unresolved component of the moist convection, while the lower half represents the opposite effects. Arakawa (2004) refers to the upper half of the loop as the "control" or "large-scale forcing" and the lower half as the "feedback" or "cumulus adjustment". The cumulus parameterization which aims for a two-way interaction between the resolved and unresolved part and thus a closed loop, requires closure assumptions. The choice of such closure assumptions are by no means an obvious question (Arakawa, 2004; Arakawa and Jung, 2011).

The artificial distinction in the model physics, representing deep convection in GCMs and CRMs, has led to several efforts to unify all physical parameterizations in the models (e.g. Arakawa, 2004; Arakawa et al., 2011; Arakawa and Jung, 2011). Such a unification would allow a continuous transition of model physics from one kind to the other as the resolution changes and improves interactions between different physical processes (Arakawa et al., 2011). The approach used in the model physics of the ALARO-0 model, which is currently used for the operational NWP at the RMI, is a key example of such a consistent treatment of deep convection at various resolutions from fully subgrid to fully explicit. Gerard et al. (2009) proposed and developed a new physics parameterization package for deep convection and clouds, to be used specifically from the mesoscale to the convection-permitting scales (so-called "gray-zone" scales). This Modular Multiscale Microphysics and Transport (3MT) physics package is based on an integrated sequential treatment of resolved condensation, deep convection and microphysics, using prognostic variables. In the down-scaling chain toward CRM simulations, many modelers avoided simulations in this "gray zone", corresponding to grid spacings between 10 km (convection-parameterized) and 4 km (convection-permitting), as some assumptions used in parameterizations of deep convection
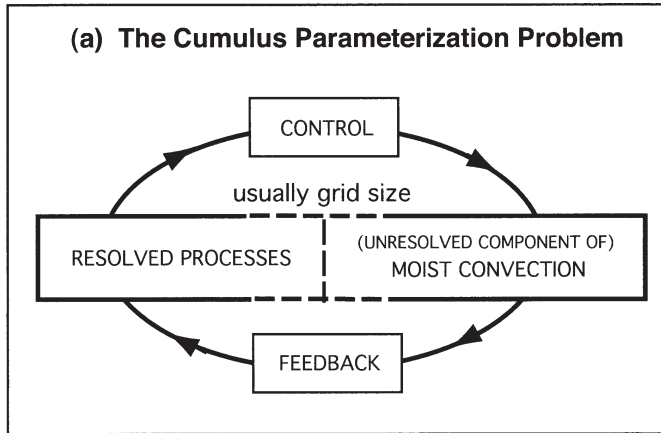
Figure 1.4: *A schematic diagram showing interactions between the resolved and unresolved processes of moist convection. The formulation of the right half of the loop represents the cumulus parameterization problem (adopted from Arakawa, 2004).*

are violated and deep convection is insufficiently resolved to be modeled explicitly (Prein et al., 2015). However, this 3MT physics parameterization package has found to be suitable scale-aware to allow simulations at the gray-zone scales. Model results of an operational forecast using this revised and modular structure of the physical parameterizations, indeed demonstrated consistent and realistic results at resolutions ranging from a few tens of kilometers down to less than 4 km (Gerard et al., 2009). The next Chapter 2 provides a further detailed description of ALARO-0 and 3MT.

1.2.2   Added value

As already alluded, the main potential of dynamical downscaling using RCMs can be found in their spatial resolution which is suggested to *add value* compared to the lower resolution data from global reanalysis or GCMs. In this context, added value means that higher resolution models simulate the observed local-to-regional climate processes more realistically and therefore provide more appropriate information compared to other lower resolution models (Mayer et al., 2015).

The spatial resolution of RCM simulations has indeed steadily increased over the last decades. As an illustration, the ensemble assessments over Europe carried out in the PRU-DENCE (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects) and ENSEMBLES projects applied main resolutions of around 50 km and 25 km, respectively. The ongoing EURO-CORDEX initiative, which is part of the global COordinated Regional climate Downscaling EXperiment framework (CORDEX), which aims to produce an ensemble of several dynamical and statistical downscaling RCMs for Europe, uses spatial resolutions of $0.44°(\approx 50$ km) up to $0.11°(\approx 12$ km). It thereby complements the coarser resolution datasets of the former PRUDENCE and ENSEMBLES activities (Maraun et al., 2010; Berg et al., 2013b; Jacob et al., 2014). Furthermore, re-

cently with the advances in high-performance computing, CPMs operating at the kilometer scale (i.e. resolutions < 4 km), are becoming available on climate time scales (e.g. Kendon et al., 2012; Prein et al., 2013b; Ban et al., 2014; Chan et al., 2014; Fosser et al., 2015). For decades, these CPMs, where most of the error-prone convection parameterizations are switched off, have been widely used for NWP.

The kilometer scale at which these CPMs operate, allows to explicitly resolve deep convection as well as an improved representation of fine-scale orography, variations of surface fields and boundary layer processes, which are crucial for the initiation of convection in complex terrain (Ban et al., 2014). Furthermore, precipitation is one of the climate variables most sensitive to model formulation, being strongly dependent on several parameterization schemes and their interplay with the resolved model dynamics. Hence, it is not surprising that it is shown that for this variable the high-resolution models significantly add value compared to the global ones (Maraun et al., 2010; Rummukainen, 2010). More specifically, CPMs are found to improve the representation of the precipitation distribution including the extreme events, and this over regions of complex topography and at small spatial and temporal scales (e.g. subdaily precipitation) (Maraun et al., 2010; Prein et al., 2013a, 2015).

As the downscaling principle assures that RCMs should not alter the simulated climate on scales that can be skillfully reproduced by the resolutions of the global model, the added value of the downscaling thus appears on the finer scales, such as for example mesoscale structures and extremes (Rummukainen, 2010). However, if the latter is true as a principle, it does not imply that only small-scale or mesoscale processes may be changed by the RCM. More specifically, if a RCM improves mesoscale structures such as fronts, it can indeed lead to differences at the larger scales (e.g. better representation of cyclone development) which, in turn, have an impact also on the larger scales. In reanalyses-driven simulations the large scales are well-constrained in the global model by the observations. But, in GCM-driven simulations RCMs may have the potential to improve the large scales.

Generally, a comprehensive and exact assessment of the added value of an RCM can thus be done by driving the model with reanalysis data. The latter indeed provides quasi-observed boundary conditions, which allows to isolate the RCM downscaling skill by excluding the systematic biases usually present when coupling the nested model to a GCM. Furthermore, these reanalyses are able to reproduce the actual day-to-day sequence of weather events, so that a more clear assessment of the downscaling skill is indeed possible (Maraun et al., 2010).

1.2.3   Uncertainty

The main sources of *uncertainty* of the downscaled RCM model results are: (i) model formulation, (ii) uncertainty in anthropogenic climate forcing factors, and (iii) natural variability. Uncertainty due to model formulation includes the numerical schemes, parameterizations and resolution, coupling strategy, and uncertainty due to natural variability can be attributed to internal variability of the chaotic climate system dependent on initial conditions, and natural forced variability due to, e.g., solar forcing (Maraun et al., 2010).

As mentioned previously, parameterizations represent a simplification of the real world and thus lead to inherent uncertainty in the regional climate modeling. More specifically, the deep convection parameterization schemes are a key source of uncertainty in the simu-

lation of precipitation, as these schemes are often inherited from low-resolution GCMs or originally developed for operational short-range NWP models (Richard et al., 2002; Maraun et al., 2010).

The time scale under consideration strongly determines the relative roles of the different sources of uncertainty. On longer time scales, the signal to noise ratio between the climate change signal and natural variability increases, and uncertainty due to model formulation becomes dominant. In this way, variations in RCM formulation play a dominant role at fine-scales, and particularly for changes in precipitation extremes in summer. Nevertheless, in the context of future changes, natural variability is still found to be important for summertime precipitation and precipitation extremes, such that a single 30 year climate projection is not robust. A climate projection represents indeed just one possible realization of the future climate, conditional on a given scenario of natural and anthropogenic forcing (Maraun et al., 2010).

The range of uncertainty due to model formulation and natural variability can be assessed by ensemble simulations based on different GCMs and RCMs (multi-model ensembles), perturbed parameterizations (perturbed physics ensembles) and different initial conditions. Over the last decade, there has been considerable international effort to quantify uncertainty in regional climate modeling through the inter-comparison of multiple RCMs. Examples of such international initiatives are the European projects PRUDENCE, ENSEMBLES, and the most recent project CORDEX, which study the uncertainty due to structural errors of different GCMs and/or RCMs (Maraun et al., 2010).

1.3   Key research goals

Related to these three conceptual issues (parameterizations, added value, and uncertainty), which are often raised within the RCM community, we formulate the three key research goals of this thesis:

1. *Describe in detail the Belgian operational ALARO-0 NWP model with its revised physical parameterizations in the context of regional climate modeling.* A detailed study and validation of the application of the ALARO-0 NWP model for regional climate purposes in Belgium, has since its use for regional climate simulations in 2010, never been carried out, making it a very important novelty of this thesis.

2. *Investigate at which temporal and spatial scales the downscaling results add value.* As mentioned previously, the downscaling principle aims that RCMs should not alter the simulated climate on scales that can be skillfully reproduced by the resolutions of the global model, but should rather add value on the finer scales, such as for example mesoscale structures and extremes.

3. *Assess qualitatively the uncertainty of the regional downscaling climate change results.* Multi-model ensemble simulations, allowing to quantify the spread of uncertainty due to model formulation and natural variability, is not feasible at a small institute such as the RMI with limited computing resources. Therefore, the uncertainty is explored in a qualitative manner by comparing our downscaled future climate model

results in its proper context, i.e. w.r.t. other RCM climate projections, which ideally use the same scenario of natural and anthropogenic forcing.

These goals are addressed by two main research steps, focusing on extreme precipitation and air pollution dispersion during winter smog episodes in Belgium:

1. Validate the Belgian operational ALARO-0 NWP model for climatological time scales, by driving the model with "perfect boundary conditions" coming from global reanalyses.

2. Apply the ALARO-0 model for a dynamical downscaling of climate change projection, by driving the model with a GCM scenario.

Both statistical and dynamical downscaling have their advantage and disadvantages and should ideally be used in a complimentary way. However, statistical downscaling methods have important shortcomings relative to dynamical downscaling, where all relevant variables are modeled simultaneously and the full dynamical range in space and time is preserved. Hence, if physical consistency between variables is desired and non-stationarity of absolute values and biases is present, a dynamical downscaling is preferable (Mayer et al., 2015). Therefore, in this thesis we will limit ourselves to the *dynamical downscaling method*, and more specifically the one-way nesting approach. This approach where a RCM instead of a computational demanding high-resolution or variable-resolution GCM is used, is a very attractive way forward in regional climate modeling and regional impact assessments at a small institute such as the RMI where only limited computing resources are available.

As discussed previously, CPMs in general, and the ALARO-0 model in particular, find in many aspects their origins in NWP. CPMs had been used for decades in NWP, before they started recently to be used for climate simulations. Their parameterizations are often originally developed for the operational short-range NWP models, which is also the case for the 3MT physics package in the ALARO-0 model. The physics packages might have been adjusted to optimize forecasts for a specific region and a particular weather regime, and may need adjustments for its generalization to other regions and altered climate (Richard et al., 2002). Therefore, it is of crucial importance to validate the parameterization within the ALARO-0 model before running climate change experiments.

The 3MT physics package in the ALARO-0 model has been specifically designed for the gray-zone scales. However, the new package has been found to behave "multiscale", i.e. providing consistent NWP results for resolutions ranging from $\approx$ 10 km down to less than 4 km (Gerard et al., 2009). Furthermore, CPMs are found to add value over their driving global models in the simulation of extreme precipitation events at small spatial and temporal scales. Hence, a profound validation of the climate downscaling results from the ALARO-0 model will allow to examine to which extent the model adds valuable fine-scale temporal and spatial details (i.e. range of atmospheric and corresponding temporal scales as shown in Fig. 1.1), while being consistent with the lower resolutions in the mesoscale.

In this thesis, the dynamical downscaling of climate projections with the ALARO-0 model is performed according to the A1B scenario as described by the Intergovernmental Panel on Climate Change (IPCC). This scenario is one of the Special Report on Emission Scenarios (SRES), that cover a wide range of the main driving forces of future emissions,

from demographic to technological and economic development (Nakićenović et al., 2000). Although, the IPCC recently defined a new set of scenarios, i.e. the so-called Representative Concentration Pathways (RCPs), we will use the "old" A1B scenario as it is still the current scenario used at the RMI. Furthermore, the A1B scenario is one of the three SRES scenarios (A2, A1B and B1) that was mostly used as forcing for the projections within the ENSEMBLES project (Niehörster et al., 2008).

As mentioned previously, a climate projection with one model provides just *one* realization of the future climate, given the scenario of natural and anthropogenic forcing. Although, a future climate projection with one model does not allow to quantify the uncertainty range around the projection, we aim to assess its spread in a rather qualitative way. Furthermore, the ALARO-0 downscaling simulations as performed for this research, reach much higher horizontal and vertical resolutions, compared to the state-of-the-art GCM and RCM resolutions of roughly 1 to 2 degrees ($\approx$ 100-200 km) and $\approx$ 12 km, as used in the Coupled Model Intercomparison Project Phase 5 (CMIP5) and EURO-CORDEX project, respectively. In this thesis, the ALARO-0 simulations with a vertical resolution of 46 model levels, are carried out up to high spatial resolutions of 4 km, corresponding to the finest atmospheric (micro)scales (Fig. 1.1). Finally, this great potential of high horizontal and vertical resolutions provides relevant climate information that can be used as a forcing for impact studies on for example the Urban Heat Island effect (UHI) (e.g. Hamdi et al., 2014, 2015), extreme precipitation and the meteorological conditions which are unfavorable for the dispersion of air pollution.

1.4   Outline

This thesis is structured in two main parts. The first part (including Chapters 2 to 4) validates the ALARO-0 model for Belgium in a climate setup, and more specifically on the validation of extreme precipitation. The second part (including Chapters 5 to 6) investigates the application of the model for present and future climate impacts on extreme precipitation and the unfavorable meteorological conditions for the dispersion of air pollution. Except for the general introduction and conclusion, each chapter of this thesis is written as a research paper. At the beginning of each of the chapters which are based on published work or work that soon will be submitted, the paper and current status of the paper on which they are based, are indicated.

Chapter 2 addresses research goal 1, by describing in detail the ALARO-0 NWP model as used for the regional climate simulations in this thesis. Furthermore, the chapter discusses the major technical issues related to the nesting technique both from a general point of view as well as in the context of the ALARO-0 simulations that are carried out for this research.

The following Chapter 3 and Chapter 4 address research goal 1 and research goal 2. Both chapters validate the ALARO-0 model for (extreme) precipitation in Belgium at daily and subdaily temporal scales, respectively. For this, the model is driven by reanalyses or so-called "perfect boundary conditions", allowing to assess purely the model errors coming from the nested model. This setup allows: (i) to investigate the added value of the

dynamical downscaling at multiple spatial resolutions and temporal scales, (ii) to validate the new physics parameterization scheme, called 3MT within the ALARO-0 model, and (iii) to study the multiscale characteristic of 3MT in a climate context, as the multiscale performance of 3MT has previously only been validated in a NWP context.

Chapter 5 addresses research goal 1 and research goal 3. In this chapter, potential future changes in extreme precipitation for Belgium are studied from downscaled GCM data using the ALARO-0 model at 4 km resolution. It is first verified whether or not the climatology for the control period is correctly represented by the ALARO-0 model. Next, future changes for Belgium are assessed for mean and extreme precipitation both at daily and subdaily timescales. Finally, the uncertainty of the spread around our downscaling results are discussed qualitatively w.r.t. other results found in literature, including a recent multi-model intercomparison study, in which the ALARO-0 projection has taken part.

Chapter 6 addresses research goal 1 and research goal 3. In this chapter, two different stability indices which are based on meteorological conditions determining the dispersion of air pollution, are analyzed. This methodology, which is commonly used for alerts of winter smog peaks in a context of operational weather impact , is here for the first time applied and validated with the high-resolution climatological data from the ALARO-0 model. In a first step the statistics from both indices are evaluated for present climate conditions. A bias correction is then applied to assess the uncertainty around the derived response in the meteorological conditions for the dispersion of air pollutants under climate change.

Finally, Chapter 7 summarizes the results in a general conclusion and provides some perspectives for further research, as well as some limitations of this work.

2

ALARO-0: FROM NUMERICAL WEATHER PREDICTION TO CLIMATE MODELING

"Regional Climate Models (RCMs) are not a new concept. They are at their core limited area models that are used in Numerical Weather Prediction (NWP)."
– Rummukainen (2010)

2.1   Introduction

The ALARO-0 model [ALADIN and AROME (Application de la Recherche à l'Opérationnel à Meso-Echelle) combined model, first baseline version released in 1998] is developed and maintained mainly through a collaboration between the Royal Meteorological Institute of Belgium (RMI) and the Regional Cooperation for Limited Area modeling for Central Europe (RC LACE). The model finds its origin in Numerical Weather Prediction (NWP) and it runs operationally in a number of countries of the Aire Limitée Adaptation Dynamique développement INternational (ALADIN) and HIgh Resolution Limited Area Model (HIRLAM) consortia (Austria, Belgium, Czech Republic, Croatia, Hungary, Norway, Portugal, Romania, Sweden, Slovenia, Slovakia, and Turkey), for the national NWP applications, the first of them already since 2008. More recently, the model is also used for climate runs. In the research department of the RMI, the model is since 2010 used for regional climate simulations. This chapter discusses in detail the ALARO-0 model, the experimental setup, and technical specifications of the numerical simulations that are carried out in this thesis.

2.2   Model description

ALARO-0 is a new model version of the ALADIN model, which is the Limited Area Model (LAM) version of the Action de Recherche Petite Echelle Grande Echelle Integrated Forecast System (ARPEGE-IFS) (Bubnová et al., 1995; ALADIN international team, 1997). Since the 1990s the ALADIN model has been widely used in the NWP community and, more recently, in regional climate modeling (e.g. Radu et al., 2008; Skalák et al., 2008). The model uses a diagnostic-type deep convection and microphysics parameterization which is based on Bougeault (1985) with upgrades from Gerard and Geleyn (2005). The new physical parameterizations within the ALARO-0 model as proposed by Gerard et al. (2009), were specifically designed to be used from the mesoscale to the convection-permitting scales (so-called "gray-zone" scales) and are centered around an improved convection and cloud scheme (Fig. 2.1). ALARO-0 is based on the ALADIN hydrostatic and spectral dynamical core.

Furthermore, the ALARO-0 model utilizes: a Semi-Lagrangian Horizontal Diffusion scheme called SLHD (Váňa et al., 2008), some pseudo-prognostic Turbulent Kinetic Energy (TKE) scheme (pTKE, i.e. a Louis-type scheme for stability dependencies, but with memory, advection and auto-diffusion of the overall intensity of turbulence) and a statistical sedimentation scheme for precipitation within a prognostic-type scheme for microphysics (Geleyn et al., 2008). The physics package of the ALARO-0 model is coupled to the dynamics of the ALADIN model (Bubnová et al., 1995) via a physic-dynamics interface based on a flux-conservative formulation of the equations proposed by Catry et al. (2007). The model can be run with different schemes to impose the lateral-boundary conditions (Davies,

1976; Radnóti, 1995; Termonia et al., 2012). For the model simulations used in this thesis, the version of Radnóti (1995) is used (De Troch et al., 2013). For this thesis we use the version of the ALARO-0 model (cycle36) that was adopted for the operational applications in the RMI of Belgium in 2010.
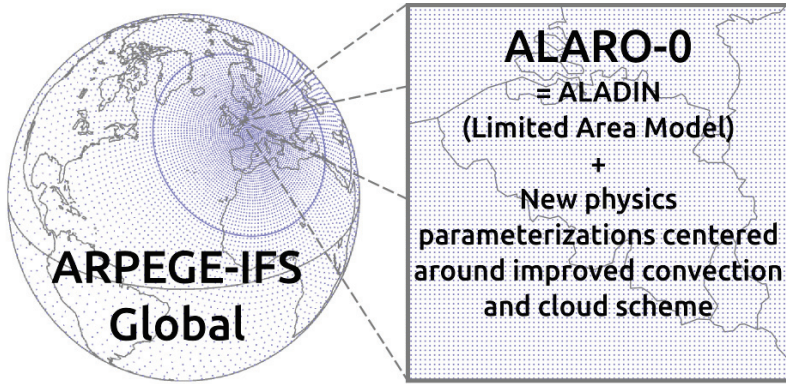
Figure 2.1: Schematical representation of the Limited Area Model (LAM) ALARO-0.

2.2.1   Deep convection parameterization

As mentioned before, the deep convection parameterization in ALADIN is based on the classical diagnostic-type scheme from Bougeault (1985) with upgrades from Gerard and Geleyn (2005). The scheme of Bougeault (1985) uses a mass-flux approach, which replaces the cloud population within a grid box by a single "equivalent updraft", and which assumes that the convective activity implies a reorganization of moisture, heat and momentum over a single vertical in the model grid (Gerard and Geleyn, 2005). The basic principle behind this classical mass-flux approach is schematically illustrated in Fig. 2.2.

For grid cell resolutions mostly below the Rossby radius of deformation for convective phenomena, the mass-flux based parameterizations generally assume that convective clouds cover only a small portion of the grid cell, so that σ_c, which is the fractional area covered by all convective clouds in the grid cell, can be assumed to be negligible with respect to "one" (i.e. $\sigma_c << 1$) (Fig. 2.2 (a)). With this assumption, the temperature and water vapor to be predicted, essentially correspond to those for the cloud environment and the value over the updraft area has no impact on the grid cell mean (Gerard and Geleyn, 2005; Arakawa and Jung, 2011). As illustrated by the green arrows in Fig. 2.2 (a), the relevant physical processes are the "cumulus-induced" subsidence in the environment and the detrainment of cloud air into the environment. However, this "cumulus-induced" subsidence is only a hypothetical subsidence. This subsidence is a component of the subgrid-scale eddy, which has by definition its own mass budget closed within the same grid. This does not imply that the true subsidence is confined within that cell. The true subsidence is the sum of the green and red vertical arrows in Fig. 2.2 (a), which normally tend to compensate each other. In such a case, the true subsidence occurs in another grid cell, which may be far away (Arakawa and Jung, 2011). Hence, the fact that the true subsidence from the updrafts is happening in a multitude of grid cells, must be taken into consideration by the parameterization schemes.

Therefore each individual grid box realization of the parameterization has a statistical view of the true subsidence or so-called "compensating subsidence" happening inside its area.
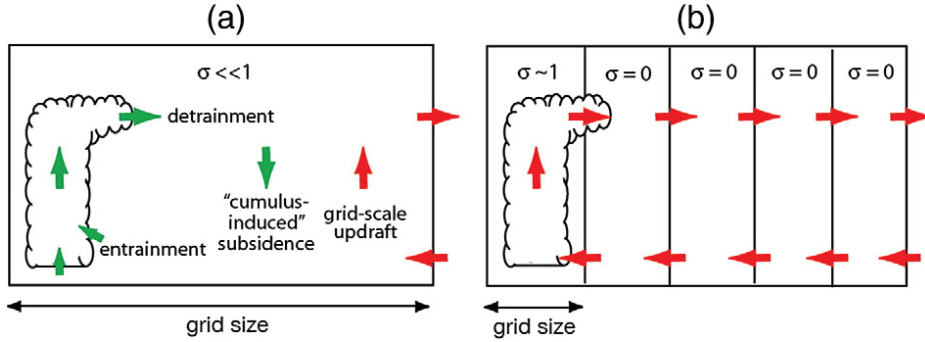
Figure 2.2: Schematic illustration of circulations associated with clouds for (a) coarse and (b) fine resolutions (adopted from Arakawa and Jung, 2011).

As long as the updraft computation can also be considered as statistical with respect to its population of updrafts of various depths and sizes, it seems not to matter much that the compensating subsidence is computed on the basis of a purely local closure. But when the mesh size becomes smaller, the cloud may eventually occupy the entire grid cell, so that there is no "environment" within the same cell and only a few updraft realizations happen inside each grid cell. As can be seen from Fig. 2.2 (b), it can in this case no longer be assumed that the convective area fractions σ_c are negligible with respect to "one", and the whole concept of the classical convective parameterization schemes collapses (Arakawa and Jung, 2011; De Troch et al., 2013).

The key concept behind the physics parameterization package of ALARO-0 lies in the precipitation and cloud scheme called Modular Multiscale Microphysics and Transport (3MT), developed by Gerard and Geleyn (2005); Gerard (2007); Gerard et al. (2009). In the 3MT scheme the problem of the classical mass-flux based convective parameterizations is addressed by combining three key features of the scheme: (i) the separately computed deep convective condensation and large scale condensation are merged as single input for a "prognostic-geometric" set of microphysical computations (sedimentation, auto-conversion, collection and melting-evaporation during fall), (ii) the convective detrainment is not diagnosed independently but becomes the result of the combined computations of closure, entrainment and condensation and (iii) the closure assumption (core of the physics-dynamics coupling) is a prognostic-type one with memory of the updraft area fraction and of the updraft vertical velocity of previous time-steps. These three interrelated characteristics of 3MT induce a good multiscale performance of 3MT, in particular in the so-called "gray-zone". The latter can be defined as the range of horizontal mesh-sizes for which the precipitating convection is partly parameterized and partly simulated by the resolved motions of the model. In case that the classical diagnostic-type schemes of e.g. ALADIN, is used at the gray-zone scales, this ambivalence results in double-counting or double-void situations, leading to several negative "gray-zone syndromes". More specifically, in convective situa-

tions drizzle appears nearly everywhere, and the precipitation maxima are too intense and too scattered: this happens especially over mountainous areas (see Chapter 3 and De Troch et al., 2013).

The unique feature of the ALARO-0 model, and in particular the physics parameterization package 3MT, to obtain consistent results at different spatial model resolutions has been intensively tested in the NWP community. For example, (Gerard et al., 2009) tested the multiscale aspect of 3MT for a case study on 2 May 2008, characterized by small-scale convective cells of medium intensity over central Europe. Figure 2.3 shows the results from the forecasts using three physical setups (i.e. 3MT, the classical diagnostic-type scheme, and no convection scheme) for three different spatial resolutions (i.e. hydrostatic: 9 km and 4.5 km, and nonhydrostatic: 2.3 km). For all three setups the same ALARO-0 schemes were used for the microphysical calculations, either only for the "resolved" part or for the 3MT accumulated-condenstation-input specific occurrence. The results indeed demonstrate the good multiscale feature of 3MT.

The multiscale characteristic of the 3MT scheme in a *climate context* is studied in detail in the next chapter (Chapter 3) that aims to elaborate on the relative importance of resolution versus parameterization formulation on the model skill to simulate realistic extreme daily precipitation. This is achieved by comparing at varying horizontal resolutions 30-yr (1961-1990) daily cumulated summer precipitation from the ALADIN model and the ALARO-0 model with respect to observations.
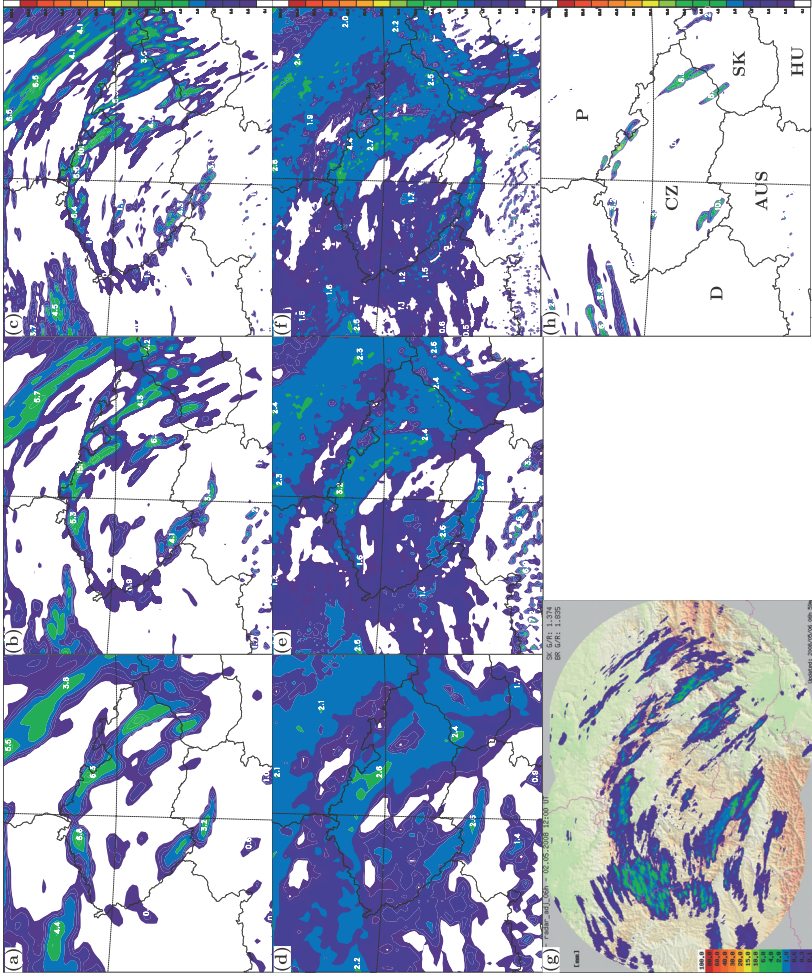
Figure 2.3: *Accumulated precipitation over central Europe between 0600 and 1200 UTC 2 May 2008. Forecasts from initial conditions of 0000 UTC at (a),(d) 9- and (b),(e) 4.5-km resolution (hydrostatic) and at (c),(f),(h) 2.3-km resolution (nonhydrostatic). (g) Scaled radar composite image. (a)-(c) 3MT, (d)-(f) diagnostic, and (h) no convection scheme (adopted from Gerard et al., 2009).*

2.2.2 Surface parameterization

For the land-surface parameterization the ALARO-0 model relies historically on the ISBA scheme (Interactions between Soil, Biosphere and Atmosphere) developed by Noilhan and Planton (1989) and Noilhan and Mahfouf (1996). The ISBA scheme describes the exchanges of heat and water between the low-level atmosphere, the vegetation, and the soil. ISBA is a relatively simple scheme, but it embraces the most important components of the land surface processes. The simplicity of the scheme is achieved by the calibration of several important coefficients with more sophisticated models and experimental data. The scheme includes the treatment of soil heat content, soil water content, water interception by vegetation and aerodynamic transfer processes in the atmospheric surface layer (Noilhan and Mahfouf, 1996).

During the last decade, a more sophisticated land surface scheme called SURFEX (SURFace EXternalisée) has been developed. SURFEX is an externalized surface scheme that can be run either in a coupled mode, in which case the atmospheric forcing is provided by the host atmospheric model (ALARO-0 in our case), or in a stand-alone mode where the atmospheric drivers are derived either from observations or model output and fed to the surface scheme such that it is decoupled from the atmospheric part of the model (Hamdi et al., 2014). For Belgium, the SURFEX land surface scheme has been mainly used for urban climate studies (e.g. Hamdi et al., 2014, 2015) and surface data assimilation (e.g. Duerinckx et al., 2015).

In SURFEX, each grid box consists of four adjacent surfaces: vegetation, urban areas, sea or ocean and lake, which are each associated with a specific parameterization. Horizontal interaction does not exist between the different surface tiles. The physiographic characteristics of each of these surfaces is provided by the global ECOCLIMAP database (Masson et al., 2003; Champeaux et al., 2005). Sea tiles use the Exchange Coefficients from the Unified Multicampaigns Estimates (ECUME) parameterization (Belamari and Pirani, 2007). Inland waters use the classical formula of Charnock (1955). The vegetated areas are parameterized with the ISBA scheme and for the parameterization of urban surfaces the Town Energy Balance (TEB) single-layer urban canopy model from Masson (2000) is used (Hamdi et al., 2014).

2.2.3 Radiation parameterization

For radiation there are two different parameterization schemes available within ALARO-0. The ALARO-0 physics has been developed with the Action de Recherche Petite Echelle Grande Echelle (ARPEGE) Calcul Radiatif Avec NEBulosité scheme for radiation (ACRANEB) based on Ritter and Geleyn (1992). This is a two-stream approximation with a Net Exchange Rate (NER) formulation for solving the thermal part. All the computations consider two spectral bands (for the solar and thermal part) with the contribution of three gases (water vapor H_2O, carbon dioxide CO_2 equivalent and ozone O_3). The gaseous optical depths are computed by means of a Padé fit of the equivalent scale width computed by the Malkmus formula (Malkmus, 1967). To cope with the high model levels, the scheme has been extended by using a Voigt-line profile (Geleyn et al., 2005). Recently, a new version of the ACRANEB scheme (i.e. ACRANEB2) has been developed (Mašek et al., 2015).

The other radiation scheme than can be used in ALARO-0 is the European Centre for Medium-Range Weather Forecasts (ECMWF) Fouquart-Morcrette Radiation (FMR) scheme. The FMR scheme is a shortwave radiation scheme (Fouquart and Bonnel, 1980; Morcrette, 1990) with 6 spectral bands, whereas the longwave radiation with 16 spectral intervals is computed by the Rapid Radiative Transfer Model (RRTM) code (Mlawer et al., 1997) using climatological distributions of ozone and aerosols.

One main difference between the ACRANEB and FMR scheme can be found in their computation approach to solve the problem of computational expensiveness of radiation schemes. FMR is called intermittently to save computing costs. More specifically, only the shortwave flux dependency on the zenithal solar angle is updated at every time step. The rest of the radiation computations are updated with a lower temporal frequency of 1 hour or 15 minutes. ACRANEB, on the other hand, is in itself designed for cost effectiveness and is called every time step (Hamdi et al., 2014).

2.2.4 Assessment on options in the surface- and radiation parameterization schemes

Table 2.1 gives an overview of the three different setups in the surface- and radiation parameterization schemes that are used for the ALARO-0 model simulations in this thesis. The greater part of the ALARO-0 simulations carried out in this thesis, have used the "default" settings given by setup A, namely the ISBA scheme for the surface and the ACRANEB scheme for radiation. The current operational version of the ALARO-0 model at the RMI is also running with this configuration, i.e. the ACRANEB and ISBA scheme.

		Radiation scheme	
		ACRANEB	**FMR**
Surface scheme	**ISBA**	Setup A - Chapter 3, 4, and 5	/
	SURFEX	Setup B - Chapter 6	Setup C - Chapter 6

Table 2.1: *Overview of the surface- and radiation schemes used in the different chapters of this thesis.*

Setup B and C both use SURFEX, but rely on a different radiation scheme (ACRANEB and FMR, respectively). As discussed previously in the General introduction (Chapter 1), the nested model can be driven at its boundaries by global model data coming from either reanalyses or either a GCM. In this thesis, the GCM driving data for the ALARO-0 model simulations is given by the GCM CNRM-CM3, which uses the FMR radiation scheme (Salas-Mélia et al., 2005). Hence, to account for a more consistent use of the radiation scheme within the driving GCM CNRM-CM3 and the nested ALARO-0 model, the GCM CNRM-CM3 driven simulations with ALARO-0 in Chapter 6 are performed with the FMR scheme. The reanalysis driven ALARO-0 simulations in this Chapter 6 are carried out with the "default" ACRANEB radiation scheme. Furthermore, for the simulations using SURFEX for the surface parameterization, the TEB scheme, which accounts for the simulation of the interactions with urban areas, is always switched on. The sensitivity and relative impact to the

use of the more sophisticated surface scheme SURFEX and a different radiation scheme is assessed through comparison of daily precipitation and daily minimum 2-meter temperature from simulations using the three setups (A to C) w.r.t. the observations (see also Giot et al., 2014). All three simulations are performed with the ALARO-0 model at 4 km horizontal resolution, using the ERA-Interim reanalysis as Lateral Boundary Conditions (LBC). The 30-yr mean (1981-2010) spatial distribution of precipitation and minimum temperature for each of the three model setups as well as for the observations (colored circles) are shown in Fig. 2.4 and Fig. 2.5, respectively. Differences in the mean spatial distribution between the different setups are shown in the bottom rows of the figures.
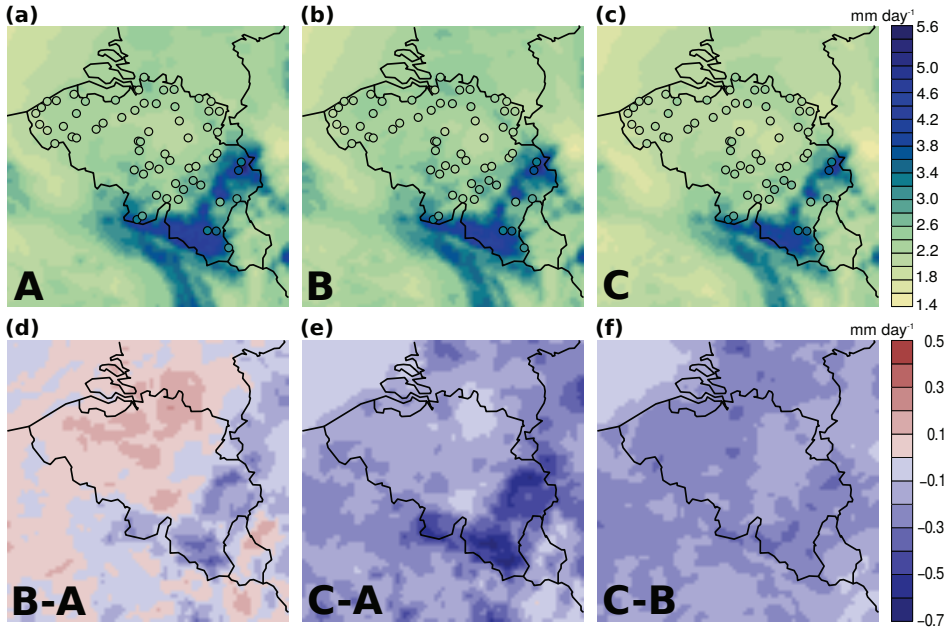
Figure 2.4: *(top) 30-yr mean (1981-2010) spatial distribution of daily precipitation for the different model setups (i.e. A-C) as described in the text. Colored circles show the 30-year mean observed values for daily precipitation from 63 climatological stations. (bottom) Differences in daily precipitation between the different setups.*

Figure 2.6 shows the observed and modeled frequencies for each of the setups for summer (June-July-August, JJA) and winter (December-January-February, DJF) daily precipitation and minimum 2-meter temperature. The modeled frequencies are calculated for the closest model grid points to the station locations. Overall, it can be seen from Figs. 2.4 to 2.6 that all model setups are able to reproduce reasonably well the observed values. For example, the observed precipitation maximum in the south east of the country is well captured by all three model simulations (Fig. 2.4). Only for the lowest minimum temperatures during the winter (DJF), all models show an underestimation (Fig. 2.6 (d)). However, some specific aspects in the sensitivity related to the surface- and radiation parameterization scheme are visible. As can be seen from Fig. 2.6 (a) and (c) summer (JJA) precipitation is more sensitive to the radiation scheme, while for minimum temperature the effect of a different

surface scheme is more pronounced. The use of the FMR radiation scheme (setup C) implies systematically lower precipitation amounts w.r.t. observations than with the ACRANEB scheme (setup A and B) which has specifically been designed for ALARO-0 (Fig. 2.4 (e) and (f) and Fig. 2.6 (a) and (b)). Furthermore, Fig. 2.4 (d) and (e) demonstrate that precipitation values from the model simulations with SURFEX (setup B and C) are lower (higher) at higher (lower) altitudes than with ISBA (setup A). The use of the more sophisticated surface scheme SURFEX (setup B and C) shows more spatial detail for minimum temperature compared to ISBA (setup A), which shows a smoother spatial distribution. This is also reflected by the use of TEB within SURFEX in setup B, resulting in higher and more realistic minimum temperatures in urban areas than it is the case with ISBA (setup A).
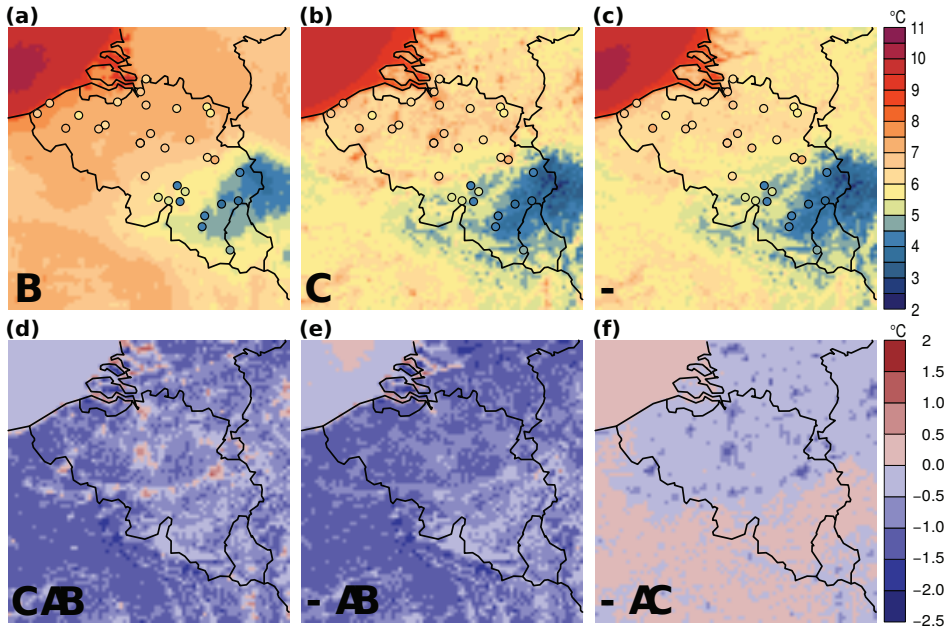
Figure 2.5: *As in Fig. 2.4, but for daily minimum 2-meter temperature. Colored circles show the 30-year mean observed values for daily minimum 2-meter temperature from 33 climatological stations.*

   This assessment on the use of different land-surface schemes and radiation schemes shows that depending on the meteorological variable or phenomena (such as Urban Heat Island (UHI)) one aims to study, one or another setup is preferable. Except for Chapter 6, the focus in this thesis is mainly limited to precipitation as climatological variable. The above results demonstrate for precipitation a minor sensitivity to the surface scheme, and smaller differences in precipitation w.r.t. the observations with the ACRANEB scheme, which has been specifically designed for ALARO-0, than with the FMR scheme. From this, it can be concluded that w.r.t. the other setups, the "default" settings of setup A using the ACRANEB and ISBA scheme, is an acceptable setup for the simulation of precipitation with the ALARO-0 model. In the following chapters (Chapters 3 to 5) this will be further elaborated through extensive validation of the precipitation characteristics from the ALARO-0

model using the 3MT, ISBA and ACRANEB parameterization schemes.
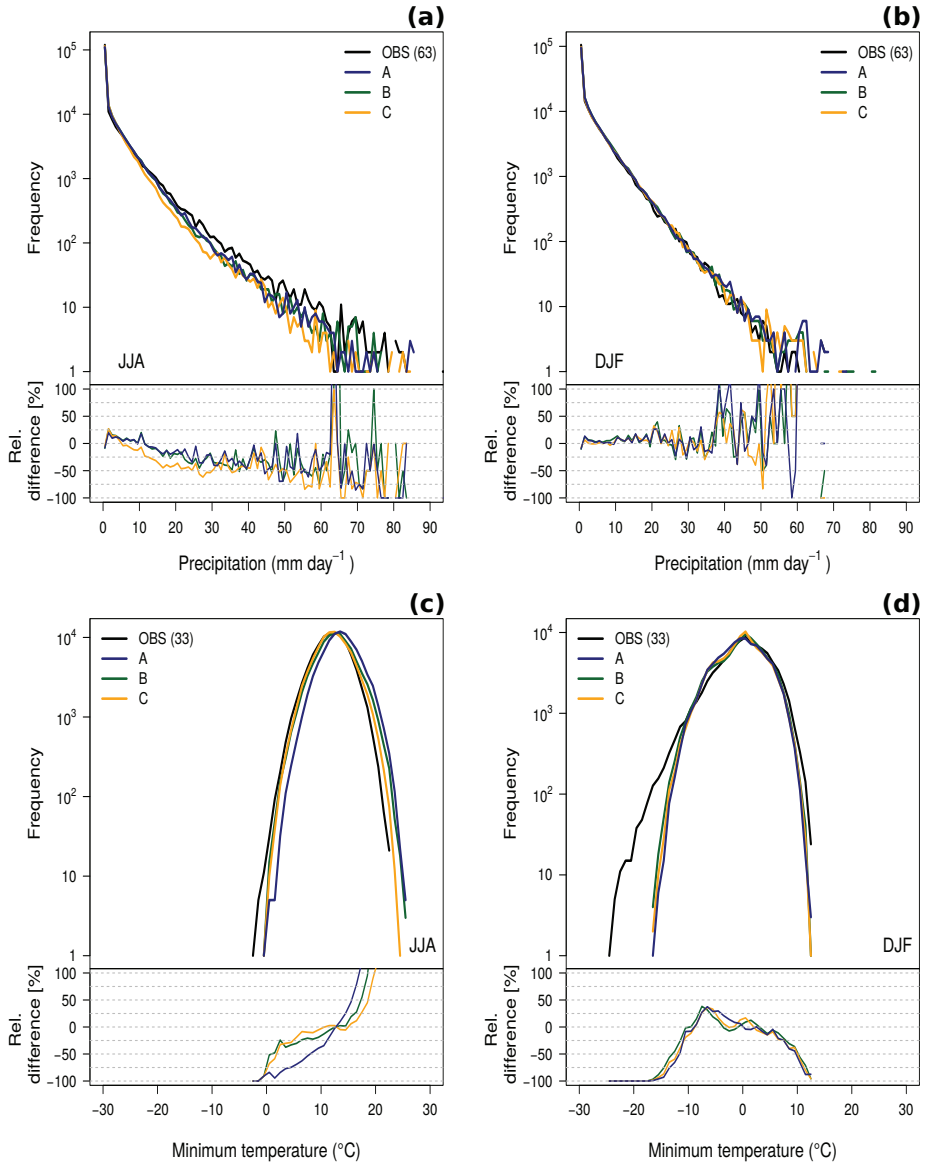
Figure 2.6: *Frequencies of observations and the different model setups (i.e. A-C) as described in the text. Frequencies are computed with the 30-yr (1981-2010) daily summer (June-July-August, JJA) and winter (December-January-February, DJF) cumulated precipitation ((a) and (b)) and minimum 2-meter temperature ((c) and (d)) given for each observation station separately and are displayed on a logarithmic scale. Relative differences ((setup-obs)/obs) are also shown in the bottom of each figure.*

2.3   Dynamical downscaling

As mentioned in the General introduction (Chapter 1), the use of nested LAMs or RCMs as a climate downscaling technique involves a number of technical issues. These issues are discussed in the next sections both from a general point of view as well as in the context of the simulations that are performed in this thesis.

2.3.1   Lateral Boundary Condition problem

In contrast to NWP, which is an initial value problem, regional climate modeling can be considered as a boundary value problem. There are multiple LBC issues related to the nesting technique. First, there is the drawback that one is obliged to impose imperfect LBCs inducing various errors at the boundaries of the LAM (e.g. Warner et al., 1997; Termonia et al., 2009). Mathematically, the methods to treat the LBC in nested LAMs or RCMs are "ill-posed" problems, signifying that a unique solution does not exist and that it is thus not possible to specify exactly the right conditions. Furthermore, differences in the spatial and time resolution, as well as in the process description between the driving and regional model, can distort wave propagation and reflection properties and lead to numerical instabilities. To address this ill-posedness, most RCMs developed the so-called "relaxation" method. The idea of this method is to build in a relaxation or adjustment zone of a particular width of number of grid points, in between the global and nested model. In this way, boundary conditions and the solution computed in the RCM are merged in a way as to dampen spurious numerical features (Giorgi and Mearns, 1999; Rummukainen, 2010). For the simulations in this thesis a relaxation zone consisting of 8 grid points is used, irrespective of the spatial resolution.

Furthermore, there is the problem that possible errors in the large-scale circulation produced by the driving model will be transmitted to the nested model. This problem is often referred to as the "garbage in, garbage out" problem; meaning that the quality of the simulated regional climate is conditional on the quality of the LBCs which are provided by the underlying global data or global model simulation. The problem is in particular relevant when one drives the regional model with a GCM, seen their inability to accurately simulate the regional features (Giorgi and Mearns, 1999; Rummukainen, 2010).

Despite these LBC issues, past and current applications with RCMs have shown that the one-way nesting strategy is a workable solution (Giorgi and Mearns, 1999). In order to minimize the effects of the LBC problems, Giorgi and Mearns (1999) recommend to first validate the model for the current climate using analyses of observations, i.e. the so-called "perfect boundary conditions".

Interesting work has been carried out by de Elía et al. (2002) and Denis et al. (2002) with a perfect-model approach or so-called Big-Brother Experiment (BBE). The BBE approach is schematically represented in Fig. 2.7. The first step of the BBE approach consists in running a global high-resolution model to produce a high-resolution reference dataset. Then, the small scales existing in that reference dataset are filtered to generate a low-resolution dataset needed to drive the nested RCM. Therefore only the large and medium scales would be used to feed the Little-Brother RCM such as to mimic the nesting of a RCM by a GCM. The reference dataset (before filtering) however would contain small scales

against which the RCM small scales could be validated. Hence, the objective is to assess to what extent the Little-Brother simulation is able to reproduce the Big-Brother small scales when it is driven with only the large and medium scales of its Big Brother. The BBE has the main advantage that the differences can be solely attributed to the nesting strategy (i.e. such as the specification of boundary conditions and model domain) since the model resolutions, physics, dynamics and numerics are the same and therefore the model errors are also the same. However, the high computational cost of running the global high-resolution simulation is an important disadvantage of the BBE. For this reason, both de Elía et al. (2002) and Denis et al. (2002) applied a modified version of the BBE, that does not employ a global model but the same RCM for both the Big and Little Brother. The Big Brother was run over a much larger domain than the Little Brother but still smaller compared to a global domain (Denis et al., 2002).
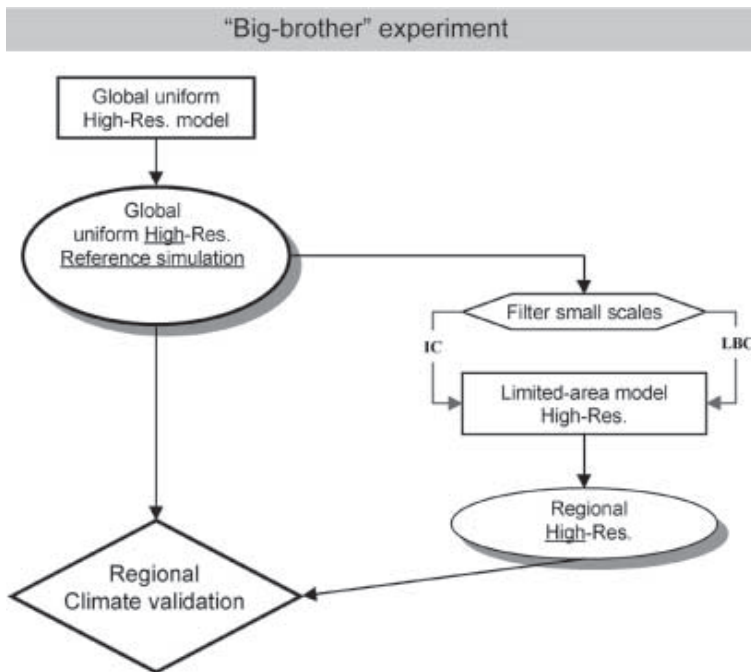
Figure 2.7: The Big-Brother Experiment flowchart. Rectangles are the models and ovals are their corresponding datasets. The diamond represents validation of the Little-Brother regional-scale features against those existing in the reference Big-Brother dataset. The Initial Conditions (IC) and Lateral Boundary Conditions (LBC) for the Limited Area Model (LAM) (right branch) are spatially filtered such that the small scales are removed (adopted from Denis et al., 2002).

Both studies of de Elía et al. (2002) and Denis et al. (2002) show that, in a downscaling with a one-way nesting, a LAM or RCM is able to regenerate the correct amount of variability at the scales smaller than the ones of the driving model in which the high-resolution variability had been removed by filtering. However, de Elía et al. (2002) found that the LAM is not capable of reproducing the correct details with sufficient precision required by the root

mean square errors (RMSE), i.e. that the variables locally in space and time do not fully reproduce the ones of the perfect model run. Whereas de Elía et al. (2002) concentrated on the short-term evolution of weather systems and quantified the models' ability to simulate the data in a deterministic day-by-day basis by means of RMSE, Denis et al. (2002) focused on climate timescales and demonstrated the ability of high-resolution RCMs to gain accuracy in a climatic-statistical sense.

Therefore, for studying the climate of weather extremes it is rather the statistics of the extremes that are important, provided the large-scale evolution is consistent with the large-scale flow of the driving model. This is an important additional criterion in deciding to use RCMs with respect to global ones and it is further discussed in the next section.

2.3.2   Long run versus daily reinitialization

For long-range runs, at temporal scales of multiple decades, there is indeed the problem that the internal climate of the RCM or LAM can start to diverge from the large-scale atmospheric circulation given by the driving global model (Nicolis, 2003; Qian et al., 2003; Nicolis, 2004). The reason for the deviations of the large-scale atmospheric circulation of the RCMs from that of the driving field can be attributed to (i) the errors, missing processes or the coarse representation of surface forcings such as the topography and land-sea mask in RCMs, and (ii) the chaotic nature of the weather simulated by the RCMs, in which different solutions can emerge from simulations started with slightly different initial conditions but using the same lateral boundary conditions, the latter often referred to as the internal variability (Lucas-Picher et al., 2015). One can deal with these deviations by either (i) interrupting the model runs of the LAM after a few days and restart them, while taking into account a spinup period which allows that the physics can adjust, or (ii) carrying out uninterrupted model runs over long periods, allowing the LAM to find its own climate equilibrium (Qian et al., 2003).

In the second case, one can for instance apply a large-scale nudging (LSN) or spectral nudging where a fraction of the time-variable large-scale atmospheric states from the driving global fields is imposed during the integration of the RCM. In this way, the large scales of the RCM are forced to follow both the large scales provided by the driving global model at the lateral boundaries as well as the large-scale conditions in the domain's interior (Lucas-Picher et al., 2015).

Within the RCM community there is still no consensus on the relevancy of LSN. Several studies have demonstrated the positive effects of LSN with an improvement of the large-scale atmospheric circulation of the RCM, especially at the outflow lateral boundary where inconsistencies between the RCM and its driving field can affect the simulated climate (Miguez-Macho et al., 2004; Castro et al., 2005). In a recent study, Lucas-Picher et al. (2015) compared two RCM simulations that are driven by the ERA-Interim reanalysis, but where one RCM uses additional LSN. The authors investigate the ability of both RCM simulations to reproduce weather regimes over North America. Their results showed that a classic RCM, only driven at the lateral boundaries, simulates the mean statistics of the weather regimes well, but on a day-to-day and seasonal scale the classic RCM simulation, simulated many days where the weather regimes are different from that of the driving field. To reproduce as best as possible daily weather regimes and seasonal anomalies, LSN is

recommended by the authors. Besides these improvements by means of LSN, it has also been shown that nudging induces disadvantages such as for example affecting the simulation of precipitation extremes. Furthermore, one can argue that the use of LSN makes the simulation less internally consistent and can downplay potential feedbacks from simulated local processes on to the larger scale (Rummukainen, 2010). As stated by Rummukainen (2010), LSN could be an option when relatively large domains are used, since small domains as such offer less opportunity for the regional model solution to diverge from its boundary conditions (see also further in Section 2.3.3).

In this thesis, we will study whether the internal climate variability generated by the higher resolution of the RCM and its model physics as identified by Denis et al. (2002) and de Elía et al. (2002) reproduces the correct statistics. Hence, we use the nesting approach originating from NWP, consisting of a pure downscaling with reinitializations based on the concept of one-way nesting. Lucas-Picher et al. (2013) demonstrated that dynamical downscaling with reinitializations has lower systematic errors than with a standard continuous model configuration. The downscaling approach as used in this thesis is demonstrated in Fig. 2.8.
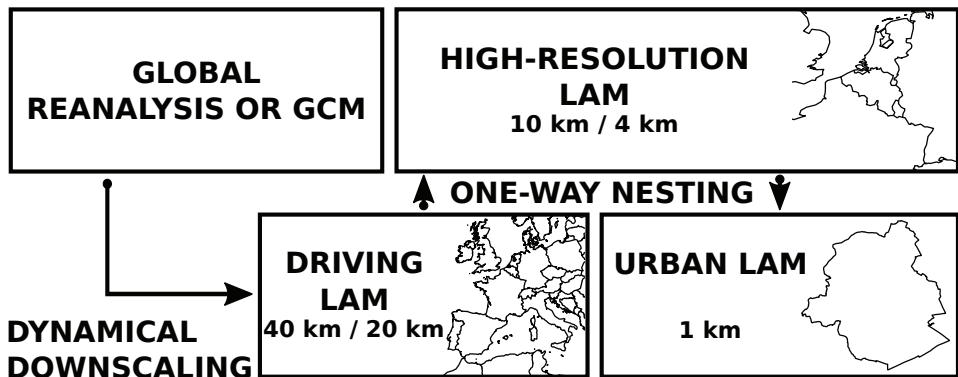
Figure 2.8: Schematic overview of the downscaling approach as used in this thesis.

Global climate model data from reanalyses (either 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) or ERA-Interim, Uppala et al., 2005; Dee et al., 2011) or the GCM CNRM-CM3 (Salas-Mélia et al., 2005) are interpolated to a 40 or 20-km resolution domain that encompasses most of western Europe. These 6-h files serve as initial and boundary conditions for 48-h ALADIN or ALARO-0 runs at 40 or 20 km resolution. These are started at 0000 UTC every day. The (3 h) output from these first runs serves as input for the high-resolution 10-km and 4-km runs on a domain centered on Belgium. However, to exclude spinup problems, the first 12 h are not taken into account. So we have 36 h of data left for the 4-km and 10-km runs (which thus start at 1200 UTC). Finally, we again dismiss the first 12 h of the runs, to arrive at 24 h of output for the regional climate simulations at 4-km and 10-km resolution, and then integrate/reinitialize over each subsequent 24-h period.

In general, the climatology of a regional atmospheric model is determined by a dy-

namical equilibrium between two factors: (i) the information provided by the LBCs, and (ii) the internal model physics and dynamics. The atmospheric spinup time of the model thus corresponds to the time taken by the lateral boundary information to get through the model domain and generate this dynamical equilibrium. The atmospheric spinup time of the traditional continuous long-term climate simulations varies depending on the domain size, season, and circulation strength, but it is typically of the order of several days (Giorgi and Mearns, 1999). Hence, our setup with daily reinitializations justifies our relatively short spinup time of 12 h. Furthermore, in the case that SURFEX is used as land-surface scheme, the high-resolution 4-km runs can be further downscaled to an urban scale such as e.g. the Brussels Capital Region. In that case, SURFEX is used in offline mode where the 4-km resolution atmospheric forcing is simply projected on a 1-km grid by searching for the closest grid point. However, the daily reinitialization will limit the equilibration of the surface physics (soil moisture and temperature), which is particularly desirable in long-term regional climate modeling (Giorgi and Mearns, 1999). Therefore, for the simulations where SURFEX is employed, the soil variables evolve freely after initialization and are never corrected or nudged in the course of the simulation (Hamdi et al., 2014, 2015). However, for the simulations using the ISBA scheme, the surface variables are daily reinitialized. Therefore, as a future outlook it would be beneficial to test this setup with the ISBA scheme where the soil variables evolve freely.

The technical aspects of our downscaling approach using a one-way nesting, gives rise to two other important issues related to the LBC provision (as discussed previously in Section 2.3.1): (i) the resolution difference between the driving data and the nested model, and (ii) the temporal updating frequency of the LBCs. Denis et al. (2003) investigated the sensitivity of a one-way nested RCM to these two issues by means of a BBE. With an RCM at 45 km resolution, their results demonstrated for spatial resolution jump ratios up to 12 between the resolution of the nesting LBC data and the RCM, a reliable regional climate for most studied fields. For the LBC update interval, 12 h appears to be sufficient. However, an update frequency of 6 h was showed to be significantly better, and is strongly recommended by the authors since there is little additional computational cost related when doing so. Furthermore, the authors found no improvement by increasing the update frequency from 6 to 3 h. Although, these results are carried out with one RCM at a resolution of 45 km and using a specific domain size, it gives an indication to what extent the resolution jump ratios and update frequency as used for our downscaling simulations are in line with the findings of Denis et al. (2003).

Table 2.2 summarizes for our downscaling simulations the ratios in spatial resolution between the different global driving datasets and the nested 40-km and 20-km ALARO-0 model. Except for the 20-km simulation driven by the GCM CNRM-CM3, the resolution jump ratios are all smaller than the maximum ratio of 12 as has been suggested by Denis et al. (2003). The resolution jump ratios between the coarse 40-km and 20-km nested models and the high-resolution nested models are also not higher than 12. Finally, according to the results of Denis et al. (2003), our update frequency of 6 h between the global driving data and the nested models appears to be sufficiently high.

	40-km nested model	20-km nested model
ERA-40 reanalysis $\approx$ 125 km resolution (Uppala et al., 2005)	≈ 3	≈ 6
ERA-Interim reanalysis $\approx$ 80 km resolution (Dee et al., 2011)	2	4
GCM CNRM-CM3 $\approx 2.8° \approx$ 300 km resolution (Salas-Mélia et al., 2005)	≈ 8	15

Table 2.2: *Overview of the resolution jump ratios between the different global driving datasets and the nested 40-km and 20-km ALARO-0 model as used in this thesis.*

2.3.3   Domain size and spatial resolution

The choice of domain size and resolution are two other important parameters that can strongly affect the RCM numerical solutions. Generally, both elements are determined by a compromise between physical and computational considerations. On the one hand the domain should be large enough so that the small scale features can fully develop and include relevant regional forcings, and the resolution should be high enough to sufficiently capture the scale and effects of such forcings (e.g. topography). However, an increase in model domain size and resolution inevitably imply an increase in computational cost (Giorgi and Mearns, 1999; Leduc and Laprise, 2009). On the other hand, if no LSN is applied the model domain should not be too large too avoid departures from the driving data (Leduc and Laprise, 2009).

Giorgi and Mearns (1999) formulate some guidelines for the choice of model domains. First, the authors state that it is useful to choose a domain in which the area of interest is as far as possible from the lateral relaxation zone. This to avoid that the solution over the area of interest would be affected by possible LBC errors (see previously Section 2.3.1). Secondly, the model domain should encompass all regions that include forcings and circulations which directly affect climate over the area of interest. Finally, it is also preferable to place the lateral boundaries over the ocean rather than over land in order to avoid possible effects of unrealistic surface energy budget calculations near the boundaries (Giorgi and Mearns, 1999). Several studies have assessed the sensitivity of regional climate simulations to domain size. Leduc and Laprise (2009) evaluated the sensitivity of an RCM to its domain size by using a BBE allowing to identify and quantify the errors due to the size of the spatial domain of the RCM. Their results suggest that for domain sizes which are too small, the "spatial spinup" or characteristic distance from the lateral boundaries will be insufficient to allow the development of the small-scale features.

The choice of the domain size has also turned out to play a very important role in the RCM simulations performed in this thesis. Figure 2.9 shows the 181×181 grid point 4-km model domain as used for the high-resolution ALADIN and ALARO-0 simulations, together with a small 4-km domain of 80×80 grid points that mainly encompasses Belgium.
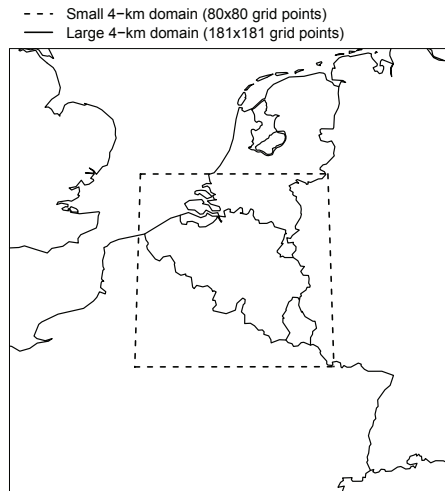
Figure 2.9: *Small and large 4-km domain. The large 181 × 181 grid point domain corresponds to the domain as used for the high-resolution 4-km ALADIN and ALARO-0 simulations.*

Precipitation fields obtained from an ERA-40 reanalysis downscaling with ALARO-0 at both domains are compared with each other. The spatial distribution of 30-yr (1961-1990) mean summer precipitation (JJA) for both domains are shown in Fig. 2.10.
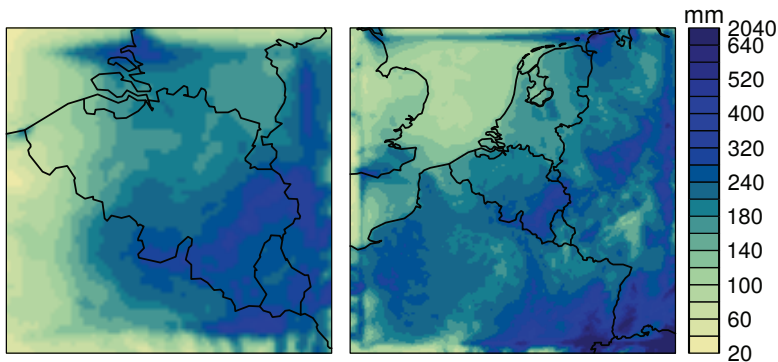
Figure 2.10: *Spatial distribution of 30-yr (1961-1990) mean summer precipitation (June-July-August, JJA) from an ERA-40 downscaling with ALARO-0 at 4 km resolution at a small 80 × 80 grid point domain (left) and large 181 × 181 grid point domain (right).*

Spurious effects at the lateral boundaries are visible for both domains. However, as it is also suggested by Giorgi and Mearns (1999) the small domain is too small so that these spurious effects enter the region of interest (i.e. Belgium) (Fig. 2.10, left). Furthermore, the different domain sizes also affect the actual precipitation values. For both domains the orographic precipitation effects in the southeast are captured, nevertheless precipitation amounts for the small domain are generally lower than for the large domain. Figure 2.11 shows histograms of the 30-yr (1961-1990) daily summer precipitation (JJA) for both do-

mains together with the observed frequencies obtained from 93 climatological stations that cover all of Belgium. The modeled frequencies are given for the closest model grid points to the station locations. As could be expected, the precipitation amounts from the large domain correspond clearly better to the observations than those from the small domain. From the relative differences in the bottom figure it is visible that precipitation values from the small domain are underestimated (overestimated) for the low (high) precipitation amounts.
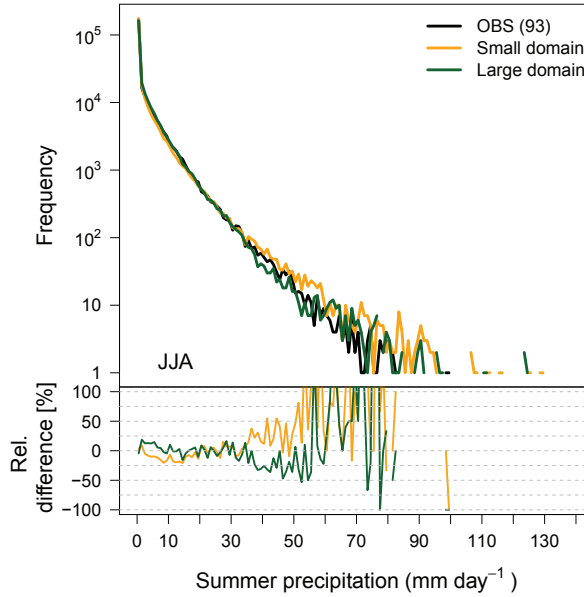
Figure 2.11: *Frequencies of observed and modeled 30-yr (1961-1990) daily summer precipitation (June-July-August, JJA), given for each observation station separately and displayed on a logarithmic scale. Relative differences ((setup-obs)/obs) are also shown in the bottom figure.*

Seth and Giorgi (1998) also found that the domain size and location of the lateral boundaries have an effect on the simulation of summer precipitation in North America as well as on its sensitivity to initial soil moisture. The authors demonstrated that the small domain captures better the observed precipitation amounts due to the fact that the boundaries constrain the interior solution more to the driving field. However, the sensitivity of precipitation to the initial soil moisture appears to be more realistic in the larger domain than in the small domain. The latter yields an unrealistic response to internal forcings which are not consistent with the applied large-scale forcing (Seth and Giorgi, 1998).

Overall, it is clear that the domain size and location of the lateral boundaries play a very important role in the RCM numerical solutions, where the choice of the domain size may simply rely on a trial-and-error approach and an assessment of how sensitive the solution is on the placement of the lateral boundaries (Giorgi and Mearns, 1999).

The spatial resolution of the RCM should be high enough so that scales of forcings that one wants to study as well as relevant scales of motions are captured. It is also important

that the models' resolution provides useful information in case that the model output is used for specific applications (Giorgi and Mearns, 1999). As extreme precipitation and climate impacts for Belgium are the main subjects of this thesis, the spatial resolution of the simulations goes up to high resolutions of 4 km. This spatial resolution of 4 km is a compromise between (i) the minimum resolution to capture the convection processes, which are mostly relevant for extreme precipitation events, and (ii) the computation time of 30-yr simulations on the 181×181 grid point domain. At the RMI computing infrastructure, one day of a 36-h run takes ≈ 9 hours. An overview of the different numerical simulations with the ALARO-0 and ALADIN model in this thesis are presented in Table 2.3. The LBCs, horizontal resolution and corresponding time steps of the different simulations are also given.

Name	LBC	Δx (km)	Δt (s)
(ERA40-)ALR04 ERAINT-ALR04 CTL SCN	(ERA40-)ALR40 ERAINT-ALR20 Chapter 5: CTL-ALR40, Chapter 6: CTL-ALR20 Chapter 5: SCN-ALR40, Chapter 6: SCN-ALR20	4	180
(ERA40-)ALR10 ALD10	(ERA40-)ALR40 ALD40	10	450
ERAINT-ALR20 CTL-ALR20 SCN-ALR20	ERA-Interim reanalysis GCM CNRM-CM3 (control climate) GCM CNRM-CM3 (future climate, A1B scenario)	20	900
(ERA40-)ALR40 ALD40 CTL-ALR40 SCN-ALR40	ERA-40 reanalysis ERA-40 reanalysis GCM CNRM-CM3 (control climate) GCM CNRM-CM3 (future climate, A1B scenario)	40	300

Table 2.3: Overview and specifications of the numerical simulations in this thesis. The LBCs used to drive the ALARO-0 model, and the applied horizontal grid spacings (km) and time steps (s) are listed.

Furthermore, as discussed in the General introduction (Chapter 1), the spatial resolution is closely related to the physical parameterization. For example, when the resolution of the model is refined far beyond 10 km, the model physics starts to solve the deep convection and cloud microphysical processes explicitly (Giorgi and Mearns, 1999). As described previously, the ALARO-0 model uses the physical parameterization package 3MT that counts for the cloud and precipitation processes. This package is specifically developed in such a way that regardless of the spatial resolution it determines which cloud fraction is stratiform or convective (subgrid) and should be resolved or parameterized respectively. This multi-scale characteristic of 3MT thus allows that the model can be applied at different spatial resolutions without the need of different schemes for different spatial resolutions, or without the need of any particular tuning. Since the multiscale performance of 3MT had previously only been validated in a NWP context up to a spatial resolution of 4 km (see Gerard et al., 2009), the multiscale characteristic of 3MT is in detail evaluated in a climate context in the next Chapter 3.

2.4 Conclusion

This chapter discussed in detail the ALARO-0 model, the experimental setup, and the technical specifications of the numerical simulations that are carried out in this thesis. The Belgian operational ALARO-0 NWP model, which is a new version of the ALADIN model, and its parameterization schemes for deep convection, surface and radiation were described in detail. The unique characteristic of the ALARO-0 model is centred around an improved physical parameterization package for precipitation and clouds, called 3MT. The basic concepts of 3MT and its unique multiscale feature in the context of NWP were given. This yet highlighted the importance of the new physics parameterization scheme for the simulation of convective and extreme precipitation.

Furthermore, the different options in the surface and radiation parameterization schemes of the model were discussed in this chapter. Historically, the ALARO-0 model relies for the land-surface parameterization on the ISBA scheme. However, during the last decade, the more sophisticated land surface scheme called SURFEX has been developed. For radiation there are two different parameterization schemes available within ALARO-0: the ACRANEB scheme and the ECMWF FMR scheme. The greater part of the ALARO-0 simulations carried out in this thesis have used the "default" settings using the ISBA and ACRANEB scheme, which is also the configuration as used for the current operational NWP applications of the model. However, for the analysis of Chapter 6, the simulations are performed with the SURFEX and FMR scheme. Therefore, the sensitivity and relative impact to the use of the more sophisticated surface scheme SURFEX and a different radiation scheme was assessed. This was done through comparison of 30-yr daily precipitation and daily 2-meter minimum temperature from ERA-Interim driven ALARO-0 simulations at 4 km resolution, using the different options in the surface and radiation scheme. The results from this assessment demonstrated for precipitation a minor sensitivity to the surface scheme, and smaller differences in precipitation w.r.t. the observations with the ACRANEB scheme, which has been specifically designed for ALARO-0, than with the FMR scheme. From this, we could conclude that w.r.t. the other setups, the "default" settings using the ACRANEB and ISBA scheme, is an acceptable setup for the simulation of precipitation with the ALARO-0 model. In particular, because, except for Chapter 6, the focus in this thesis is mainly limited to precipitation as climatological variable.

The experimental setup used for the regional climate simulations in this thesis are based on the widely used "nesting technique". This technique, which originates from NWP, consists of a pure downscaling with reinitializations based on the concept of one-way nesting. The technical issues related to this nesting technique have been discussed. One of these issues is related to the LBC problem, where possible errors in the large-scale circulation produced by the driving model will be transmitted to the nested model. In order to minimize the effects of this problem, it is recommended to first validate the model for the current climate using analyses of observations, i.e. the so-called "perfect boundary conditions". Therefore, in the next chapters (Chapter 3 and 4), LBCs from reanalyses of ERA-40 as well as ERA-Interim will be used to drive the ALARO-0 model for the validation of (extreme) precipitation in Belgium at daily and subdaily temporal scales as well as at multiple spatial resolutions. This setup will indeed allow us to validate in a climate context the main feature of the ALARO-0 model, i.e. the new 3MT physics parameterization package, and its

multiscale characteristic.

3

MULTISCALE PERFORMANCE OF THE ALARO-0 MODEL FOR SIMULATING EXTREME DAILY SUMMER PRECIPITATION CLIMATOLOGY IN BELGIUM

> *"[...] to minimize the effects of model resolution, [...] the most general but also most challenging approach is to design a scheme that would give internally consistent results on a wide range of resolutions."*
> – Giorgi and Marinucci (1996)

3.1   Introduction

Extreme precipitation events have a large impact on societies through damage caused by floodings, landslides and snow events. Precipitation is thus an important meteorological variable in weather prediction and climate studies. Herrera et al. (2010) studied the ability of regional climate models (RCMs) to reproduce the mean and extreme precipitation regimes over Spain using a state-of-the-art ensemble of RCM simulations. The RCMs show a good agreement with the observed mean precipitation regime, but for the extreme regimes the models reveal important limitations.

As described in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (AR4 IPCC), the model skill to simulate realistic extreme daily precipitation strongly depends on the spatial resolution and convective parameterization of the model (Randall et al., 2007). However, it is not straightforward to quantify the relative contribution of an increase in spatial resolution versus an improvement in physical parameterization of deep convection on the overall performance of the model.

On the other hand, precipitation is one of the most sensitive quantities to the different parameterization schemes of the climate models and to their interplay with the dynamics of the atmosphere represented in the models. For this variable it has been shown that RCMs are able to add significant information to the driving global simulations, both in space and time (e.g. Jones et al., 1995; Durman et al., 2001; Jones et al., 2004). In general terms, the RCMs produce an intensification of precipitation with respect to the driving Global Climate Model (GCM), related to the intensification of the hydrological cycle (Jones et al., 1995; Durman et al., 2001; Buonomo et al., 2007). Giorgi and Marinucci (1996) assessed several numerical experiments using a RCM driven by analyses to investigate the model sensitivity of different precipitation parameterizations to model resolution and to the resolution of topographic forcing. The authors performed two months of simulations over Europe for January and July 1991 at three different spatial resolutions of 200, 100 and 50 km and with various topography configurations. For all simulations the continental-scale average precipitation as well as the precipitation frequencies and intensity distributions showed a greater sensitivity to resolution than to topographic forcing. However, precipitation was also found to be sensitive to the parameterizations used, in particular during summer. As stated by the authors, these findings indicate that on the continental scale, when increasing the resolution, the effects of physical forcing (e.g. better representation of topography and coastlines) may be masked by the direct sensitivity of the model parameterization to resolution itself. Hence, without specific tuning of the parameterization, an increase in model resolution does not necessary result in an improvement in the simulation of precipitation (Giorgi and Marinucci, 1996). In a more recent study of Lynn et al. (2010), a RCM with different physics components at two different spatial resolutions was tested. Their results demonstrated a sensitivity of the RCM to the choice of the convective parameterization, leading to significantly different summer

precipitation outcomes. The authors conclude that these differences are due to differences in the convective parameterizations and not because of the change in spatial resolution of the model.

The aim of this chapter is to elaborate on the relative importance of resolution versus parameterization formulation on the model skill to simulate realistic extreme daily precipitation. This is achieved by comparing at varying horizontal resolutions the ALADIN model with an improved version of the model that has been updated with physical parameterizations, the so-called ALARO-0 model (see Chapter 2 for a detailed model description). The version of the ALARO-0 model used here was adopted in 2010 for the operational applications in the Royal Meteorological Institute of Belgium (RMI). Since then the model has undergone systematic verification with respect to observations at 7 km resolution. Gerard et al. (2009) tested the new parameterizations within the ALARO-0 model in a 1-day case-study over Belgium, which was characterized by heavy convective precipitation. From this study an improvement of ALARO-0 at varying horizontal scales has been demonstrated.

The 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) (Uppala et al., 2005) is used as large-scale coupling data to drive the coupled models, ALARO-0 and ALADIN. As suggested by Giorgi and Mearns (1999), atmospheric reanalyses, such as the ERA-40 reanalysis, can be used in climate studies to provide the so-called "perfect boundary conditions" for RCMs (e.g. Csima and Horányi, 2008; Déqué and Somot, 2008; Skalák et al., 2008; Heikkilä et al., 2011; Hamdi et al., 2012). These reanalyses are produced by means of data assimilation methods in order to find optimal estimates for past atmospheric states that are consistent with meteorological observations and the model dynamics.

In a recent study of Hamdi et al. (2012) the use of high-resolution dynamical downscaling of ALARO-0 at 4 km horizontal resolution is explored by means of summer maximum surface air temperature over Belgium. The analyses in this chapter extend the work of Hamdi et al. (2012) in the sense that, instead of temperature, precipitation is now analyzed. Daily summer precipitation from different model runs are compared with respect to station observations, with an emphasis on extreme precipitation. This approach by which model output is directly compared against station observations can be motivated by the fact that the station-level-observations provide the closest representation of extreme events (Dulière et al., 2011). Furthermore, the motivation for only considering summer precipitation is threefold: (i) other regional climate studies (e.g. Caldwell et al., 2009; Soares et al., 2012a,b) show difficulties of RCMs to simulate summer precipitation, (ii) the new parameterization scheme within ALARO-0 mostly modifies convection, which are the processes that are mostly relevant for (extreme) precipitation events in summer (Kyselý and Beranová, 2009; Soares et al., 2012a) and (iii) the relatively small scale on which these convective processes often occur better corresponds to the high-resolution ALARO-0 simulation (Kyselý and Beranová, 2009).

We add to our evaluation the ALADIN-Climate model developed by the Centre National de Recherches Météorologiques (CNRM), which took part in the European ENSEMBLES project (http://www.ensembles-eu.org/). The ALADIN-Climate model is an ALADIN model version which is specifically used for regional climate modeling. The ENSEMBLES project was finished in the end of 2009 and aimed to develop an ensemble climate forecast system to produce probabilistic scenarios of future climate in order to provide detailed, quantitative and policy-relevant information to the European society and economy

(http://www.ensembles-eu.org/). Several experiments were performed with some ten state-of-the-art European and Canadian high resolution, global, and regional climate models. The ENSEMBLES ALADIN-Climate/CNRM simulations use a long uninterrupted model run, which is a different set-up than our ALADIN- and ALARO-0 simulations. Hence, a direct comparison with the ALADIN-Climate/CNRM simulation is not possible, and these uninterrupted climate runs are merely added as a reference for regional climate modeling in order to make the present chapter complete.

The model simulations, experimental design and observational data used in this chapter are described in the next section. Section 3.3 gives a description of the applied methods and the results are discussed in the following section (Section 3.4). The results are summarized in the conclusions given in Section 3.5.

3.2   Experimental design and data

3.2.1   Experimental design

The experimental design is summarized in Table 3.1. The ERA-40 reanalysis data (Uppala et al., 2005), produced by ECMWF is dynamically downscaled using the limited area models ALADIN and ALARO-0. Both models are described in detail in the previous chapter (Chapter 2). The same land surface model ISBA (Interactions between Soil, Biosphere and Atmosphere; Noilhan and Planton, 1989) is used in both the ALARO-0 model as well as in the ALADIN model.

	Reference	Daily cumulated precipitation [day]	Model	Daily cumulated precipitation [day]
a. Effect of downscaling	Station observations	8 LT [day] → 8 LT [day+1]	ERA-40 ALR40 ALD40	6 UTC [day] → 6 UTC [day+1]
b. Multiscale performance of ALARO-0	Station observations	8 LT [day] → 8 LT [day+1]	ALR40 ALD40 ALR10 ALD10 ALR04	6 UTC [day] → 6 UTC [day+1]
c. Reference for regional climate modeling	Station observations	8 LT [day] → 8 LT [day+1]	CNRM	mean [0-24 UTC]

Table 3.1: *Overview of the experimental design.*

In a first part of this chapter the improvement of the downscaling by means of the ALADIN- and ALARO-0 model is examined. This is done by comparing recent past (1961-90) summer precipitation data from an ALARO-0 and ALADIN simulation performed at

40 km spatial resolution (ALR40 and ALD40, Fig. 3.1) with summer precipitation from the driving ERA-40 reanalysis data (Uppala et al., 2005).
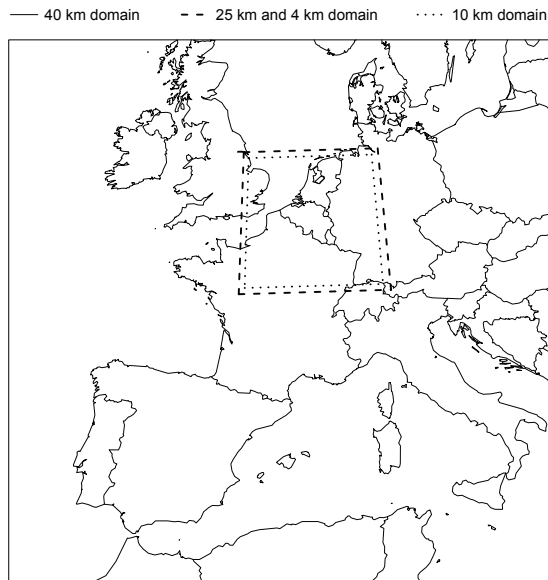
Figure 3.1: Domains corresponding to the different simulations at 40-, 25-, 10-, and 4-km horizontal resolution.

Despite the fact that reanalysis data products are more continuous in space and time than station data, they inevitably contain biases. A number of evaluations for ERA-40 re-analysis precipitation have been performed (e.g. Zolina et al., 2004; Ma et al., 2009). The ERA-40 precipitation has distinct regional limitations, most of them are generally related to the coarse horizontal resolution of the ERA-40 model, on one hand, and to its strong model dependency, on the other (Ma et al., 2009). All physical parameterizations within ERA-40, including those of precipitation, were run on a spatial resolution of about 125 km (Zolina et al., 2004; Ma et al., 2009). The model diagnostics precipitation in ERA-40 is produced by parameterized microphysical processes in clouds, which are formed at supersaturation by convective or large-scale processes (Ma et al., 2009). Total precipitation is then simply the sum of the convective precipitation generated by convective clouds, and large-scale strati-form precipitation, associated with the frontal or dynamical systems (Zolina et al., 2004). Hence, ERA-40 precipitation is a pure model product and due to the poor skill of operational Numerical Weather Prediction (NWP) models to account for all important physical mech-anisms that affect the atmospheric water cycle, it appears to be one of the most uncertain forecasted parameters in the reanalysis (Zolina et al., 2004; Ma et al., 2009; Heikkilä et al., 2011). The 6-hourly forecasts from the ERA-40 reanalysis are used to calculate daily cu-mulated summer precipitation between 6 UTC and 6 UTC of the next day. For the coupling to the regional model we use a linear interpolation in time. This may produce errors at the lateral-boundaries on our small domains (Fig. 3.1), but as shown by Termonia et al. (2009),

such errors only occur very rarely, and the impact on the statistics of extreme precipitation should be very minor.

To explore further the multiscale performance of ALARO-0, as found by Gerard et al. (2009), but now for climate timescales, we evaluate in a second step recent past simulations (1961-90) of the ALADIN- and ALARO-0 model at varying horizontal resolutions against different station datasets.

(i) and (ii) The ALADIN- and ALARO-0 model driven by ERA-40 and run at a horizontal resolution of 40-km spatial resolution with 69×69 grid points on a domain that encompasses most of Western Europe (respectively ALD40 and ALR40, Fig. 3.1).
These 40 km output are then used to perform a one-way nesting on a domain centered on Belgium (Fig. 3.1) on the following spatial resolutions:

(iii) and (iv) 10 km spatial resolution on a 67×67 grid (ALD10 and ALR10);

(v) and 4 km spatial resolution on a 181×181 grid (ALR04).
The fact that we did not run any ALD04 configuration is obviously linked to the corresponding gray-zone-type resolution, where the diagnostic parameterization of convection would have become completely irrelevant (see Section 3.4: Results and discussion for the first syndromes already noticeable in ALD10).

Finally, we also include ALADIN-Climate/CNRM simulations within our analysis in order to provide a reference for regional climate modeling. One part of the performed experiments within the ENSEMBLES project aimed to validate the models for the recent past climate. The results from this experiment, including 40 years of 25-km resolution ALADIN-Climate/CNRM simulations driven by the ERA-40 reanalysis (hereafter denoted as CNRM), are used in our analysis for the period 1961-90 (http://www.ensembles-eu.org/). From the ENSEMBLES data archive we have only selected the CNRM precipitation data for the grid points that coincide with the ALR04 domain (Fig. 3.1). The precipitation data corresponds to daily means calculated for the interval 0-24 UTC. As mentioned in the introduction (Section 3.1) the model set-up of CNRM and our simulations are different. The number of vertical levels that is used in our runs with the ALADIN- and ALARO-0 model is 46 with a model top that extends up to 72 km. The CNRM simulations from ENSEMBLES have used 31 vertical levels. Furthermore, the CNRM simulations use a long-term and free run set-up, while our procedure follows a dynamical downscaling with daily reinitializations over each subsequent 24-h period during the summer period of June, July, and August, 1961-90. More details on our approach for daily reinitializations are given in the previous chapter (Chapter 2, Section 2.3.2).

3.2.2   Observations

The observation dataset comprises 93 climatological stations with daily accumulated precipitation, selected from the climatological network of the RMI of Belgium. The data has undergone a manual quality control by operators, and the stations were chosen so that continuous data for the 30-yr study period (1961-90) are available. The stations cover whole Belgium, hence representing conditions of coastal, inland, and higher orographic locations (Fig. 5.2, right).

3.3 Methods

3.3.1 Data processing and analysis

Model validation against observations can either be done with station data or gridded station data. Both validation methods have their disadvantages (Hofstra et al., 2010). Model evaluation against observations at station level often raise issues related to the scale difference between the model- and observation field (Tustison et al., 2001; Dulière et al., 2011). The model grid cell values correspond to spatially averaged values representing the area of the whole grid cell. Furthermore, the spatial variability of these averaged model fields will always be lower than the one of the observation field. These differences in spatial variability depend on the area of the grid cell as well as on the inherent variability of the field variable. Precipitation, for example, is known to have a relatively high spatial variability. To illustrate the differences in spatial variability, Fig. 3.2 shows the different grid cell areas of the models together with the 93 climatological stations (i.e. observation points).
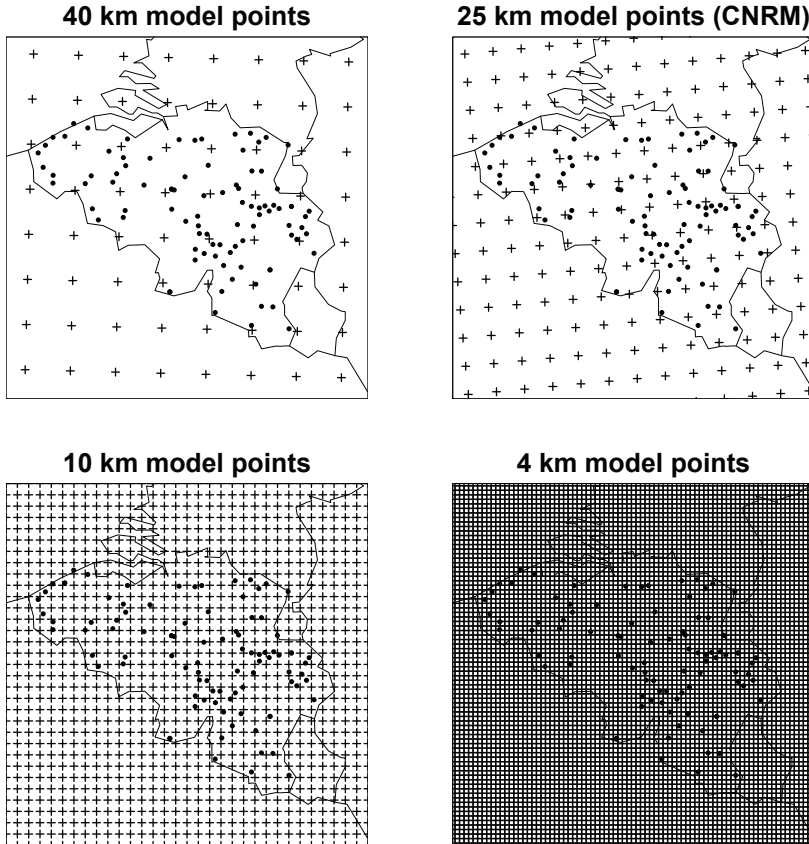
Figure 3.2: *Model grid points over Belgium for each of the horizontal resolutions for which the simulations are performed. The black dots represent the 93 climatological stations.*

The grid cell areas range here from 1600 km^2 for the 40 km horizontal resolution to 16 km^2 for the 4 km horizontal resolution (Fig. 3.2). Hence, reducing those spatially averaged model values with an originally greater heterogeneity to a single station point value, leads to an inconsistent comparison. However, for long time periods such as 30 years, we can assume that the spatial variability within a grid cell would be reduced, in such a way that the spatial variability of both model and observation field tend to converge (Dulière et al., 2011).

Another common way to overcome this scale inconsistency is the use of gridded data. The Climate Research Unit (CRU) and the European ENSEMBLES project provide daily gridded observation datasets (Mitchell and Jones, 2005; Haylock et al., 2008). However, these gridded datasets are in some regions constructed by interpolation or area-averaging of station observations from a small number of stations, which smooths and possibly affects the extreme values within the dataset (Hofstra et al., 2010). Since we aim to examine extreme precipitation events, the models are evaluated against station observations. This is done by comparison of daily observed station-level precipitation with modeled daily precipitation of the nearest grid box over land. The 93 resulting precipitation time series selected from the model simulations are not corrected for topography with respect to the altitude of the nearest station. It is difficult to apply such correction for precipitation, because of its dependency on topography, humidity, buoyancy, and other local variables (Soares et al., 2012a).

Time discrepancies between computations of daily cumulated precipitation from station observations on the one hand and model output on the other hand is an important, but rarely highlighted, problem within precipitation evaluation studies. In order to deal with this problem, the error analysis can be performed on longer than daily time scales, such as monthly, seasonal, or annual time scales (Ma et al., 2009; Soares et al., 2012b). However, in this chapter the model evaluation is done on a daily basis, requiring a consistent calculation of the daily precipitation values. Daily observed precipitation corresponds to the total accumulated precipitation between 8 and 8 Local Time (LT) of the day after. Hence, the daily model values for all simulations (ALR40, ALD40, ALR10, ALD10 and ALR04) have been calculated based upon the definition of observed daily accumulation which corresponds to 6 and 6 UTC of the day after (Table 3.1).

3.3.2   Extreme value analysis and Peak-Over-Threshold methods

The methods used for the modeling of extreme events are similar to those used in Hamdi et al. (2012). Threshold models and Peak-Over-Threshold (POT) methods are useful tools for the modeling of extreme events. A well known distribution which may describe the behaviour of the excesses or POT events is the Generalized Pareto Distribution (GPD) (Coles, 2001). Recently, several authors have modeled extreme precipitation with the GPD (e.g. Ribatet et al., 2009; Roth et al., 2012; Mailhot et al., 2013).

Consider a sequence of independent and identically distributed random variables $X_1, X_2, ..., X_i$ from an unknown distribution F. We are interested in the extreme events that exceed a certain high threshold u. The distribution function of such an extreme event X from the X_i sequence can then be defined as:

$$F_u(y) = P\{X > u + y | X > u\} = \frac{1 - F(u + y)}{1 - F(u)},$$

(3.1)

with $y > 0$. Eq. (3.1) is the conditional probability that the threshold u is exceeded by no more than an amount y, given that the threshold u is exceeded. Given that $X > u$, the GPD of the excesses $(X - u)$, is then given by:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi},$$ (3.2)

where ξ is the shape parameter and σ is the scale parameter. The GPD with parameters ξ and σ describes the limiting distribution for the distribution of excesses [Eq. (3.1)], and can be used to model the exceedances of a threshold u by a variable X. Thus, for $x > u$,

$$P\{X > x | X > u\} = \left[1 + \xi\left(\frac{x - u}{\sigma}\right)\right]^{-1/\xi}.$$ (3.3)

It follows that

$$F(x) = P\{X > x\} = \zeta_u \left[1 + \xi\left(\frac{x - u}{\sigma}\right)\right]^{-1/\xi},$$ (3.4)

where $\zeta_u = P\{X > u\}$. The parameters of the GPD are estimated by the maximum-likelihood method, following the definitions of Stephenson (2002). The level x_m that is on average exceeded once every m observations is the solution of

$$\zeta_u \left[1 + \xi\left(\frac{x_m - u}{\sigma}\right)\right]^{-1/\xi} = \frac{1}{m}.$$ (3.5)

The x_m return level, which gives the amount of extreme precipitation corresponding to a given number of observations m, is then given by

$$x_m = u + \frac{\sigma}{\xi} \left[(m\zeta_u)^\xi - 1\right].$$ (3.6)

3.4 Results and discussion

3.4.1 Effect of downscaling

As a first step we validate the effect of the downscaling of the ERA-40 reanalysis with the ALADIN- and ALARO-0 model. Figure 3.3 shows the relative frequencies calculated for daily precipitation amounts of ERA-40, ALR40 and ALD40 which are binned into bins of 1 mm day^{-1}. As a reference the relative frequencies of the observations are also shown. A logarithmic scale has been used for better representation of the extreme values. From both the ERA-40 data as well as the ALR40 and ALD40 data 93 grid points, corresponding to the closest grid points to the observation stations, have been selected. It should be noted that the ERA-40 reanalysis only have 2 grid points over Belgium. For low precipitation amounts (i.e. $< 0.95^{th}$ quantile of the observations) the ERA-40 reanalysis as well as ALR40 and ALD40 coincide well with the observations. However, for the higher rainfall rates ERA-40 starts to diverge from the observations, while ALR40 and ALD40 still approach the observations. Both 40-km models are able to reproduce rainfall rates up to 108 mm day^{-1}, while the reanalysis do not capture the higher precipitation amounts, which is related to the low spatial resolution of the ERA-40 data.
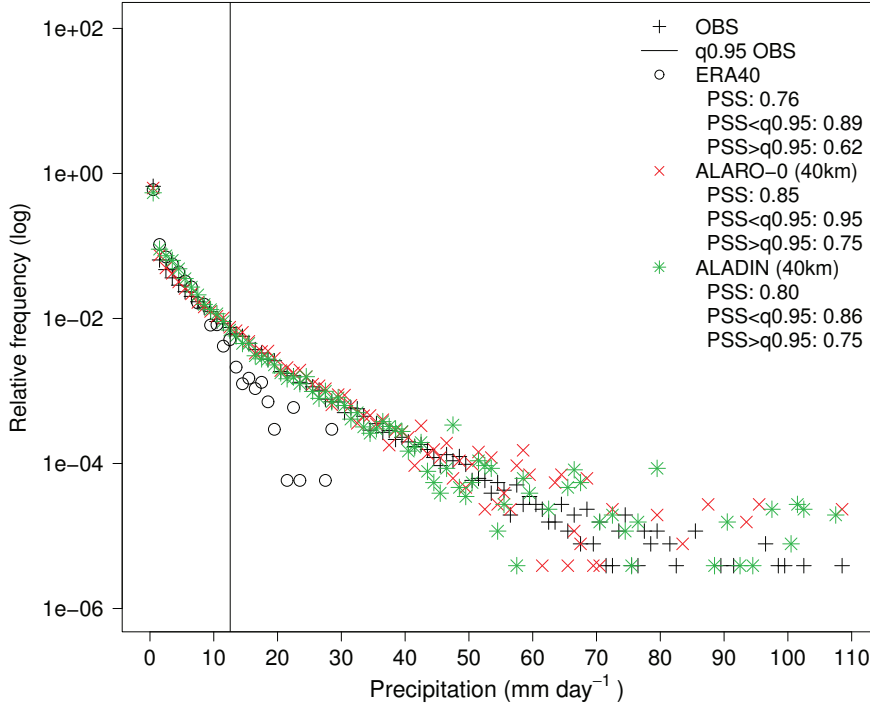
Figure 3.3: *Relative frequencies of observations, ERA-40, ALR40 and ALD40. Frequencies are computed with the 30-yr (1961-90) daily cumulated summer precipitation given for each station separately and are displayed on a logarithmic scale. Numbers for PSS correspond to the average of the Perkins Skill Score [Eq. (3.7)] calculated for precipitation amounts below and above the 0.95^{th} quantile of the observations (PSS < q0.95 and PSS > q0.95). The black line indicates the 0.95^{th} quantile of the observations.*

In order to provide a measure of similarity between observed and modeled frequencies the "Perkins Skill Score" (PSS) has been calculated (Perkins et al., 2007):

$$PSS = \sum_{1}^{n} minimum(Z_1, Z_2), \tag{3.7}$$

where n is the number of bins and $Z_{1,2}$ is the frequency of values in a given bin from the observation and model data, respectively. This metric measures how well the observations and modeled frequencies coincide, with a PSS ranging from zero for no overlap to a skill score of one for a perfect overlap. Similar to Boberg et al. (2010) and Domínguez et al. (2011), the PSS has been calculated for daily precipitation amounts going from 0 mm day^{-1} up to the 0.95^{th} quantile of the observations (PSS < q0.95) and for precipitation amounts above the 0.95^{th} quantile of the observations (PSS > q0.95). In this way the skill score is to a larger extent influenced by the more extreme precipitation values (Boberg et al., 2010). The skill scores are calculated for each station separately. The final PSS is then simply the mean value of the average of PSS < q0.95 and PSS > q0.95 over the 93 stations. The 0.95^{th} quantile of the observations, which is used as a threshold for the calculation

of the modified PSS, is also shown on Fig. 3.3. The Perkins Skill Scores for ERA-40 are relatively low, and for the higher precipitation amounts ERA-40 has a much lower PSS (PSS > q0.95: 0.62) than ALR40 and ALD40 (PSS > q0.95: 0.75). ALR40 and ALD40 perform very similar with respect to the observations and have relatively high PSS, which are close to one. To summarize, the downscaling with the ALARO-0- and ALADIN model is significantly different from the driving ERA-40 reanalysis and closer to the observations. In particular, ALR40 and ALD40 produce more extreme precipitation than their driving ERA-40 reanalysis.

3.4.2 Multiscale performance of ALARO-0

To investigate the multiscale performance of ALARO-0, 40-km, 10-km, and 4-km horizontal resolution simulations of ALARO-0 together with 40-km and 10-km horizontal resolution simulations of ALADIN are compared with respect to station observations.

3.4.2.1 Spatial and temporal distribution

Figure 3.4 shows the observed and simulated spatial distribution of the 30-yr averaged summer precipitation.
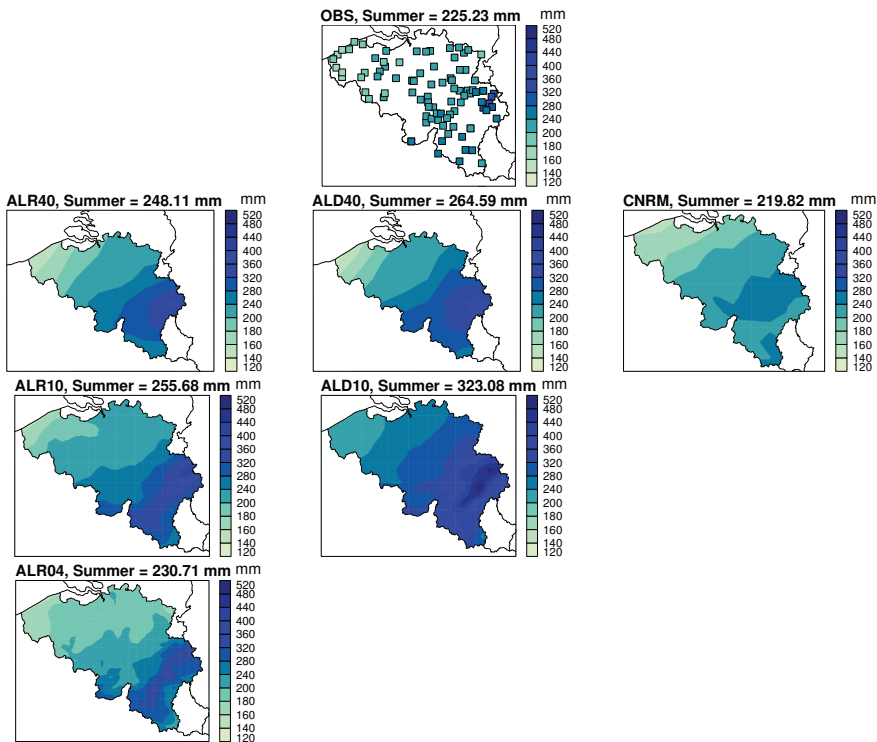
Figure 3.4: *Spatial distribution of 30-yr (1961-90) mean cumulated summer precipitation from observations and model simulations: (left) ALR40, ALR10, and ALR04; (middle) ALD40 and ALD10; (right) CNRM. The mean summer precipitation over the 93 climatological stations is given above each subfigure.*

On top of each sub-figure average values over the 93 stations for the cumulated summer precipitation are given. On average all models except for CNRM overpredict the observed cumulated summer precipitation. Both the observation- and the simulation fields show a clear topographical dependency, with a gradual increase in precipitation going from the northwest (low altitudes) to the southeast (high altitudes) of the country. The ALARO-0 and ALADIN simulation at 40 km show a very similar distribution. Obviously, the precipitation fields for the simulations with low spatial resolution are less heterogeneous than the ones with high spatial resolution. However, the 25-km spatial resolution CNRM plot illustrates less variability than the 40-km simulations: also, the local maximum in the southeast cannot been seen on the CNRM plot. For the higher resolution simulations ALARO-0 approaches much better the observations than ALADIN. For instance, ALD10 overpredicts cumulated summer precipitation with values that are, on average, over all stations almost 100 mm higher than observed. On the contrary, the average values for ALR10 and ALR04 differ only slightly from the observations, and the observed local maximum at the higher altitudes is very well simulated by both models.

The scatterplots presented in Fig. 3.5 are consistent with the spatial distributions shown in Fig. 3.4.
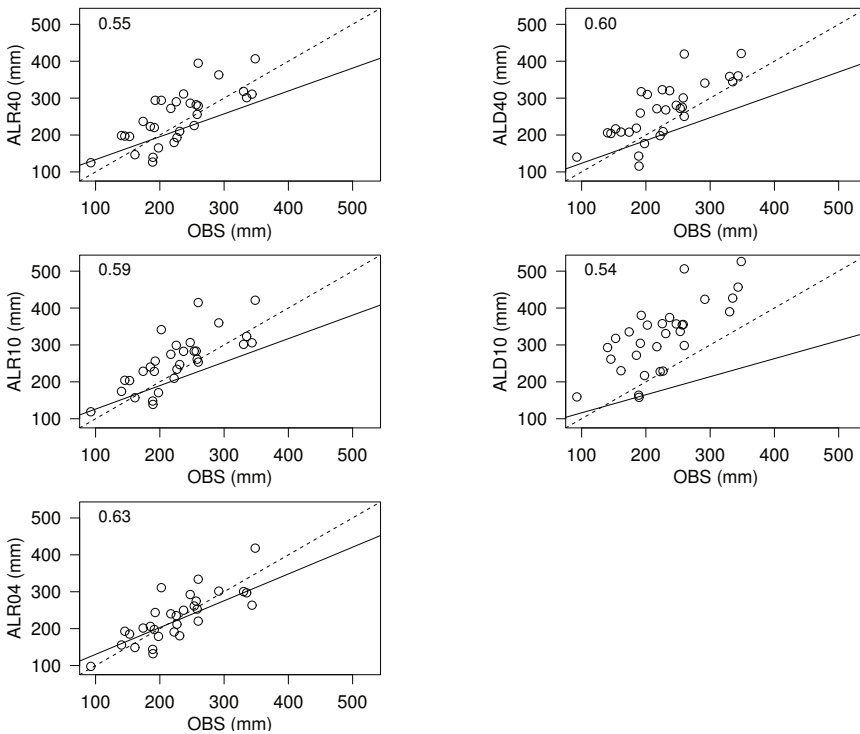
Figure 3.5: *Each point in the scatterplots represents summer cumulated precipitation for each year in the 30-yr period (1961-90) averaged for the 93 stations. The dotted (solid) black line is the diagonal (linear regression) line. The number in each scatterplot corresponds to the determination coefficient (R^2) of the linear regression.*

Each point in the scatterplots represents the summer cumulated precipitation for each year in the 30-yr period averaged for the 93 stations. Linear regression lines (solid line) and its determination coefficients (R^2) is also presented for each of the five models. Except for ALD10, summer precipitation is relatively well simulated by all models. The ALD10 model shows again a clear overestimation of observed summer precipitation. This is an indirect confirmation that, with 10-km mesh sizes, the syndromes linked to the gray-zone performance are already present.

3.4.2.2 Error statistics

The previous analysis showed the ability of the models to represent the spatial and temporal pattern of mean annual summer precipitation. To quantify this ability we have computed some important error statistics. Figure 3.6 shows the spatial distribution of the 30-yr average summer biases of the daily cumulated precipitation, as well as the mean bias over the 93 climatological stations. Average values over the 93 stations of other 30-yr mean summer statistics are also given: the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). The statistics are calculated with daily values for each station separately. Both 40-km simulations ALR40 and ALD40 again perform similar.
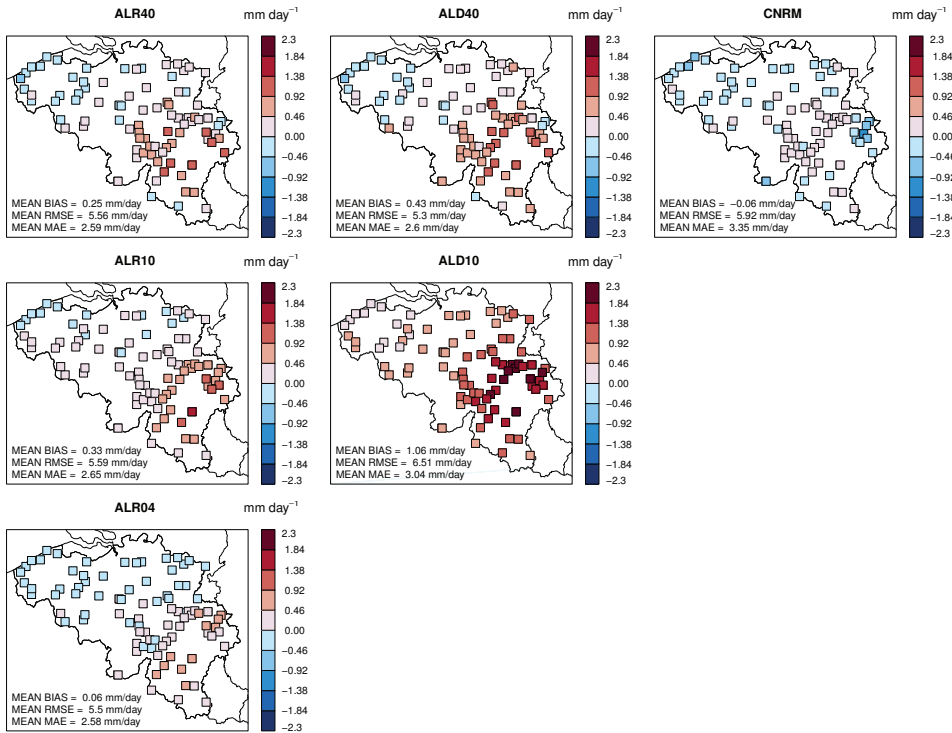
Figure 3.6: *Spatial distribution of the 30-yr (1961-90) average summer biases (model minus observed) of the daily cumulated precipitation. The numbers correspond to the spatial mean of the bias, the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE).*

Overall, the biases are remarkably lower for ALARO-0 than for ALADIN. The bias

over the 93 climatological stations between model simulations and observations is 0.25 mm day^{-1} for ALR40, 0.43 mm day^{-1} for ALD40, -0.06 mm day^{-1} for CNRM, 0.33 mm day^{-1} for ALR10, 1.06 mm day^{-1} for ALD10, and 0.06 mm day^{-1} for ALR04. The error statistics for all three ALARO-0 simulations show a similar improvement, suggesting a multiscale performance of ALARO-0. However, one should also keep in mind that error statistics are not entirely fair when validating models with different spatial resolution. Small displacements of precipitation maxima and minima in higher resolution models are highly penalized by error statistics, because of the so-called double penalty effect (Soares et al., 2012a).

The aforementioned underestimation by CNRM is confirmed by the spatial distribution of its bias. Furthermore, the coastal precipitation is by all other models generally better simulated than the inland precipitation at the higher elevations (Fig. 3.6). The larger and positive differences at the higher elevations can partly be assigned to higher uncertainties in the measurements of the observations due to rain gauge undercatchment (Buonomo et al., 2007). However, this overestimation can also be attributed to the driving ERA-40 data or to the model formulation. The positive biases are indeed more strongly pronounced for ALD10 which uses the old diagnostic parameterization scheme (Fig. 3.6). All three ALARO-0 simulations (40-, 10-, and 4-km horizontal resolution) produce the lowest deviations from the observations, with a tendency to slightly overestimate (underestimate) in the southern (northern) part of the country. ALARO-0 values for RMSE and MAE lie in the same range as those for ALADIN, indicating that the low mean biases of ALARO-0 are possible due to cancellation effects arising from the bias computation. Nevertheless, the overall errors of the ALARO-0 simulations are still smaller than those of ALD10.

In order to get an understanding of the trend of frequency and intensity of extreme precipitation, density curves and frequency and quantile distributions of all six simulations have been created (Figs. 3.7 to 3.9). The densities in Fig. 3.7 have been calculated with the square root of the daily precipitation since the majority of the precipitation rates are less than 10 mm day^{-1}. All models tend to overestimate the amount of "drizzle" and low precipitation (i.e. < 1 mm day^{-1}). In the 1-2 mm day^{-1} range, both ALADIN simulations as well as CNRM overestimate the observed density almost by 2 times, while ALARO-0 starts to approach closely the observed density (Fig. 3.7, middle). The latter continues to do this up to the right end tail of the observed density curve (Fig. 3.7, right). Perkins et al. (2007) use Probability Density Functions (PDFs) for the evaluation of simulated daily precipitation over Australia from 14 different climate models. Similarly to the density curves of ALADIN and CNRM, the PDFs in Perkins et al. (2007) show for all models an overestimation of "drizzle", with most models overestimating the observed density of rainfall in the 1-2 mm day^{-1} range by 2-3 times.

The relative frequencies, shown in Fig. 3.8, are again calculated for daily precipitation amounts of the observations and model data, which are binned into bins of 1 mm day^{-1}. For the low precipitation rates all models manage to reproduce the observed frequencies relatively well. Once the 0.95^{th} quantile of the observations (indicated by the vertical black line) is exceeded, CNRM shows an increasing departure from the observations with frequencies left shifted from the observations. ALARO-0 and ALADIN at 40-km horizontal resolution reveal again a similar result, while for the higher 10-km resolution a clear difference between both models is apparent. The small overestimation of ALD10 for the low

precipitation rates persists and becomes larger for the higher rates. The model clearly rains too often, both with very small and very high quantities of rainfall. On the other hand, the frequencies of ALR04 and ALR10 nicely follow the observations, showing their ability to capture the occurrence of extreme and rare precipitation events, with values around 100 mm, quite well. As a measure for similarity between the observed and modeled frequencies the PSS [Eq. (3.7)] are also given in Fig. 3.8. The overall PSS, as well as PSS for precipitation amounts below and above the 0.95^{th} quantile of the observations, is higher for ALARO-0 than for ALADIN and CNRM.
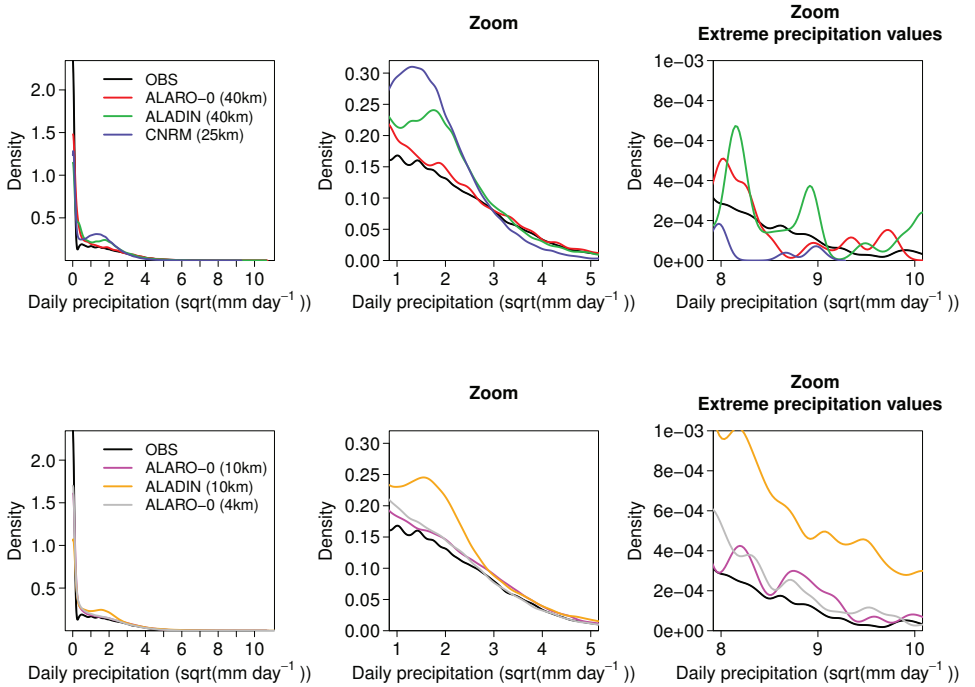
Figure 3.7: *Density curves of (top) observations, ALR40, ALD40, and CNRM and (bottom) observations, ALR10, ALD10, and ALR04. Densities are computed with the 30-yr (1961-90) daily cumulated summer precipitation given for each station separately. The x axes represent the square root of the daily precipitation since the majority of the precipitation rates are less than 10 mm day^{-1}.*

The quantile distributions confirm the ability of ALR04, ALR10, and even ALR40, to reproduce extreme rainfall rates (Fig. 3.9). Only the highest 99.9 quantile (i.e. strongest events) is slightly overestimated by ALARO-0. It is evident that such events, which are situated in the very end of the distribution, might correspond to outliers. Consistently with the frequency plots, the higher quantiles are over- and underestimated by ALD10 and CNRM, respectively.
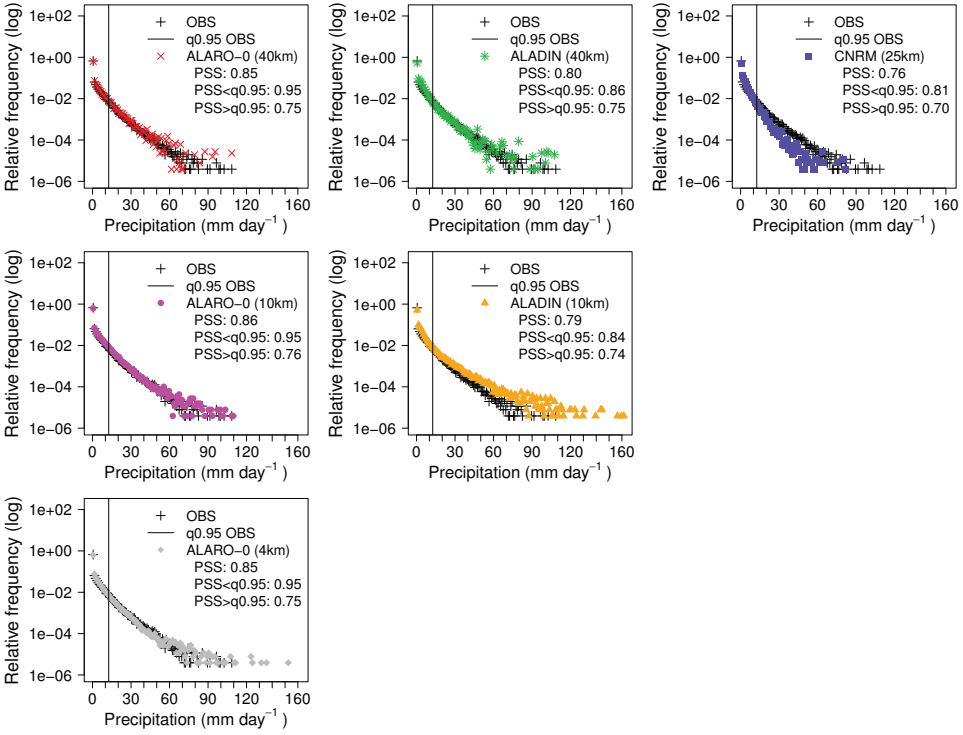
Figure 3.8: *As in Fig. 3.3, but for observations and model simulations: (left) ALR40, ALR10, and ALR04; (center) ALD40 and ALD10; (right) CNRM.*
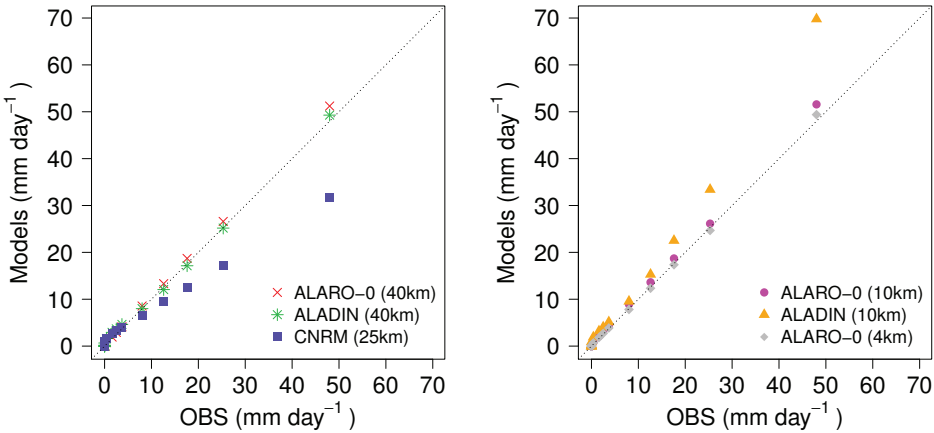
Figure 3.9: *Quantiles (2.5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95, 97.5, 99, 99.9) of observations vs. (left) the ALR40, ALD40, and CNRM models and (right) the ALR10, ALD10, and ALR04 models. Quantiles are computed with the 30-yr (1961-90) daily cumulated summer precipitation given for each station separately.*

Previous results can be qualified in the context of other regional downscaling studies; however, a direct comparison is difficult because of differences in study area and model design. Soares et al. (2012a) performed a dynamical downscaling of 20 years of ERA-Interim reanalysis (1989-2008) for Portugal using the Weather Research and Forecast model (WRF). Two WRF high-resolution simulations (9 and 27 km) and ERA-Interim reanalysis are compared with station observations. For summer precipitation, their results show a different frequency distribution for the 9-km and 27-km simulation. The 9-km frequencies of summer precipitation follow well the observed frequencies and show a clear improvement compared to the driving reanalysis. In a more recent study, Chan et al. (2013) evaluated for southern United Kingdom 17 years (1991-2007) of ERA-Interim driven RCM runs at 50 km, 12 km and 1.5 km resolution. For the 50-km and 12-km simulations the non-hydrostatic Hadley Centre Global Environmental Model version 3 (HadGEM3-RA) was used. The highest resolution runs at 1.5 km were one-way nested into the 12 km simulation, and used another model version, with no convection parameterization, nor a prognostic cloud scheme. Comparison of moderate and extreme daily precipitation values from the three model runs with observations, showed a clear improvement in both the winter and summer bias going from 50 km to 12 km or 1.5 km. However, at the daily level their results showed no clear evidence that the 1.5-km simulation performs better than the 12-km simulation, or vice versa. For instance, at 1.5 km the model simulates too many days with extreme JJA precipitation. The authors attribute these biases to the fact that the 1.5-km simulation is "convection permitting", so that even at 1.5 km, convection is still under-resolved, and the explicitly-resolved convective motion and vertical mass flux become too intense.

Our results show a coherent performance of the ALARO-0 model across all resolutions and the good model performances as displayed in Figs. 3.7 to 3.9 can be practically attributed to the quality of the physics parameterizations unrelated to the increase of the resolution. Finally, the persistent positive biases of the ALADIN model ALD10 are in accordance with other studies where recent past (1961-90) ALADIN simulations at 10-km horizontal resolution, driven by ERA-40 data, are validated against gridded observations (see Csima and Horányi, 2008; Skalák et al., 2008). According to Skalák et al. (2008), these positive (summer) precipitation biases can be linked with the tendency of the model "to precipitate" more often than in the station observations.

3.4.2.3   Extreme Value Analysis

The Extreme Value Analysis (EVA) has been performed for each station separately, using the 30-yr daily summer data. The use of a GPD as a model for threshold excesses assumes independent excesses (Coles, 2001). In practice this is rarely the case. Exceedances over a certain threshold often occur in clusters. In order to account for these clusters of POT events, the data have been declustered by selecting the maximum value within each cluster. The independence of two clusters of POT events is determined by a combination of the threshold and the separation time between both clusters. However, the choice of a suitable threshold and separation time is relatively arbitrary. The threshold has to be high enough in order to ensure *extreme* events and to avoid dependency between the events, but a threshold which is too high prevents statistical significance due to a loss of information (Kyselý and Beranová, 2009; Heikkilä et al., 2011). Similar to the study of Heikkilä et al. (2011), the

threshold has been defined for each station separately as the 0.95^{th} quantile of daily summer precipitation, so that spatial differences in the precipitation amount (see Fig. 3.4) are taken into account.
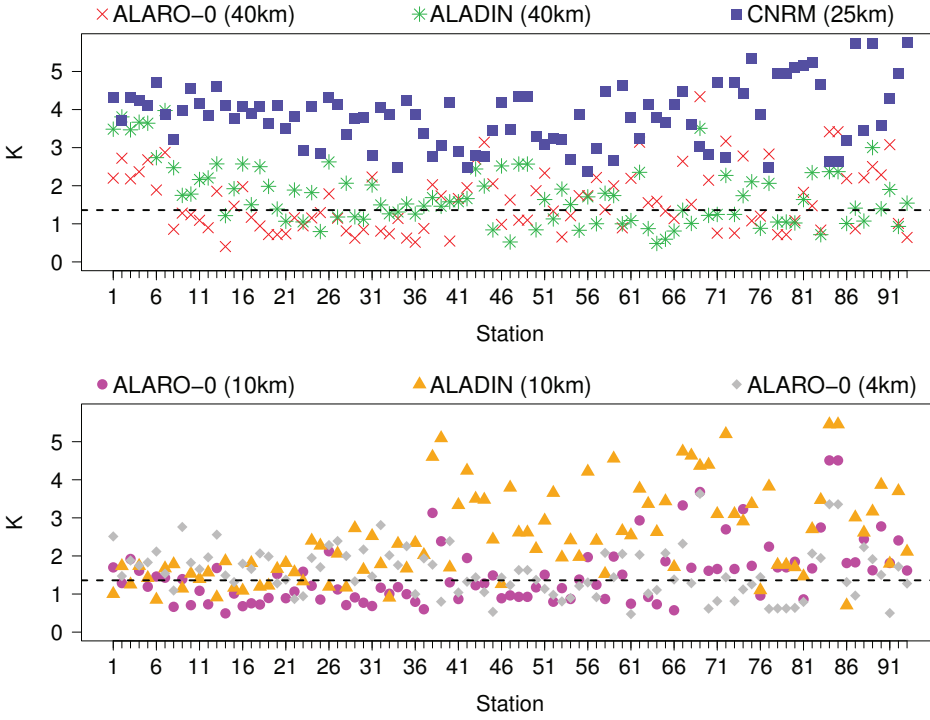
Figure 3.10: *K statistic from a Kolmogorov-Smirnov test [Eq. (3.9)]. The 93 stations (abscissa) are shown by ascending altitude (from left to right). The test is performed on the POT events of the observations vs. the (top) ALR40, ALD40, and CNRM and (bottom) ALR10, ALD10, and ALR04 model simulations. The horizontal dotted line represents the critical K level with significance $\alpha = 0.05$.*

The results obtained by using cluster maxima defined with different separation times (e.g. 1, 2, or 4 days) do not differ much from the results when the original non-declustered data have been used (not shown). Hence, in accordance with another study on extreme precipitation of Kyselý and Beranová (2009), two POT events are considered to be independent when the minimum separation time between both events is one day.

To investigate if the underlying probability distribution of the (declustered) POT events of the observations and models significantly differs, a Kolmogorov-Smirnov (K-S) test has been applied. The K-S test statistic is defined as the maximum absolute difference between to distribution functions:

$$D_{n1,n2} = Max|F_{n1}(x) - F_{n2}(x)|, \tag{3.8}$$

where $F_{n1}(x)$ and $F_{n2}(x)$ are the empirical distribution functions of the observations and the model, respectively, and n_i refers to the number of samples. The null hypothesis (H_0) that the distribution of the observed POT events equals the distribution of the modeled POT

events, is rejected at significance level $\alpha = 0.05$ if

$$K = \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} D_{n1,n2} > K_\alpha, \tag{3.9}$$

where K_α is the critical α-level of the Kolmogorov distribution:

$$Pr(K \le K_\alpha) = 1 - \alpha. \tag{3.10}$$

Figure 3.10 shows for each station the K statistic of the observations and models. In general, the K values for the ALARO-0 model at all three spatial resolutions are much smaller than for ALADIN and CNRM. H_0 is accepted at the 95%-level at 35 and 16 stations for ALD40 and ALD10, respectively. For ALARO-0 at 40, 10, and 4 km, H_0 is accepted at 46, 47, and 38 locations, respectively. Compared to ALD10, there are for ALARO-0 more stations at the high altitudes for which the distribution of the POT events equals the observed distribution of the POT events. This indicates that an increase in resolution does not necessarily contribute to a better representation of orographic precipitations. In the case of CNRM, H_0 is rejected for all stations. Thus, consistent with the results from the frequency and quantile distributions, the K-S test confirms that the ALARO-0 simulations yield more reliable statistics of the extreme events.
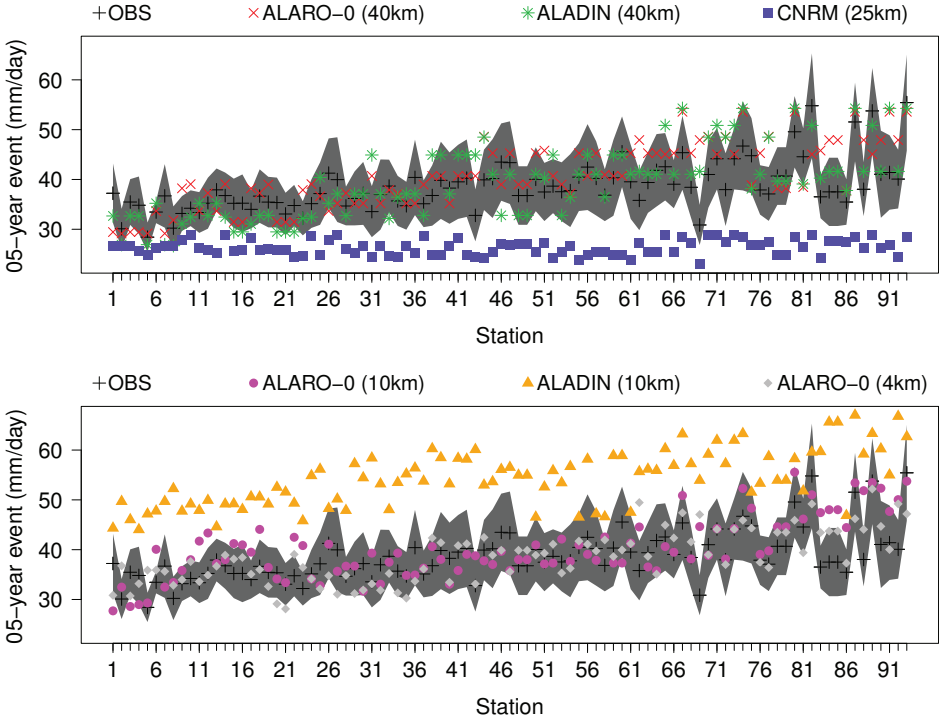
Figure 3.11: *The 5-yr return levels of the POT models for the observations and model simulations: (top) ALR40, ALD40, and CNRM and (bottom) ALR10, ALD10, and ALR04. The 93 stations (abscissa) are shown by ascending altitude (from left to right), and the shaded area represents the 95% confidence interval of observed return levels.*

The GPD equation [Eq. (3.2)] is then fitted through the selected cluster maxima of the observations and the six model simulations ALR40, ALD40, CNRM, ALR10, ALD10, and ALR04. The 5- and 20-yr return levels of the POT models for the observations and six simulations are shown in Figs. 3.11 to 3.12. The return levels x_m are calculated by Eq. (3.6) using the declustered data with 1 day separation time and a threshold u, defined as the 0.95^{th} quantile. Since the return levels x_m are calculated on an annual basis, the value for m equals 92 observations, corresponding to the number of summer days within one year of the study period. The return levels are for both return periods generally larger at the higher elevations. The 95% confidence levels of the observed return levels are also indicated. It appears that for most stations the return levels of ALARO-0 lie within the 95% confidence range of the observed return levels. In contrast to ALARO-0, ALD10 and CNRM are not able to produce the observed 5- and 20-yr return events. Their estimated return levels lie for a great number of stations outside the observed confidence interval.

In line with what Hamdi et al. (2012) found for summer maximum temperature, previous results from the EVA show for ALARO-0 at the high resolutions of 4 km and 10 km, as well as at 40-km horizontal resolution, a clear improvement in simulating extreme summer precipitation. Extreme events are also often investigated by means of climate indices (e.g. Herrera et al., 2010; Domínguez et al., 2011; Dulière et al., 2011; Soares et al., 2012b). In order to complete the EVA, two main precipitation indices have been calculated: the number of wet days and the number of very heavy precipitation days. Both indices are explained below and are calculated for each year (i.e. summer season) and each climatological station.
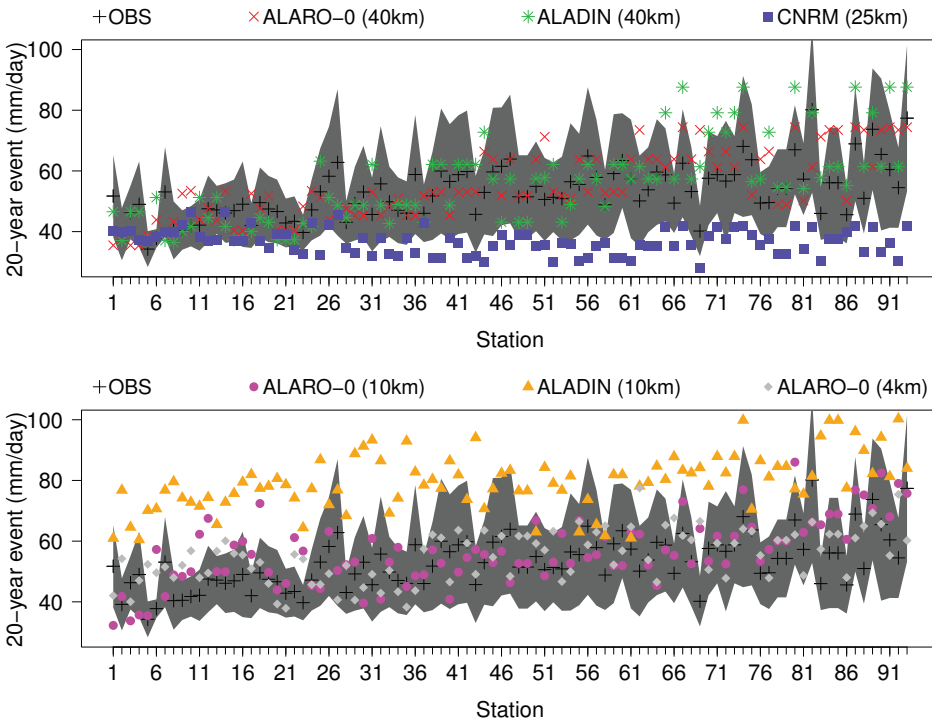
Figure 3.12: As in Fig. 3.11, but for 20-yr return levels.

3.4.2.4   Number of wet days (WD)

The number of wet days (WD) for the observations and models are defined as the annual count of days when precipitation is >1 mm. Figure 3.13 shows the ratio of WD in models to observations.
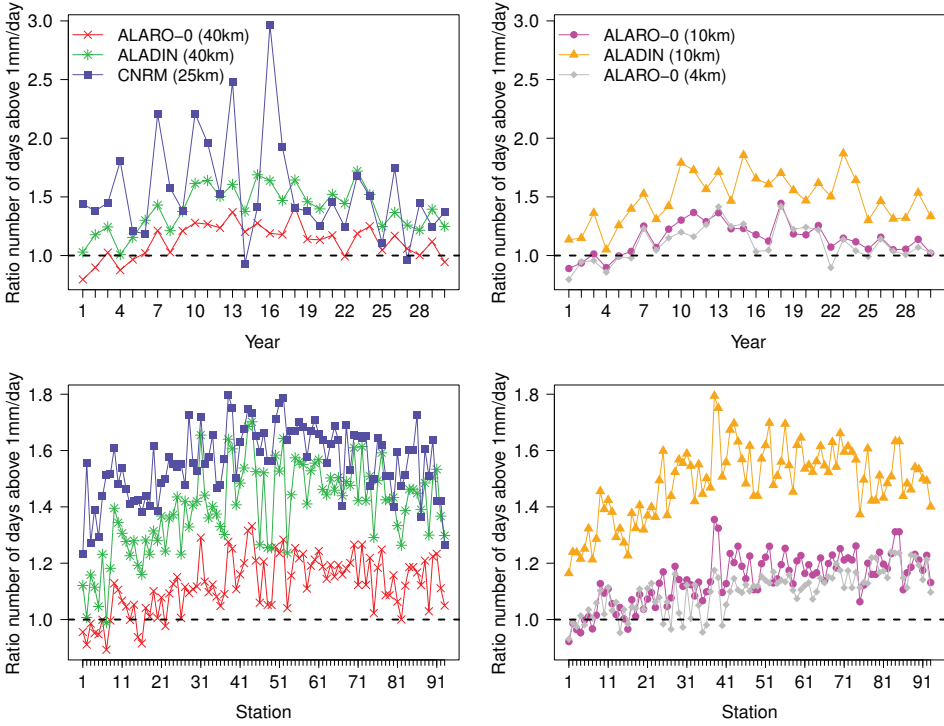
Figure 3.13: *(top) Spatial mean of ratio of number of days above 1 mm day^{-1} (i.e. WD) in models to observations. (bottom) Temporal mean of ratio of number of days above 1 mm day^{-1} (i.e. WD) in models to observations. Station sequence as in Fig. 3.11.*

As the model values represent a whole grid box, we could assume that the models, and especially the lower resolution models, will poorly reproduce the indices at the station points. However, the low resolution ALR40 model (left) reproduces relatively well the observed WD. On the other hand, ALADIN and CNRM show an overestimation for WD. This can be explained by the fact that precipitation may occur more systematically at the model grid box level, which gives rise to a WD even when no precipitation has been observed at the station location. Compared to ALADIN and CNRM, the ALARO-0 model (at 4-, 10-, and 40-km horizontal resolution) is able to better reproduce the number of wet days.

3.4.2.5   Number of very heavy precipitation days

The number of very heavy precipitation days is derived by annual counting of days with precipitation rates >20 mm. The temporal as well as the spatial mean of the number of very heavy precipitation days are consistent with the results from foregoing EVA. Overall,

ALR04, ALR10, and ALR40 can reproduce the number of days with precipitation >20 mm day^{-1} very well (Fig. 3.14). ALR04 and ALR10 have the highest correlations, and for three out of the 93 stations ALR10 predicts exactly the same number of days with heavy precipitation rates as have been observed.
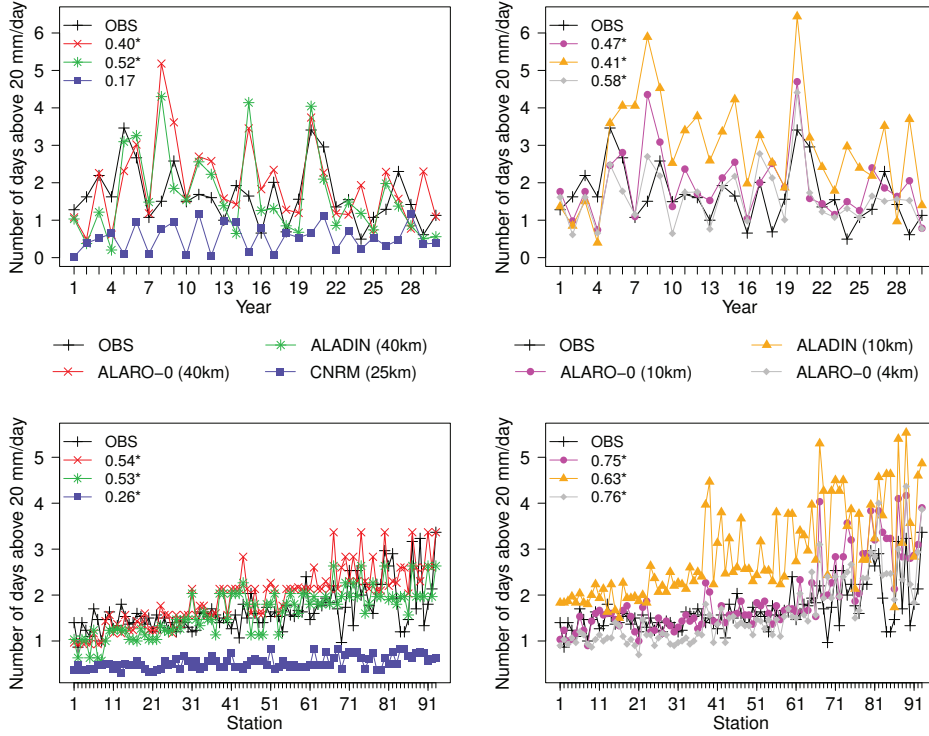
Figure 3.14: *(top) Spatial mean of number of days above 20 mm day^{-1}. (bottom) Temporal mean of number of days above 20 mm day^{-1}. The numbers correspond to Pearson correlation coefficients. Significant correlation coefficients at significance level of 0.05 based on t statistics are indicated with an asterisk. Station sequence as in Fig. 3.11.*

3.5   Conclusion

Extreme value analysis, using the Peak-Over-Threshold method and the GPD, was performed in order to explore the relative importance of resolution versus parameterization formulation on the simulation of extreme daily summer precipitation. The results show that dynamical downscaling of the ERA-40 reanalysis using the ALARO-0 model adds value to the prediction of extreme daily summer precipitation when compared to the ERA-40 results. Hence, running a LAM with the adapted parameterization, which was originally motivated to perform in the convection-permitting resolutions, statistically outperforms the global data in the output of extreme precipitation events of the ERA-40 reanalysis. The main strength of these tests is that, by the choice of the setup, we are considering the pure effect of the downscaling, without being obliterated by issues such as spectral nudging. Moreover, the model regenerates the precipitation instead of letting it evolve from its initial state. The re-

gional nature keeps the computing cost within reach of a typical small center, like the RMI, while reproducing the correct statistics of the extreme precipitation events consistently with the large-scale forcing imposed by the initial conditions and the lateral boundaries. Furthermore, it should be stressed that the present model version has been developed and tuned in a context of NWP, is used as a 12-member component of the Grand Limited Area Model Ensemble Prediction System (GLAMEPS), and has been taken as such to downscale ERA-40 data. This can be seen as an extra indirect validation of the NWP applications running ALARO-0, in the sense that the model has a more correct climatology of convective rain. It is clear that there are several components, such as the physics-dynamics interaction, the interaction between model physics, and the numerics, that may influence the climatology of the precipitation. However, it is difficult to isolate the importance of these components, and it is beyond the scope of this thesis to address the relative impact of the different parameterization updates within ALARO-0, it should be kept in mind, though, that all of these factors play a crucial role in the model performance at gray-zone resolutions.

ALARO-0 simulations at 40-, 10-, and 4-km horizontal resolution with a new parameterization scheme of deep convection and microphysics, as well as 40- and 10-km horizontal resolution output from the ALADIN model, with an old parameterization scheme were compared with respect to station observation data. We find for ALARO-0 at high spatial resolutions of 10 and 4 km an improvement in the spatial distribution of summer precipitation, such that the distinct local maximum at the highest elevations is well resolved by the model, a feature which is strongly overestimated by the ALADIN model at 10-km resolution. Furthermore, the results from the EVA suggest that the new parameterization scheme of ALARO-0 contributes to the improvement in the modeling of extreme precipitation events at varying horizontal resolutions, rather than the increase in spatial resolution. Thus, the nature of the parameterization is more important than the resolution, which confirms previous findings of Lynn et al. (2010) and Hamdi et al. (2012).

In the next chapter (Chapter 4), the ability of the ALARO-0 model to reproduce characteristics of subdaily precipitation at multiple resolutions, will be assessed.

4

ASSESSMENT OF SUBDAILY PRECIPITATION FROM THE ALARO-0 MODEL AT DIFFERENT SPATIAL RESOLUTIONS

"Current Regional Climate Models (RCMs) show skill in capturing the statistics of the daily precipitation distribution but do not well represent subdaily precipitation and the diurnal cycle of convection."

– Maraun et al. (2010)

4.1 Introduction

In the previous chapter (Chapter 3) it was investigated how the ALARO-0 model performed in simulating daily extreme summer precipitation for Belgium. However, for impact studies of extreme precipitation events, decision makers often require current and projected future climate information at the local scale and at higher temporal resolutions than the daily scale (Maraun et al., 2010; Mahoney et al., 2013). Over the last decade considerable efforts were made to further develop and improve Regional Climate Models (RCMs) by increasing their complexity and resolution. The state-of-the art RCMs typically have horizontal grid spacings ranging from 50 to approximately 10 km (Prein et al., 2013a, 2015). Nevertheless, there still remains a discrepancy between the spatial scale of RCM precipitation and the site-specific data needed for many impact studies (Maraun et al., 2010). In addition to this scale gap, numerous processes at these resolutions still cannot be resolved on the model grid and therefore must be parameterized. These parameterizations, and in particular the deep convective parameterizations, are a key source of model errors in the simulation of precipitation (Prein et al., 2015) . The parameterization of deep convection is a challenging task for climate modelers because the triggering emerges from an interplay of processes acting at scales from the microscale to the synoptic scale (Prein et al., 2015). Furthermore, the simulation of precipitation in RCMs is also highly sensitive to other aspects of the model formulation, such as the horizontal resolution, the numerical scheme, and other physical parameterizations (Maraun et al., 2010).

Several studies have demonstrated that the deep convection parameterizations are responsible for common errors in the models. At the daily scale, RCMs underestimate the number of dry days (e.g. Fowler et al., 2007), overestimate the low-precipitation event frequency (i.e. so-called "drizzle effect"), and at the subdaily scale the models show a poor representation of convective processes such as the diurnal cycle (e.g. Dai et al., 1999; Bechtold et al., 2004; Brockhaus et al., 2008; Hohenegger et al., 2008) and an underestimation of hourly precipitation intensities and high intensity events (Maraun et al., 2010; Kendon et al., 2012).

For decades, convection-permitting models (CPMs), where most of the error-prone convection parameterizations are switched off, have been widely used for Numerical Weather Prediction (NWP) and idealized case studies (e.g. Mass et al., 2002; Guichard et al., 2004; Hohenegger et al., 2008; Gerard et al., 2009). The kilometer scale at which these CPMs operate allows to explicitly resolve deep convection as well as an improved representation of fine-scale orography, variations of surface fields and boundary layer processes, which are crucial for the initiation of convection in complex terrain (Ban et al., 2014). This has indeed led to significant improvements of quantitative precipitation forecasts. Recently, with the advances in high-performance computing, CPMs are also used

in a climate context (e.g. Kendon et al., 2012; Prein et al., 2013b; Ban et al., 2014; Chan et al., 2014; Fosser et al., 2015). However, the use of such convection-permitting resolutions in long-term climate studies is currently still limited (Kendon et al., 2014). In particular because increased resolution implies additional computational costs and storage (Schwartz et al., 2009; Prein et al., 2015). Most studies are generally limited to small domains, and often run for a single season or shorter multiyear simulations (Kendon et al., 2014). This raises the question which horizontal grid spacing is sufficient for such CPM climate simulations.

In a NWP context, several studies made an attempt to address this issue, i.e. to find the optimal grid spacing of CPM forecasts that maximizes the forecast quality, value and realism (Schwartz et al., 2009). Weisman et al. (1997) investigated the upper limit on the horizontal grid spacing of convection-permitting simulations using idealized squall line simulations. Their findings suggest that a resolution of 4 km is sufficient to reproduce most system-scale aspects of squall-line-type convective systems over a 6-h period (Weisman et al., 1997). Furthermore, Schwartz et al. (2009) assessed the impact of horizontal resolution by comparing output from a single deterministic 2-km model with forecasts from a 4-km resolution ensemble control member. Except for the difference in horizontal resolution, the configurations of both sets of forecasts were identical and no parameterization for deep convection was used. In order to provide a reference for the high-resolution output, forecasts from a 12-km NWP run that uses a convection parameterization, were also considered. The authors found for the convection-permitting, high-resolution models significant added value for next-day forecasts compared to the 12-km forecasts. It is suggested that the improved model behavior from the 4-km resolution is nearly as advantageous as the 2-km grid spacing. Hence, as concluded by the authors, the added value of further increasing the spatial resolution is much smaller compared to the benefits obtained when NWP models with convective parameterization and at $\approx$ 10-km grid spacing are reconfigured to $\approx$ 4 km resolution using an explicit formulation of convection. However, some operational centers still apply parameterization schemes at 4 km, though in modified forms, because of a concern that unrealistic forecasts will be produced if no convective parameterization is used (Schwartz et al., 2009). A study of Roberts and Lean (2008), for example, assessed the performance of 12-, 4-, and 1-km versions of a Met Office's NWP model with the aim to examine the improvement to forecast skill from increased resolution alone. Only at the 1 km resolution the model was run without any convective parameterization scheme. Their results showed that the 1-km model is more skillful than the lower resolution models, specifically giving a more accurate distribution of the rain and a better prediction of high accumulations. As pointed out by the authors, the 4-km model does not achieve the same level of performance because of inherent difficulties in representing convection at that resolution (Roberts and Lean, 2008). Given this issue of 4-km suitability for convection permitting forecasts, Prein et al. (2015) hence suggest that it seems prudent to use horizontal grid spacings of less than 4 km for CPM climate simulations.

Although there is only a small number of CPM climate studies using different models, they have demonstrated clear improvements in the issues related to the deep convective parameterizations of RCMs. Kendon et al. (2012) and Chan et al. (2014) compared 20-yr ERA-Interim driven simulations over southern United Kingdom from a RCM at 12 km resolution with a parameterized convection and a CPM at 1.5 km resolution where convection

is explicitly solved. Results from both studies showed that rainfall is much more realistic in the 1.5 km simulation, with a much better representation of subdaily extreme events during summer and a clear improvement in its spatial and temporal structures and the diurnal cycle. Similarly, Fosser et al. (2015) conducted two high-resolution simulations at 7 and 2.8 km resolution over southwest Germany for 30 years. The authors examined with respect to observations the differences between both resolutions in the representation of precipitation at subdaily timescales and found for the highest resolution a significant improvement in the representation of hourly intensities and diurnal cycle of precipitation. Furthermore, several studies revealed no general improvement in daily mean precipitation between CPMs and RCMs (e.g. Li et al., 2012; Chan et al., 2013; Prein et al., 2013a; Chan et al., 2014; Ban et al., 2014). Hence, the added value of CPMs can mostly be found: (i) where/when deep convection is a dominant process (e.g. midlatitude summer), (ii) on small spatial and temporal scales (e.g. subdaily precipitation), (iii) in regions with strong spatial heterogeneities (e.g. coastal and urban areas), (iv) in the timing of the onset and peak of convective precipitation (i.e. diurnal cycle) and (v) for high precipitation intensities (Prein et al., 2013a, 2015).

To extend the investigations of previous work, this chapter examines the added value of the limited area model ALARO-0 through analyses of subdaily precipitation characteristics at multiple spatial resolutions. Furthermore, the central question in this assessment is whether the multiscale performance of ALARO-0 in the simulation of daily summer precipitation, as was found in the previous chapter (Chapter 3), is also retained in the simulation of subdaily precipitation. Most studies assessing the benefits of CPMs in the representation of convective and subdaily precipitation generally compare simulations with horizontal grid spacings where deep convection is parameterized with those with resolutions where the convective parameterizations are turned off. However, to ensure a fair and consistent comparison between both resolutions, fundamental differences in the model physics and in the approach of the treatment of deep convection should be avoided. Although, it is suggested to use spatial resolutions of less than 4 km for the CPM climate simulations, the highest spatial resolution considered in this chapter is limited to 4 km. This can be justified by the fact that (i) studies with models running at grid spacings below 4 km report a weak grid sensitivity to the simulation of precipitation, and (ii) the use of physical parameterizations result in a similar or even larger spread. Hence, Prein et al. (2015) states that it appears more urgent to address aspects of physical parameterizations before further refining the horizontal resolution, with prioritizing the design of scale-aware physical parameterizations (Prein et al., 2015). The ALARO-0 model uses the Modular Multiscale Microphysics and Transport (3MT) physics package for clouds and deep convection. As described previously (see Chapter 2), the so-called "multiscale" characteristic of 3MT indeed allows a consistent use of the same model physics at different spatial resolutions without the need of different schemes or schemes that need to be switched off or on when changing the models' spatial resolution. Therefore, priority has indeed been given to the scale-awareness of the physical parameterization, which is a fundamental characteristic present in ALARO-0, rather than increasing the spatial resolution to resolutions below 4 km.

Subdaily precipitation values from reanalyses driven ALARO-0 simulations at 40-, 10-, and 4-km resolution are analyzed through assessment of the diurnal cycle, intensity and frequency characteristics and scaling properties such as the linear behavior of the Generalized

Extreme Value (GEV) parameters and the Clausius-Clapeyron (CC) relation. The next sections (Sections 4.2 to 4.3) provide a description of the data and methods, and in Section 4.4 the results on the representation of subdaily precipitation are discussed. We end this chapter with the conclusions in Section 4.5.

4.2   Data

4.2.1   Model data

This chapter uses the same 30-yr (1961-90) ERA-40 driven 40-km, 10-km, and 4-km ALARO-0 simulations as in Chapter 3 (abbreviated as respectively ERA40 - ALR40, ERA40 - ALR10, ERA40 - ALR04 in the figures). Compared to the ERA-40 reanalysis (Uppala et al., 2005), the more recent ERA-Interim reanalysis has significant improvements in the global hydrological cycle in terms of water vapor, clouds and precipitation (Betts et al., 2009). Therefore, in addition to the ERA-40 driven runs, a dynamically downscaling of the ERA-Interim reanalysis (Dee et al., 2011) was performed. The experimental design is similar to the ERA-40 driven runs, i.e. the ERA-Interim reanalysis is dynamically downscaled using the ALARO-0 model at a spatial grid of low spatial resolution of 20-km resolution corresponding to a 149 × 149 grid point domain that encompasses most of Western Europe. This 20 km output are then used to perform a one-way nesting on a 181 × 181 4-km resolution domain centered on Belgium (see Chapter 2, Section 2.3.2 and Chapter 3, Section 3.2.1 for more details on the experimental setup). The 30-yr (1981-2010) 4-km resolution ERA-Interim driven ALARO-0 model results are also used here for the analysis (abbreviated as ERAINT - ALR04 in the figures).

4.2.2   Observations

Sufficiently long time series of observed precipitation at high-temporal resolutions are crucial for model validation of extreme precipitation events at subdaily time scales. However, a dense network of such point observations from gauge measurements are generally difficult to find (Ban et al., 2014). Gridded hourly radar data could provide a good alternative to gauge measurements, as they have excellent resolutions both in time and space. Rainfall estimates from a 10-yr (2005-2015) dataset of volumetric weather radar measurements from a radar located in the southeast of Belgium (Wideumont), have recently been processed by Goudenhoofdt and Delobbe (2015). As consistent long-term gridded radar datasets are not yet available for our study period, and attenuation of the radar signal can lead to underestimation of the higher precipitation intensities (Chan et al., 2014), the observation data is here limited to available gauge data at high temporal resolutions.

The observation network of the Royal Meteorological Institute of Belgium (RMI) consists of several automatic weather stations (AWSs) which measure amongst other meteorological variables, 10-min precipitation. Nevertheless, the measurement period of these AWSs covers only approximately the last 10 years and thus does not correspond to our study period and is too short for studying heavy precipitation characteristics. An exception is the station of Uccle in Belgium (longitude: 4.358°E, latitude: 50.798°N), where an unique and long-term time series of 10-min precipitation is available and is collected with

high accuracy by the RMI. The measuring accuracy is homogeneous; until 2008 the rainfall is recorded by the same measuring instrument (a Helmann-Fuess pluviograph) at the same location since the start of the measurements in 1898. Furthermore, the precipitation data are recalibrated by the institute to the daily precipitation amounts from a non-recording pluviometer (or rain gauge). This recalibration ensures that the sum of the 10-min measurements of the pluviograph equals the daily observed totals as measured by the pluviometer, and it partly corrects the 10-min data for measurement biases by wind effects and wetting losses (Willems, 2000; Ntegeka and Willems, 2008). Because of several difficulties in operating mechanical pluviographs, the pluviograph records are since July 2008 replaced by data from modern automatic pluviometers (i.e. AWS). In order to limit the lack of homogeneity in the 10-min time series, the 10-min controlled AWS precipitation data are still recalibrated against the daily pluviometer observations. However, the recalibration process is since then done with a new tool, so that the precision of the final recalibrated 10-min data has become 0.01 mm instead of 0.1 mm as it was before. Nevertheless, the new recalibration procedure ensures that the difference between the recalibrated 10-min values and the daily pluviometer measurements is smaller than 0.1 mm (M. Journée, personal communication, 2015).

Finally, in order to analyze the dependence of extreme hourly precipitation on temperature, daily mean temperatures from the climatological station in Uccle are also used.

4.3 Methods

4.3.1 Data processing and analysis

For the evaluation of the modeled subdaily precipitation, the 10-min Uccle rainfall records as well as the hourly model series are aggregated to durations of 1-hour up to 24-hour precipitation amounts for the study periods 1961-90 and 1981-2010. For the data aggregation, a moving time window of a width equal to the respective duration was conducted over the 1-hourly datasets. Precipitation intensities are then calculated by dividing the total rainfall amount falling during 1-, 2-, 3-, 6-, 12-, or 24 hour by the respectively duration.

Daily mean temperature values used for the assessment of the CC relation (see Section 4.3.4), are simply derived by averaging observed and modeled daily maximum- and minimum 2-meter temperature values. Observations of daily maximum (minimum) temperatures are measured between 8 and 8 Local Time (LT) of the day after (before). The modeled daily maximum- and minimum temperatures are calculated accordingly the definition of the observations.

Since the improvements in the simulation of subdaily precipitation are mostly found during summer, when the highest precipitation intensities are usually related to convective showers, we mostly focus our analyses on the summer season (i.e. June-July-August, JJA). For some analyses, results for winter (December-January-February, DJF) are also given. Since the ERA-40 driven simulations were only done for the summer period 1961-90, the results for winter are based upon the ERA-Interim driven 4-km ALARO-0 model results for the period 1982-2010.

4.3.2 Extreme Value Analysis

Generally, the characteristics of extreme (sub)daily precipitation events are investigated by means of a theoretical extreme value distribution. One candidate distribution is the Generalized Pareto Distribution (GPD), that describes the maximum (sub)daily values by the exceedances of a large enough threshold u (see previous Chapter 3). As an alternative to describe extremes within a certain period, one can also consider the GEV distribution. Hence, to assess the added value of ALARO-0 for hourly precipitation extremes, we assume that the (sub)daily annual precipitation extremes follow a GEV distribution. According to the extremal types theorem (Coles, 2001), this distribution has often been used to model block (annual or seasonal) maxima of observed and modeled precipitation (Hanel and Buishand, 2010). The GEV *cumulative* distribution function $F(x, \mu, \sigma, \gamma)$ is given by:

$$F(x, \mu, \sigma, \gamma) = \exp\left\{ -\left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\gamma} \right\} \qquad \text{for } \gamma \neq 0,$$

$$\text{and for } 1 + \gamma \left(\frac{x - \mu}{\sigma} \right) > 0; \tag{4.1}$$

$$F(x, \mu, \sigma, \gamma) = \exp\left\{ -\exp\left[-\left(\frac{x - \mu}{\sigma} \right) \right] \right\} \qquad \text{for } \gamma = 0. \tag{4.2}$$

with x the sampled maxima, and μ, σ and γ the location, scale and shape parameter, respectively. The parameters satisfy $-\infty < \mu < +\infty$, $\sigma > 0$, and $-\infty < \gamma < +\infty$. The GEV distribution combines three asymptotic extreme value distributions into one single distribution, determined by the shape parameter γ: type I, light-tailed Gumbel distribution ($\gamma = 0$); type II, heavy-tailed Fréchet distribution ($\gamma > 0$); and type III, bounded Weibull distribution ($\gamma < 0$) (Nikulin et al., 2011; Russo and Sterl, 2012).

As an illustration, Fig. 4.1 shows the GEV *probability density* distribution of the hourly observed summer annual maxima intensities in Uccle for the period 1961-90. The values in the legend correspond to the GEV parameter estimates. The red dotted line indicates the location parameter (μ) which specifies the center ($\approx$ mean) of the distribution, but does not influence the standard deviation and higher order central moments. The scale parameter (σ) determines the width of the distribution or the size of deviations around the location parameter, and the shape parameter (γ) controls the rate of tail decay, with positive values implying a heavy upper tail (Hanel and Buishand, 2010).

There are two common methods for fitting the GEV distribution to the annual extremes: L-moments and maximum likelihood. The L-moments method is much more computationally efficient and also for small samples the shape parameter generally has better sampling properties than the maximum likelihood method (Nikulin et al., 2011). Because of this and the fact that our study period is relatively short (i.e. 30 years), the L-moments method is used here for fitting of the GEV distribution to summer annual maxima precipitation intensities for different durations d.

A goodness-of-fit test is carried out to check whether the GEV distribution fits the sampled annual/seasonal extremes satisfactorily. The same approach is followed as in Kharin and Zwiers (2000). A Kolmogorov-Smirnov (K-S) goodness-of-fit is applied, which measures the overall difference between two (cumulative) distribution functions. The K-S statistic D is defined as the maximum absolute difference between two cumulative distribution

functions (see also previous Chapter 3):

$$D = \max_{-\infty < x < \infty} | S_n(x) - F(x) |, \tag{4.3}$$

where $F(x)$ is the fitted cumulative distribution function and $S_n(x)$ is an empirical cumulative distribution function estimated from a sample of size n as the proportion of data values less than or equal to x. The null hypothesis (H_0) that the seasonal extremes are realizations from the GEV distribution $F(x)$ is rejected when D exceeds a certain critical value. Since the GEV distribution parameters are estimated from the data, these critical values taken from statistical tables would result in a too conservative test, i.e. H_0 is rejected less frequently than indicated by the significance level (Nikulin et al., 2011). Therefore, the critical value is determined by a parametric bootstrap procedure. In a first step, 1000 samples of size 30 (i.e. number of years in the study period) are randomly generated from the fitted GEV, and for each of the 1000 samples, values of D are derived. The 95th percentile of the resulting collection of Ds is then used as the critical value for the rejection of H_0 that the annual/seasonal maxima are drawn from the GEV distribution at the 5% significance level (Kharin and Zwiers, 2000).
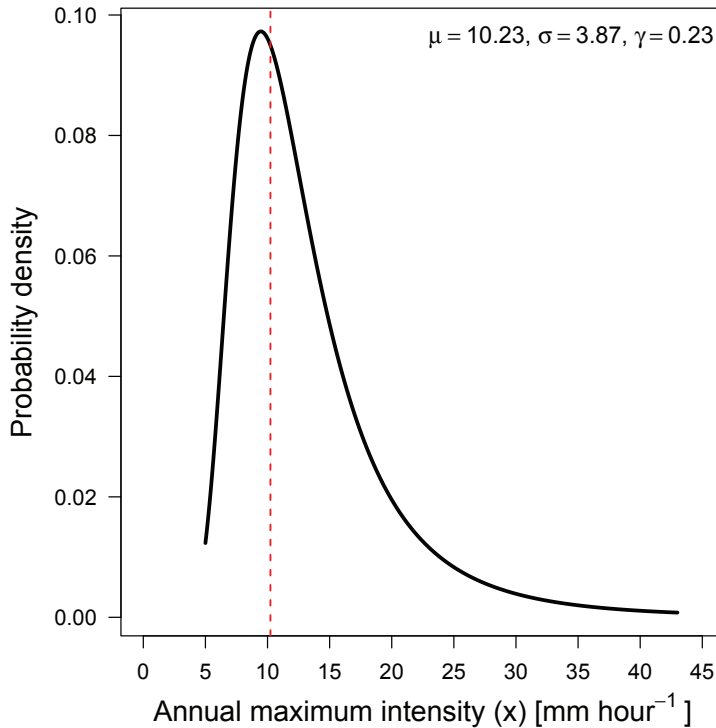
Figure 4.1: Generalized Extreme Value (GEV) probability density distribution of the hourly observed summer annual maxima intensities in Uccle for the period 1961-90. The values in the legend correspond to the GEV parameter estimates. The red dotted line indicates the location parameter (μ) which specifies the center ($\approx$ mean) of the distribution.

The return period T associated with return level x_T is the average time (expressed in

years) between two successive exceedances of x_T:

$$T = \frac{1}{1 - F(x_T)}. \tag{4.4}$$

After fitting the GEV distribution to the sampled extremes the T-year return level, x_T can be estimated by inverting the GEV cumulative distribution function [Eqs. (4.1) to (4.2)]:

$$x_T = F^{(-1)}\left(1 - \frac{1}{T}\right) = \mu - \frac{\sigma}{\gamma}\left\{1 - \left[-\log\left(1 - \frac{1}{T}\right)\right]^{-\gamma}\right\}, \quad \text{for } \gamma \neq 0; \tag{4.5}$$

$$x_T = F^{(-1)}\left(1 - \frac{1}{T}\right) = \mu - \sigma\log\left[-\log\left(1 - \frac{1}{T}\right)\right], \quad \text{for } \gamma = 0. \tag{4.6}$$

The estimated GEV parameters are then used to calculate return levels x_T for each rainfall duration d:

$$X \sim \text{GEV}[\mu(d), \sigma(d), \gamma] \tag{4.7}$$

$$\leftrightarrow x_T(d) = \mu(d) - \frac{\sigma(d)}{\gamma}\left\{1 - \left[-\log\left(1 - \frac{1}{T}\right)\right]^{-\gamma}\right\} \tag{4.8}$$

As the estimated shape parameters (γ) are for all durations and datasets different from 0, only the case for $\gamma \neq 0$ is considered in what follows. For a fixed duration d and return period T, the return level $i_T(d)$ is given by the Intensity-Duration-Frequency (IDF) relationship:

$$i_T(d) = \frac{\mu - \frac{\sigma}{\gamma}\left\{1 - \left[-\log\left(1 - \frac{1}{T}\right)\right]^{-\gamma}\right\}}{d^{\eta}}. \tag{4.9}$$

Furthermore, an important rainfall feature of particular interest for extremes is that, to a first approximation, the IDF curves display a power law dependence on averaging duration d and return period T. This scaling property of rainfall can be related to the fact that the estimated GEV parameters such as the location parameter μ and scale parameter σ, have a power law of the aggregation times (Willems, 2000). This means that if the parameter values are known for one particular aggregation time, they are also known for all other aggregation times only by applying a scaling factor. This can be demonstrated by equalizing Eq. (4.8) and Eq. (4.9):

$$\mu(d) = \mu d^{-\eta}, \tag{4.10}$$

$$\sigma(d) = \sigma d^{-\eta}, \tag{4.11}$$

with η the scaling factor. From this scaling property, Intensity-Duration-Frequency (IDF) relations can directly be derived. The above methodology is used to investigate whether the observed scaling properties are reproduced by ALARO-0 for different durations and at multiple spatial resolutions (De Troch et al., 2014a).

4.3.3   Upscaling and areal reduction factors

To compare the model data with the station point of Uccle, the nearest model grid box is selected. It is well known that the comparison of point observations and grid box mean values is not a perfect way of validation (Ban et al., 2014). To account for these issues in scale inconsistency, the high-resolution 4-km and 10-km data are upscaled through bilinear interpolation to a common 40-km analysis subdomain covering Belgium (Fig. 4.2).
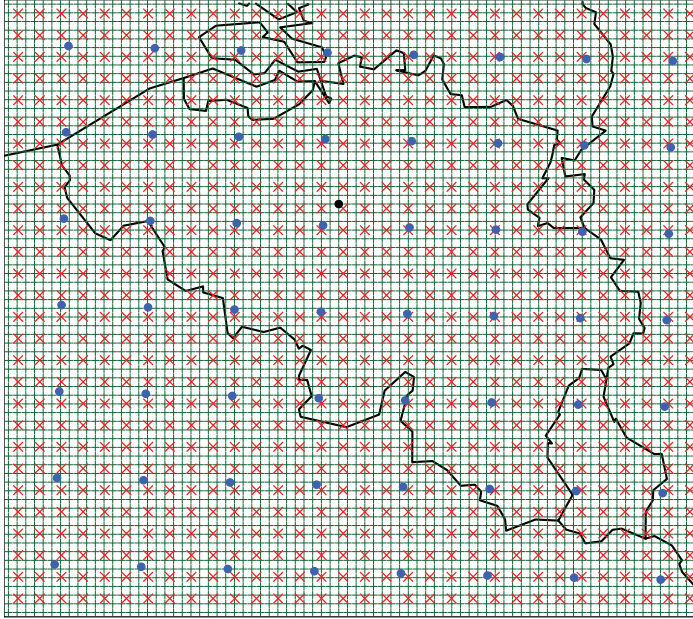
Figure 4.2: *High resolution 4-km (green plus signs) and 10-km (red crosses) data points are upscaled towards a low-resolution 8×7, 40-km resolution subdomain (blue dots) covering Belgium. The location of the 4-km grid point closest to the station of Uccle is also shown, and indicated by the black dot.*

In addition, to relate the maximum grid box rainfall estimates to the observed maximum rainfall rate the concept of areal reduction factor (ARF) can be used. It is generally accepted that the grid box rainfall of RCMs have the spatial characteristics of areal averages. Therefore, we may assume that the summer annual maxima intensities from the ALARO-0 model simulations represent a 40-, 10-, and 4-km climate, respectively. However, it can be expected that these maximum areal average rainfall rate will be less than the maximum rate estimated at the station point. This difference is usually referred to as the ARF (Fowler et al., 2005). To account for these differences in scale between observed and modeled summer annual maxima intensities, an ARF value can be applied to the observation point rainfall estimates of a specified duration and return period to give the areal rainfall of the same duration and return period. These ARF values are found to be site dependent, and also vary in time and with duration, size of the averaging area, and return period (Fowler et al., 2005; Willems, 2013).

Empirical ARFs for rainfall extremes have been reported for Belgium for 1-day durations and longer, and were found to vary between 0.8 and 0.9 for grid sizes ranging between 25-50 km (Willems, 2013). However, to our knowledge, no ARF values for durations lower than 24 hour for our study region have previously been estimated. Therefore, ARF values are calculated similar to Mailhot et al. (2007):

$$ARF(T,d) = \frac{x^{(g)}(T,d)}{x^{(s)}(T,d)} \tag{4.12}$$

where $x^{(s)}(T,d)$ and $x^{(g)}(T,d)$ are the return level estimates associated to events of duration d and return period T, respectively at the station scale (i.e. Uccle) and the grid box scale (i.e. ALARO-0 at 40, 10, and 4 km resolution). This ARF value is usually estimated over a certain region, corresponding to the *average* areal reduction factor between sites and grid box values, and obtained with regionalized values (Mailhot et al., 2007). Such regional averaged ARF values are then commonly applied to the return level estimates from the observations, to allow a fair comparison with those estimates from the RCM data (Fowler et al., 2005). However, in our case the ARF values are only calculated for the station location of Uccle and its surrounding grid box areas. As can be seen from Eq. (4.12), applying our ARF values to the observed return level estimates, would again result in the simulated rainfall return levels. Therefore, we do not aim to calculate the ARFs to upscale the observed point rainfall estimates in Uccle, but rather to give an indication for the order of magnitude of the differences expected between point and grid averaged intensities and how they vary with duration, area size, and return period.

Furthermore, in the hydrological community, so-called growth rates and curves are used as a common standard to characterize the difference between uncommon and extreme events. The growth rate, or curve G_T, defined as the multiple increase of the Tth-year return level over an index extreme value, is calculated as:

$$G_T = \frac{x_T}{x_i}, \tag{4.13}$$

with x_i some standard return level. This is usually chosen to be x_2, i.e. the return level corresponding to a 2-year return period (Chan et al., 2014). These growth curves allow direct comparisons between station and gridded data, because they are independent of areal reduction factors and mean biases (Chan et al., 2014).

4.3.4   Clausius-Clapeyron (CC) assessment

The CC relation expresses the capacity of the atmosphere to hold water, and is given by the rate of increase of the saturation pressure of water vapor e_s as a function of absolute atmospheric air temperature T_a (in Kelvin):

$$\frac{\partial e_s}{\partial T_a} = \frac{L_v e_s}{R_v T_a^2}, \tag{4.14}$$

where L_v is the latent heat of water vapor (2.5 x 10^6 J kg^{-1} at 0°C - this is a very weak function of temperature), and R_v is the gas constant of water vapor (461.5 J kg^{-1} K^{-1})

(Westra et al., 2014). The temperature as given in the denominator of Eq. (4.14) can be linearized around 0°C as follows:

$$T_a^{-2} = (273.15 + T_a^*)^{-2}, \tag{4.15}$$

$$\leftrightarrow T_a^{-2} = 273.15^{-2} \left(1 + \frac{T_a^*}{273.15}\right)^{-2}, \tag{4.16}$$

$$\leftrightarrow T_a^{-2} \cong 273.15^{-2} \left(1 - 2\frac{T_a^*}{273.15}\right), \tag{4.17}$$

where T_a^* is the temperature in degrees Celsius. Introducing Eq. (4.17) in Eq. (4.14) and filling in the above values gives:

$$\frac{\partial \ln e_s}{\partial T_a} \cong \frac{2.5 \times 10^6}{461.5} 273.15^{-2} (1 - 0.007 T_a^*), \tag{4.18}$$

$$\frac{\partial \ln e_s}{\partial T_a} \cong 0.073(1 - 0.007 T_a^*). \tag{4.19}$$

Furthermore, the actual saturation specific humidity q_{sat}, which is the mass of water vapor per kg of air, is given by:

$$q_{sat} = \frac{\varepsilon e_s}{p - (1 - \varepsilon)e_s} \cong \varepsilon \frac{e_s}{p}, \tag{4.20}$$

where ε is the ratio of the gas constant for dry air to that of water vapor (i.e. 0.622) and p is the atmospheric pressure (in Pa). Considering surface conditions, the pressure corresponds roughly to 10^5 Pa, and the vapor pressure e_s is only in the order of 1% of the actual pressure, so that the second term in the denominator in Eq. (4.20) can be omitted. From Eq. (4.20) it can be seen that the pressure of water vapor at saturation (i.e. relative humidity of 1) is direct proportional to the saturation specific humidity, and given Eq. (4.19) q_{sat} is thus by a good approximation exponential and increases by $\approx 7\%$ per degree at 0°C and $\approx 6\%$ per degree at 24°C (Westra et al., 2014).

Assuming constant relative humidity, the amount of water vapor in the atmosphere will thus increase at the CC rate. During extreme precipitation events it is commonly assumed that all water vapor in the air (or a constant fraction thereof) is converted to rain (Lenderink and van Meijgaard, 2010). Therefore, based on these thermodynamic principles it is expected that extreme precipitation increases relatively proportionally to surface temperature at the CC rate (Trenberth et al., 2003; Westra et al., 2014).

Recent studies using hourly precipitation observations from different locations in western Europe demonstrated that for temperatures above $\approx 10°$C, one-hour precipitation extremes increase approximately twice as fast as the CC relation (i.e. super-CC scaling of $\approx 14\%/°$C). For temperatures below 10°C the increase in precipitation extremes with temperature are according to the expected CC scaling of 7%/°C (Lenderink and van Meijgaard, 2008, 2010; Loriaux et al., 2013). Other observational studies for the United States and Australia also revealed this super-CC dependency for most of the assessed stations (Hardwick Jones et al., 2010; Mishra et al., 2012).

The cause and physical explanation for the super-CC scaling in the observations has been a point of discussion in several studies. According to Lenderink and van Meijgaard

(2008, 2010), the observed super-CC dependency is a robust property of convective extremes, while the CC scaling, as found for the daily and lower-temperature hourly analyses, describes the scaling of extremes of a stratiform origin. On the other hand, Haerter and Berg (2009) believe that the observed super-CC scaling might be a statistical artifact induced by the transition between the stratiform and convective regimes that both scale with CC. Both hypotheses agree in the sense that the super-CC scaling is somehow caused by a shift from stratiform to convective dominated extremes, but differ on the interpretation of the super-CC scaling (Loriaux et al., 2013). In an idealized modeling study, Loriaux et al. (2013) verified whether the super-CC dependency of hourly extreme precipitation is indeed robust for convective events. Therefore, the authors analyzed subhourly observed precipitation over the Netherlands, to select convective events from the dataset, and demonstrated a super-CC scaling over the entire range of dewpoint temperatures. From this, the authors conclude that the super-CC dependency is not a statistical artifact as suggested by Haerter and Berg (2009), but a robust relation for convective extreme precipitation which is likely due to enhanced convergence of moisture (Loriaux et al., 2013).

Several modeling studies also assessed the CC relation in Global Climate Models (GCMs) and RCMs (e.g. Allen and Ingram, 2002; Lenderink and van Meijgaard, 2008; O'Gorman and Schneider, 2009; Lenderink and van Meijgaard, 2010; Berg et al., 2009; Ban et al., 2014, 2015). O'Gorman and Schneider (2009) demonstrated that the rate of change in extreme precipitation with temperature vary widely among GCMs, ranging between 1.3% K^{-1} and 30% K^{-1}. At the regional scale, Lenderink and van Meijgaard (2008, 2010) for instance, found for the Netherlands that present-day modeled 99.9^{th} percentiles of hourly precipitation intensity, from 25-km output of the RCM RACMO2 and CLM, scale in accordance with the observations, i.e. a CC scaling for temperatures below $\approx 10°C$ and a two times CC relation for temperatures above $\approx 10°C$. However, for the lower percentiles and for temperatures above 22°C both models are not able to reproduce the observed CC dependency. Ban et al. (2014) derived the relationship between modeled daily mean temperature and hourly precipitation for the Alpine region using output from a 2.2-km CPM as well as a 12-km integration where convection is parameterized. Similar to the observations, the 2.2-km simulation exhibits for the 90th precipitation percentile a temperature dependence close to the CC scaling and a super-CC scaling for the most extreme hourly events (i.e. 99th and 99.9^{th} percentile). The 12-km model is able to represent the CC scaling relatively well, but it shows difficulties in reproducing the observed scaling for higher temperatures in regions of complex topography. Furthermore, Berg et al. (2009) investigated for different regions in Europe the CC relation in three ERA-40 driven RCMs with a horizontal resolution of about 50 km for the period 1961-90. In contrast to the aforementioned studies, the authors explored the scaling relationship between *daily* precipitation and daily temperature, and considered different seasons and months separately, to allow a more detailed investigation of the mechanisms responsible for the precipitation scaling. Their results show a seasonality in the temperature dependency of precipitation intensity, with a general increase roughly following the CC scaling in winter and a decrease in summer. From analyses of modeled atmospheric water vapor, cloud water and precipitation intensity, the authors demonstrated that during winter the CC relationship controls the increase in the large-scale precipitation with increasing temperature. While in summer, rather the availability of moisture than the atmosphere's capacity to hold this moisture, determines the negative scaling of daily precipitation with

increasing temperature.

It is clear that a consistent CC scaling of extreme precipitation is not generally found in the climate models. After all, besides the conceptual understanding of the potential thermodynamic relation between temperature and extreme precipitation, other factors such as microphysical contributions or dynamical processes, can also influence precipitation extremes leading to deviations from the CC scaling. The dynamic response arises because precipitation extremes are proportional to the associated pressure vertical velocity which is dependent on convective fluxes of heat and moisture, and these in turn depend on the large-scale circulation (O'Gorman and Schneider, 2009; Muller et al., 2011; Shepherd, 2014). Furthermore, since the super-CC relation is found to be a result of convective precipitation, which is often differently represented or parameterized by the models, it is not surprising that the rate of increase of extreme precipitation with temperature could widely vary among models (Ban et al., 2014). Therefore, it is interesting to assess how extreme precipitation values from the ALARO-0 model scale with temperature, and this for different timescales and seasons.

To derive the dependence of extreme hourly precipitation on temperature, we use the method of Lenderink and van Meijgaard (2008, 2010). Hence, the hourly precipitation is stratified based on the daily mean temperature in bins of 2°C width, with overlapping bins of 1°C in order to get good statistics in each bin. As we are interested in a proxy representing the temperature of the air mass, daily mean temperatures instead of hourly temperatures are used. Hourly temperatures are to a large extent influenced by variations and mixing processes in the planetary boundary layer and by radiation, and therefore may not be representative of the air mass considered (Lenderink and van Meijgaard, 2010). From the binned 1-hour precipitation data, extreme precipitation is then calculated as the 90th, 95th, 99th, and 99.9^{th} percentiles of the wet events (i.e. hours with precipitation more than 0.1 mm hour^{-1}). Similar to Ban et al. (2014), we assure the robustness of the estimated percentiles by setting a threshold of the number of precipitation cases in each temperature bin to $n = 50$ for the 90th and 95th percentile, and to $n = 100$ and $n = 500$ for the 99th and 99.9^{th} percentile, respectively. In case that the number of datapoints in a respective bin is lower than n, the percentile is not calculated. Since percentiles are calculated for different periods (e.g. seasons) or timescales (e.g. hourly, daily, monthly), the threshold value for n that sets the minimum number of datapoints to calculate the percentiles, will increase or decrease depending on whether the total number of cases in the whole dataset has increased or decreased. For percentiles calculated for hourly precipitation separately for winter and summer, the amount of data is lowered by a factor 4. Accordingly, the minimum number of hourly precipitation cases that should be in each bin, is roughly lowered by a factor of 4, corresponding to a threshold of n of 13, 25, and 125, for the 90th and 95th, 99th, and 99.9^{th} percentiles, respectively. Similarly, for percentiles obtained with daily precipitation, the threshold of n is lowered by a factor of 10, which thus sets the value of n to 5, 10 and 50 for the 90th and 95th, 99th, and 99.9^{th} percentiles, respectively. It is clear that the factor to which the threshold values for n are lowered, does not fully scale with the decrease in the number of cases in the dataset. However, the use of this threshold value rather assures robustness in the *visualization* of the results, and does not influence the results itselves.

The percentiles are calculated for the observed data of Uccle, as well as for the modeled

data using the closest grid box values from the ERA-Interim downscaling with ALARO-0 at 4 km resolution for the period 1981-2010. The ERA-Interim driven simulation is used for the CC assessment, because analyses are done for both the summer and winter season, and the ERA-40 driven simulations were only done for the summer period 1961-90.

4.4 Results and discussion

4.4.1 Diurnal cycle

Generally, the diurnal cycle of convection over land involves many processes between the surface, the boundary layer and the free troposphere such as surface exchange, turbulence, convection, and cloud-radiation interactions. Since all these processes are parameterized, the ability of a RCM or GCM to simulate correctly the diurnal cycle is an important test of the physical parameterization schemes within the models (Bechtold et al., 2004).

As mentioned previously in the introduction (Section 4.1), one of the major added value of CPMs can be found in the improvement of the diurnal cycle of summer precipitation. The mean diurnal cycle of JJA hourly precipitation obtained from observations in Uccle and from the closest model grid box values of the ERA-40 driven ALARO-0 runs at 4-km, 10-km, and 40-km resolution (ERA40-ALR04, ERA40-ALR10, ERA40-ALR40) is shown in Fig. 4.3(f). The mean diurnal cycle for the ERA-Interim driven ALARO-0 run at 4-km is also shown, together with the corresponding observed diurnal cycle for the JJA period 1981-2010. To put our result in perspective of other studies, we have added our result to an overview figure from Prein et al. (2015).

Both the high-resolution ERA40 and ERAINT results for ALARO-0, show a similar improvement in the simulation of the onset and peak of convective precipitation as all other CPMs shown in Fig. 4.3. The observed diurnal cycle of summer precipitation in Uccle is characterized by a minimum in the morning around 10 UTC and a convective maximum in the afternoon.
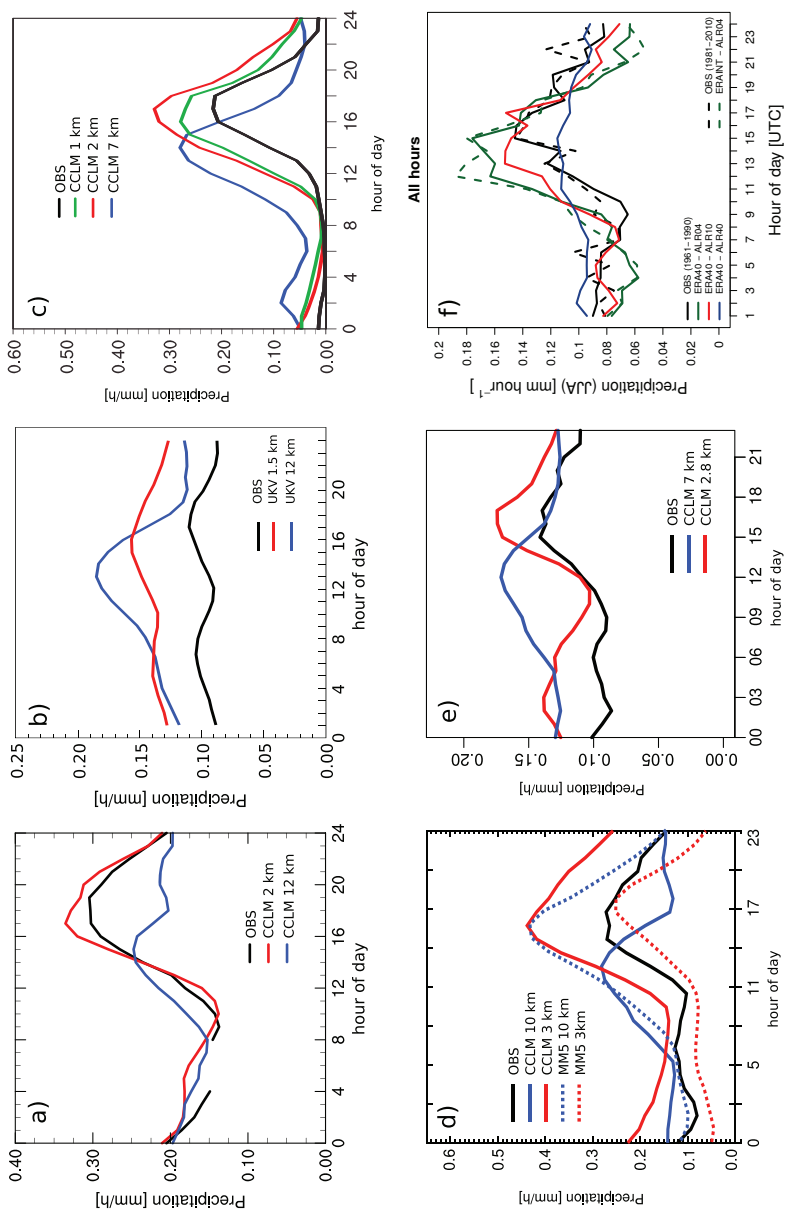
Figure 4.3: *Subfigures (a) to (e) are adopted from Prein et al. (2015). Mean diurnal cycles of (a) precipitation averaged across June-July-August (JJA) in Switzerland (Ban et al., 2014); (b) annually in Southern UK (Kendon et al., 2012); (c) July 2006 in Switzerland (Langhans et al., 2013); (d) June-July-August (JJA) in eastern part of the Alps (Prein et al., 2013a); (e) June-July-August (JJA) in Baden-Württemberg, Germany (Fosser et al., 2015); and (f) June-July-August (JJA) for all days in Uccle, Belgium.*

Compared to the low-resolution 40-km diurnal cycle (i.e. ERA40-ALR40 in blue), the timing and peak of precipitation are much better simulated by the high-resolution 4-km and 10-km runs (i.e. ERA40-ALR04, ERAINT-ALR04, ERA40-ALR10). Low resolution RCMs or GCMs are indeed known to have a too early onset and peak of convective precipitation and a too early decay of convective activity (Bechtold et al., 2004). However, compared to the observations, the high-resolution runs still show a too early onset and peak of convective precipitation. Observed precipitation starts to increase at 11 UTC and peaks around 16 UTC, while the build-up of convective rain in ALR04 and ALR10 already starts at 8 UTC, and it reaches its peak around 13 UTC.

To examine the reason for this shift in ALR04 and ALR10, we have constructed the 30-yr mean diurnal cycles for the wet hours (i.e. precipitation > 0.1 mm hour^{-1}) and for extreme precipitation hours defined as the hours with precipitation values higher than the 0.95^{th} quantile of the total dataset (Figs. 4.4(a) to 4.4(b)). It is notable that the too early onset as was seen before, has disappeared. The onset of convective rain occurs in both high-resolution simulations now at the same time as in the observations. Furthermore, in contrast to the diurnal cycle for all hours (Fig. 4.3 f)), the magnitude of the diurnal cycle is now consistently underestimated by ALR04 and ALR10, and in particular when only considering the wet hours. This suggests that the model overestimates the number of light rain events, and underestimates the number of dry cases with respect to the observations. The "drizzle effect" is a well known problem of RCMs (Kendon et al., 2012), and it has also been identified in previous chapter by the density curves obtained from the daily precipitation amounts (see Fig. 3.7 in Chapter 3).

Generally, most studies point the reason for the poor representation of the diurnal cycle of precipitation in the low-resolution models to the use of the deep convection parameterization scheme, which is for the high-resolution CPMs turned off (Langhans et al., 2013; Prein et al., 2013a; Ban et al., 2014). However, recent developments in the ALARO-0 model demonstrated that a correct description of the moist deep convection diurnal cycle does not only depend on the parameterization of convection, but also on the feedback coming from the interaction of radiation and cloud schemes (Brožková, 2015). The time shift in the diurnal cycle of ALARO-0 with respect to observations, characterized by a too early start of convection in the morning by a couple of hours and consequently too early decay of convective activity in the evening, has been a recognized problem of the model. Hence, in 2012, improvements in the 3MT physics scheme for moist deep convection were introduced, and these consisted of three major changes: (i) enhancement of the entrainment scheme by introducing a modification of the parameter of the buoyancy term, which takes into account relative humidity representing the gridcell; (ii) introduction of more memory, via the evaporation of precipitation in the previous time step of the model. More evaporation leads to higher and less entraining clouds in the next time step, which in turn allows for maintaining the convective activity longer; and (iii) adjustment in the computation of the updraft area fraction, resulting from the closure of the scheme, with respect to the mesh size of the model. These modifications indeed reduced the time shift in the onset and decay of convection. Nevertheless, the too early decay in the evening was still present. In 2014, a new model version ALARO-1, including - amongst other changes - the new radiation scheme (ACRANEB2), resulted eventually in combination with the previous improvements of 3MT

in a significant improvement of this last deficiency related to the timing of the diurnal cycle (Brožková, 2015). Since the version of ALARO-0 (cycle36), which is used for the simulations in this thesis, date from 2010, these improvements are not included, and a too early onset in the diurnal cycle is still present.
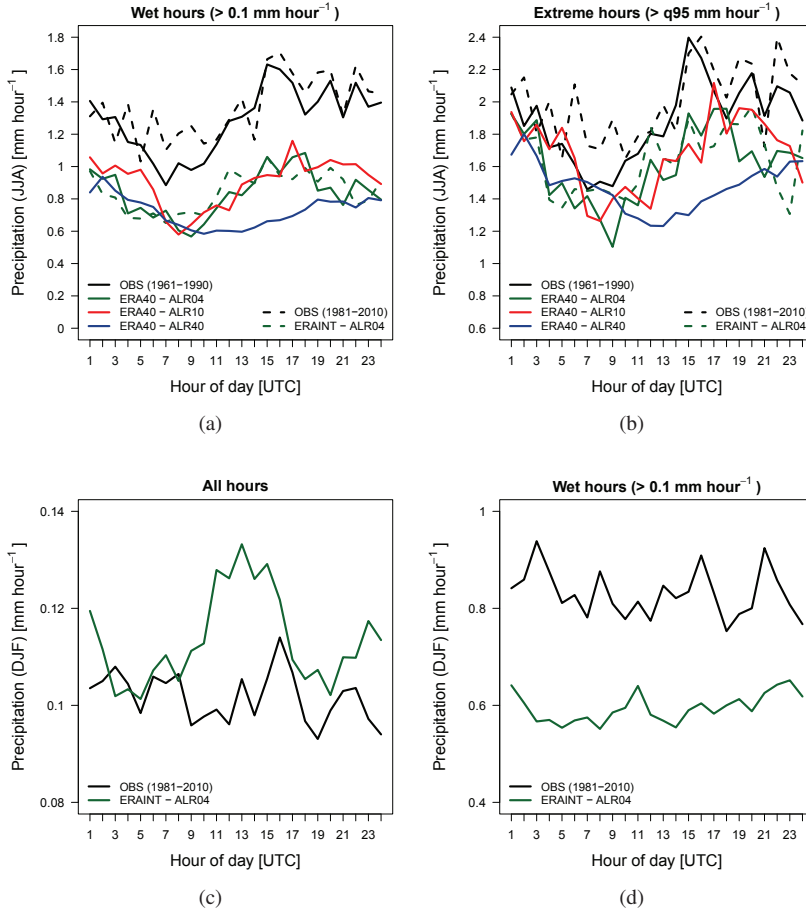
Figure 4.4: *30-yr mean diurnal cycles of summer rainfall for (a) wet hours (i.e. precipitation > 0.1 mm hour^{-1}), and (b) extreme hours (i.e. precipitation $> 0.95^{th}$ quantile), and of winter rainfall for (c) all hours, and (d) wet hours. Diurnal cycles are calculated for the observations (black) in Uccle and the closest model grid box values to Uccle (green for ERA40-ALR04 and ERAINT-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40).*

The mean diurnal cycle of winter precipitation (December-January-February) is also analyzed and shown in Figs. 4.4(c) to 4.4(d). The observed hourly precipitation intensity is quite uniform throughout the day, while the model (ERAINT-ALR04) shows a weak peak of precipitation in the early afternoon which is not visible in the observations. This peak in rainfall may again be attributed to the simulation of too much persistent light rain, since the model strongly underestimates the hourly, wet precipitation amounts throughout the whole

day (Fig. 4.4(d)).
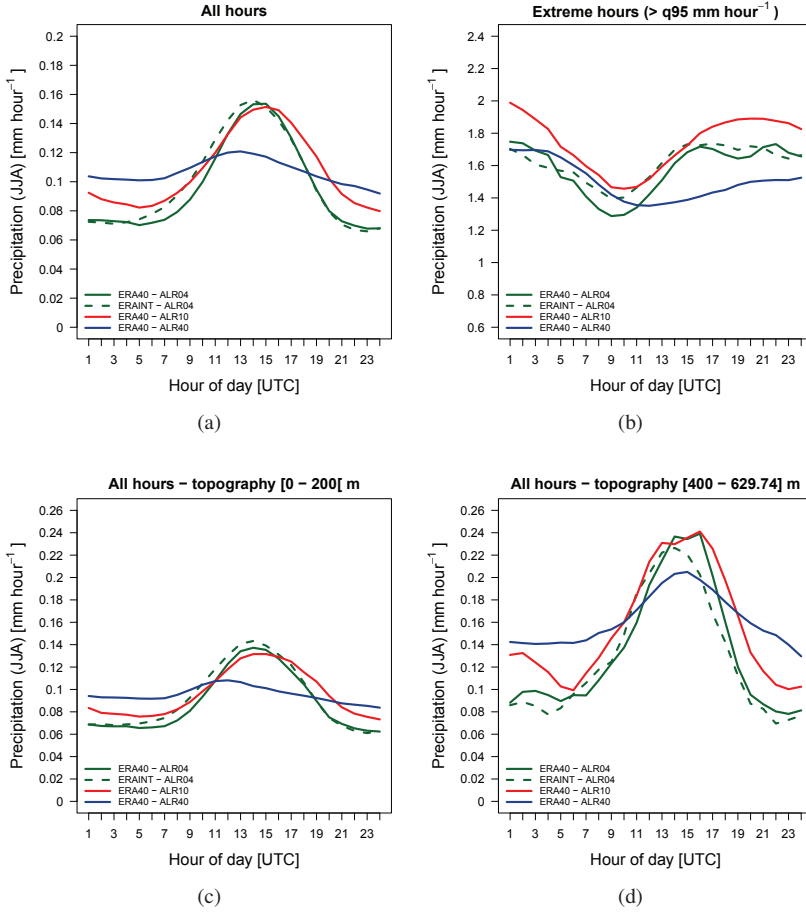
Figure 4.5: *30-yr mean diurnal cycles of summer rainfall for (a) all hours, (b) extreme hours (i.e. precipitation > 0.95^{th} quantile), (c) all hours and grid points with low topography (i.e. < 200 m), and (d) all hours and grid points with high topography (i.e. [400-630] m). Diurnal cycles are calculated for the upscaled 4-km and 10-km simulations, and 40-km simulation, across all 40-km grid boxes (8 × 7).*

Overall, the high-resolution ALARO-0 simulations show a better representation of the observed summer diurnal cycle than the low-resolution 40-km simulation (ERA40-ALR40 in blue), which clearly shows difficulties in capturing the observed convective peak, both in terms of magnitude and phase. The poor representation of the diurnal cycle by ERA40-ALR40, may be due to the fact that the amount of precipitation measured at a single point is not comparable with the amount averaged over a whole grid box, which is particularly true for small scale convection precipitation and coarse grids (Brockhaus et al., 2008). To exclude the effect of this important disadvantage of direct comparison of the diurnal cycle in a model

grid box against one single station of the same period, the diurnal cycle of the upscaled 4-km and 10-km resolution ALARO-0 data towards a common low-resolution 40-km grid (8 × 7 grid points), has been calculated. Figure 4.5 shows the 30-yr mean diurnal cycles of summer rainfall for the upscaled 4-km and 10-km simulations, and 40-km simulation, across all 40-km grid points of the common 8 × 7 subdomain.

First of all, it is important to note that the mean diurnal cycle for the regridded data is much smoother and similar to the diurnal cycles shown in Fig. 4.3(a)-(e). Using a grid-ded dataset to construct the diurnal cycle indeed smooths out all variability, with the latter clearly being present in the diurnal cycle based upon only one grid point (Uccle) as shown in Fig. 4.3(f). The mean diurnal cycles of rainfall for all hours as well as for extreme precipita-tion values, and averaged over the common 8 × 7 40-km subdomain, are shown in Fig. 4.5(a) and Fig. 4.5(b). The rainfall amounts throughout the day are given for the model only, since no gridded observation dataset for Belgium is currently available (see Section 4.2.2). To fa-cilitate the comparison with the diurnal cycles for Uccle, the same y axis is used to plot the modeled rainfall amounts. Except for the absence of the interhourly variability, the diurnal cycles for the upscaled and gridded model data are very similar to those for Uccle, and re-veal the same characteristics in timing and magnitude for the high-resolution (4 and 10 km) versus the low-resolution (40 km) simulations. This suggests a consistent improvement in the onset and peak of deep convection for ALR04 and ALR10, which is lacking in ALR40.

Finally, to assess the orographic effects on convective activity in Belgium, the mean diurnal cycles of summer rainfall are obtained for the grid points corresponding to low to-pography (i.e. 39 grid points with topography < 200 m) and high topography (i.e. 4 grid points with topography between [400-630] m) (Fig. 4.5(c) and Fig. 4.5(d), respectively). Similar to results for the Alpine region from Hohenegger et al. (2008), the ALARO-0 model reveals for all resolutions a strong dependency of the simulated daily convective develop-ment on the orographic forcing. The diurnal cycle for the grid points with low topography indicates that a weak orographic forcing hampers a clear simulation of the convective peak in the afternoon. However, simply a local effect reflected in the diurnal cycle can neither be excluded, as the 4 grid points with high topography are located next to each other.

4.4.2 Frequency and intensity characteristics

In this section we address the question whether ALARO-0 is able to reproduce the hourly precipitation statistics in terms of frequency and intensity. Figure 4.6 shows the 30-yr frequencies of hourly summer precipitation for the observations in Uccle, together with the closest model grid box values for the different simulations and resolutions (ERA-40-ALR04, ERA40-ALR10, ERA40-ALR40). The frequencies are obtained by binning the (low values of) hourly precipitation values into bins of (0.1) 1 mm hour^{-1}
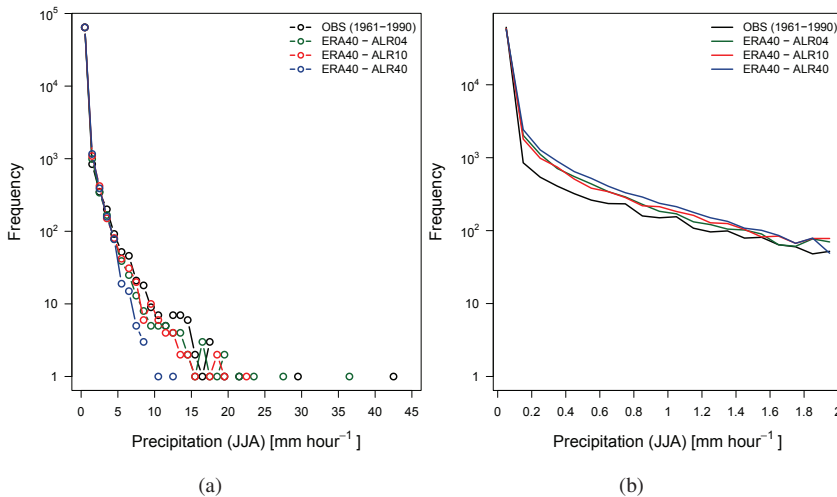
Figure 4.6: *Frequencies of (a) hourly summer precipitation, and (b) low values of hourly summer precipitation (i.e. [0-2] mm hour^{-1}) of observations, ERA40-ALR04, ERA40-ALR10, ERA40-ALR40. Frequencies are computed with the 30-yr (1961-90) hourly summer precipitation given for the station of Uccle and its nearest model grid box. Frequencies are displayed on a logarithmic scale.*

The high-resolution ALARO-0 runs reproduce the observed hourly precipitation frequencies relatively well, while the 40-km simulation clearly underestimates the frequency of hourly precipitation amounts. Both high-resolution simulations, and in particular ALR04, are able to capture the highest values of observed hourly rainfall. However, a slight underestimation of the observed values is still visible. The same findings are valid for the ERAINT-ALR04 simulation w.r.t. the observations for the 30-yr summer period 1981-2010 (see Appendix B, Fig. B.1).

As has been suggested in the previous section (Section 4.4.1), all models simulate consistently too much persistent light rain values (Fig. 4.6(b)). The slight underestimation of high-precipitation amounts and overestimation of low-precipitation values, are both long-standing issues of RCMs in the simulation of subdaily precipitation. To check on the sensitivity of this result with respect to the comparison of station point values with grid box mean values, frequencies have again been calculated for the upscaled 4-km and 10-km data to the 40-km grid. The frequency distributions are presented in Fig. 4.7, showing that the differences between the high-resolution runs (ALR04 and ALR10) and ALR40, as seen in

Fig. 4.6, remain, and even are larger.
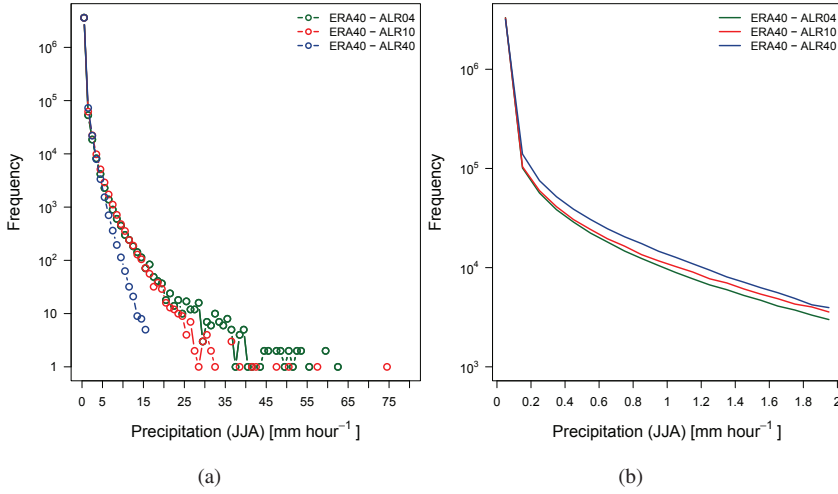
Figure 4.7: *As in Fig. 4.6, but frequencies are computed for the upscaled 4-km and 10-km simulations, and 40-km simulation, across all 40-km grid boxes of the common 8×7 subdomain.*

Hourly rainfall amounts corresponding to a range of percentile thresholds in observations and (upscaled) models are shown in Figs. 4.8(a) to 4.8(b). The 80th percentile of the hourly precipitation distribution corresponds to dry cases of 0 mm hour^{-1} in both observation datasets for the period 1961-90 and 1981-2010 (Fig. 4.8(a), black lines). The observation distribution for 1981-2010 even has a rainfall intensity of 0 mm hour^{-1} up to the 90th percentile. This is a consequence of the large number of hours with no rain such that the very high percentile thresholds (> 99) are needed to capture heavy rain (Kendon et al., 2012). In contrast, the 80th percentile of the hourly distribution of the models corresponds to light rain of less than 0.1 mm hour^{-1}, but it is still larger than 0 mm hour^{-1}. Although the simulations overestimate the light rainfall amounts, which are usually considered as numerical noise, they consistently underestimate the number of dry cases with respect to the observations. However, these differences between the observed and modeled precipitation can also be explained by the difference in the precision between the observation measurements and the modeled precipitation values, which is naturally lower for the observations (i.e. 0.1 mm until July 2008, and 0.01 thereafter) than for the model simulations. Nevertheless, the rainfall rate of the 80th percentile is lower in ALR04 and ALR10 than in ALR40. This indicates that the high-resolution models have considerably more dry spells than ALR40, which is in much better agreement with the observations. It can be seen that this result is robust for the upscaled rainfall distributions corresponding to the percentile thresholds (Fig. 4.8(b)).

For higher percentile thresholds (> 99), all models show an underestimation of the corresponding rainfall amounts. However, the tendency for underestimating the highest percentiles is significantly decreased for the highest resolutions, which is also in line with the frequency plots.
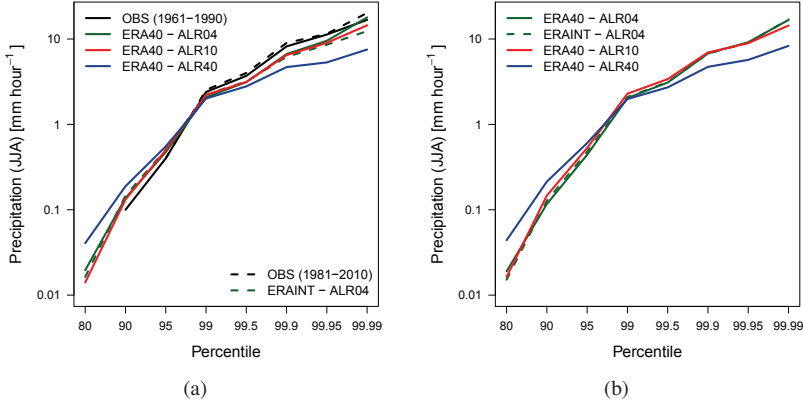
Figure 4.8: *Hourly summer rainfall intensities (mm hour^{-1}) corresponding to a given percentile threshold in (a) the observations and closest model grid box values for Uccle (ERA40-ALR04, ERA40-ALR10, ERA40-ALR40, and ERAINT-ALR04), and (b) the upscaled 4-km and 10-km simulations, and 40-km simulation. A percentile threshold corresponds to the pth percentile (80, 90, 95, 99, 99.5, 99.9, 99.95, 99.99) of the distribution of hourly rainfall for (a) Uccle, and (b) across all 40-km grid boxes of the common 8×7 subdomain and all 30 years in the respective datasets. Rainfall intensities are displayed on a logarithmic scale.*

Previous results suggest that the multiscale performance of ALARO-0 in the simulation of daily summer precipitation, as was found in the previous chapter (Chapter 3), does not hold for the simulation of subdaily precipitation. Furthermore, these results reflect that the added value in the simulation of precipitation at the subdaily timescale indeed can be found in the highest 4-km resolution ALARO-0 simulation. This finding is further investigated in the next sections, by means of two important scaling relations related to extreme precipitation at the subdaily scale.

4.4.3 Scaling properties

In general, the assessment of scaling properties of extreme (subdaily) precipitation is very important as it could provide insights in whether (and how) the scaling relations derived for present-day climate will manifest in a changing climate (Lenderink and van Meijgaard, 2010). In the next sections, two scaling properties related to extreme rainfall; i.e. the linear behavior of the GEV parameters and the CC relation, are examined.

4.4.3.1 Power law of the Generalized Extreme Value parameters

The GEV distribution [Eq. (4.1)] is fitted to the observed and modeled summer annual maxima precipitation intensities. A goodness-of-fit K-S test is applied to check whether the GEV distribution fits the summer annual maxima. As an illustration, Fig. 4.9 shows the observed summer annual maxima precipitation intensities for the different durations (gray vertical bars), together with the fitted cumulative GEV distribution function ($F(x)$, black

solid line) and the empirical cumulative distribution function estimated from a sample size of 30 ($S_{30}(x)$, red dots).
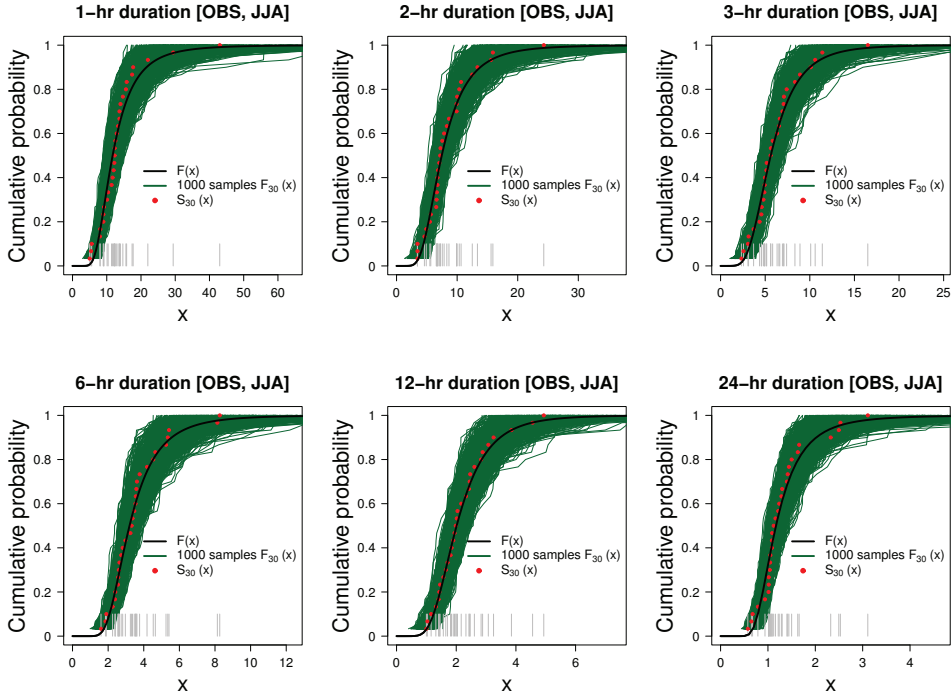
Figure 4.9: *Observed summer annual maxima precipitation intensities for the different durations (gray vertical bars), together with the fitted GEV distribution function ($F(x)$, black solid line) and the empirical distribution function estimated from a sample size of 30 ($S_{30}(x)$, red dots). The 1000 bootstrap samples of size 30, which are randomly generated from the fitted GEV ($F_{30}(x)$) correspond to the darkgreen solid lines.*

In addition the 1000 bootstrap samples of size 30, which are randomly generated from the fitted GEV are also presented ($F_{30}(x)$, darkgreen solid lines). As can be seen, the empirical cumulative distribution closely approaches the fitted cumulative distribution, and both distributions lie in the center of the 1000 bootstrapped distributions. This is also confirmed by the results from the K-S tests. The D statistic obtained from the maximum absolute difference between the empirical distribution and the fitted GEV distribution, is for all durations smaller than the critical value given by the 95th percentile from the 1000 D statistics, so that the null hypothesis that the seasonal extremes are realizations from the GEV distribution can be accepted at the 5% significance level. The same results are found for the modeled summer annual maxima precipitation intensities from ERA40-ALR04, ERA40-ALR10, and ERA40-ALR40 (not shown), indicating that the GEV distribution is a good choice for a distribution function of the observed as well as modeled summer annual maxima.
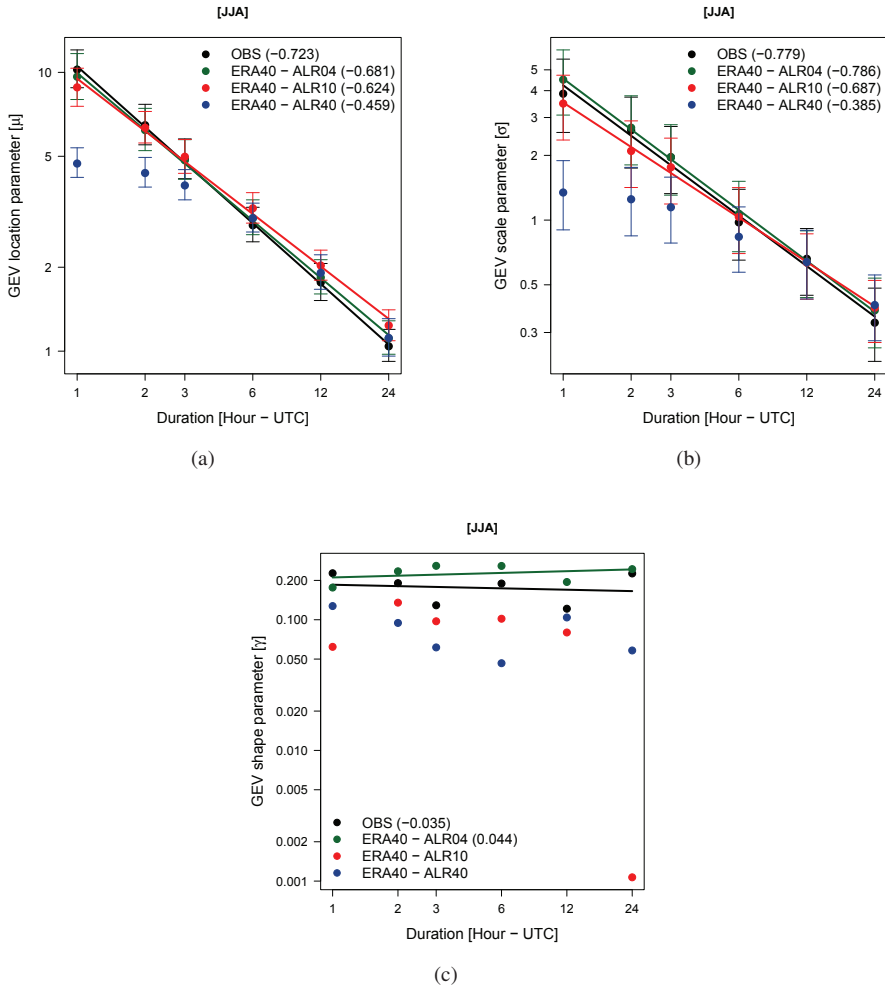
Figure 4.10: *Estimates of (a) GEV location parameter (μ), (b) GEV scale parameter (σ), and (c) GEV shape parameter (γ). GEV estimates (colored dots) are obtained by the L-moments method for different durations d of observed (black) and modeled (green for ERA40-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40) summer annual maxima precipitation intensities for Uccle. The error bars show the 95% confidence interval around the location and scale parameter estimates based on 1000 parametric bootstrap iterations and solid lines correspond to the linear regression lines between the logarithm of GEV parameter estimates and the logarithm of the different durations d, with the values of the slope given between parentheses in the legend. GEV parameter estimates and durations are displayed on a logarithmic scale.*

The GEV parameter estimates (location parameter μ, scale parameter σ, and shape parameter γ) are shown in Figs. 4.10(a) to 4.10(c) respectively, and presented by the colored dots for each duration. It can be seen that the location and scale parameter decrease with increasing duration. The decrease of the scale parameter with increasing duration reflects that the relative variability of the summer annual maxima intensities is large at short durations (Hanel and Buishand, 2010).

Furthermore, it is shown that for the observed and both high-resolution extreme rainfall amounts, the logarithm of the location parameter μ and scale parameter σ have a linear relationship with the logarithm of the duration d ranging between 1-24 hour (Figs. 4.10(a) to 4.10(b)). This linear behavior of the high-resolution extreme rainfall amounts, and in particular the 4-km values, corresponds very closely to the observed ones. The linear regression line between the logarithm of the parameter estimates and the logarithm of the durations is indicated by the solid lines, with the slopes of the regression given by the numbers in the legend between parentheses.

Table 4.1 shows for the observations and modeled data the estimated regression coefficients (i.e. slope and intercept) of the GEV parameter estimates.

GEV parameter		a	b	p-value
μ	OBS	-0.723	2.352	0.000*
	ERA40-ALR04	-0.681	2.297	0.000*
	ERA40-ALR10	-0.624	2.251	0.000*
	ERA40-ALR40	-0.459	1.748	0.002*
σ	OBS	-0.779	1.444	0.000*
	ERA40-ALR04	-0.786	1.515	0.000*
	ERA40-ALR10	-0.687	1.259	0.000*
	ERA40-ALR40	-0.385	0.447	0.001*
γ	OBS	-0.035	-1.686	0.773
	ERA40-ALR04	0.044	-1.557	0.526
	ERA40-ALR10	-1.032	-1.536	0.153
	ERA40-ALR40	-0.166	-2.307	0.315

Table 4.1: *Estimated linear regression coefficients of the GEV location (μ), scale (σ), and shape (γ) parameter estimates (i.e. a = slope, b = intercept). p-values of the F-statistic from the F-test, testing the significance of the linear regression, are also given. Linear regressions between the logarithm of the parameter estimates and the logarithm of the duration which are statistically significant at the 5% significance level, are indicated by an asterisk (i.e. p-value much smaller than 0.05).*

The significance of the linear regression is tested by the F-test. For both the observed and modeled location and scale parameters, the p-values from the F-statistic are much smaller than the 5% significance level (0.05), indicating that the null hypothesis that the regression coefficients from the linear fit are zero, can be rejected. Hence, it can be said that the linear relationship between the logarithm of both GEV parameters and the logarithm of the duration is statistically significant at the 5% significance level.

Although, the linear regression for the low-resolution 40-km location and scale parameter estimates is found to be statistically significant, the linear regression is not shown, as no clear linear dependency is visible (Fig. 4.10(a) and Fig. 4.10(b)). This is also reflected in the p-values of the F-statistic. The p-values of the low-resolution 40-km parameter estimates

(i.e. respectively 0.002 and 0.001) are indeed larger than those of the 4-, and 10-km parameter estimates (i.e. all ≈ 0). Furthermore, it can be seen from Fig. 4.10(a) and Fig. 4.10(b) that the 40-km GEV estimates deviate strongly from the observed estimations. However, from aggregation times of 6 hour, the 40-km location and scale estimates also start to approach the observed estimates. Based upon a 27-yr 10-min precipitation observation time series from Uccle, Willems (2000) also found this linear-log decrease of the GEV model parameters with the aggregation-level.

For the shape parameter (Fig. 4.10(c)), only the linear regressions for the observations and high-resolution 4-km model are displayed. This is rather illustrative, as overall there appears to be no systematic variation of γ with duration. This is also confirmed by the p-values which are all much larger than the 0.05 significance level, so that the H_0 that the regression coefficients from the linear fit are zero, cannot be rejected. However, for ERA40-ALR10 the result may be affected by the strong deviating estimate for the 24-h duration (Fig. 4.10(c)), the p-value is still found to be larger than the significance level. Hence, this suggests that γ may be considered to be constant, and that a linear regression is not an adequate fit between the logarithm of the shape parameter and the logarithm of the duration. These results, showing a linear dependency of the location and scale parameter, but no systematic change of the shape parameter with duration, are similar to the findings of Overeem et al. (2008) and Hanel and Buishand (2010).

The linear behavior of the estimated GEV location and scale parameter corresponds to the earlier described power law of the aggregation times [Eqs. (4.10) to (4.11)]. However, as it can be seen from the slope estimates in Fig. 4.10 and Table 4.1, a power law formulation with different estimates for the scaling factor η for the location and scale parameter, would be more correct:

$$\mu(d) = \mu d^{-\eta_1}, \tag{4.21}$$

$$\sigma(d) = \sigma d^{-\eta_2}, \tag{4.22}$$

with $\eta_1 \neq \eta_2$. Thus, from [Eq. (4.9)] a better IDF-relationship is then given by:

$$i_T(d) = \frac{\mu}{d^{\eta_1}} - \frac{\sigma}{d^{\eta_2}\gamma} \left\{ 1 - \left[-\log\left(1 - \frac{1}{T}\right) \right]^{-\gamma} \right\}. \tag{4.23}$$

The relations of GEV parameters as a function of duration d are then used to construct rainfall IDF curves. Figure 4.11 shows the JJA IDF-relationships from the model (green for ERA40-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40) and observational (black) data for Uccle. Intensities in the IDF-relation plots are calculated with (i) Eq. (4.8) using the parameters estimated from the GEV fit (circles), and with (ii) Eq. (4.23) using the location- and scale parameter derived from the power law given by Eqs. (4.21) to (4.22) (solid lines). The values for μ (σ) and η_1 (η_2) in Eqs. (4.21) to (4.22) correspond to the values of the intercept and slope from the linear regression of the location parameters (μ) and scale parameters (σ), respectively. As it has been demonstrated previously, the dependence of the shape parameter on duration is not significant. Therefore, for both intensity calculations (i) and (ii), one and the same mean value over all durations has been used for the shape parameter (γ). Intensities are plotted on log-log graphs, with the different panels representing different durations (1, 2, 3, 6, 12, and 24 h). First of all, it can be seen from

Fig. 4.11 that the use of the power law values for the location- and scale parameter, is a valid method for the derivation of IDF-relationships.
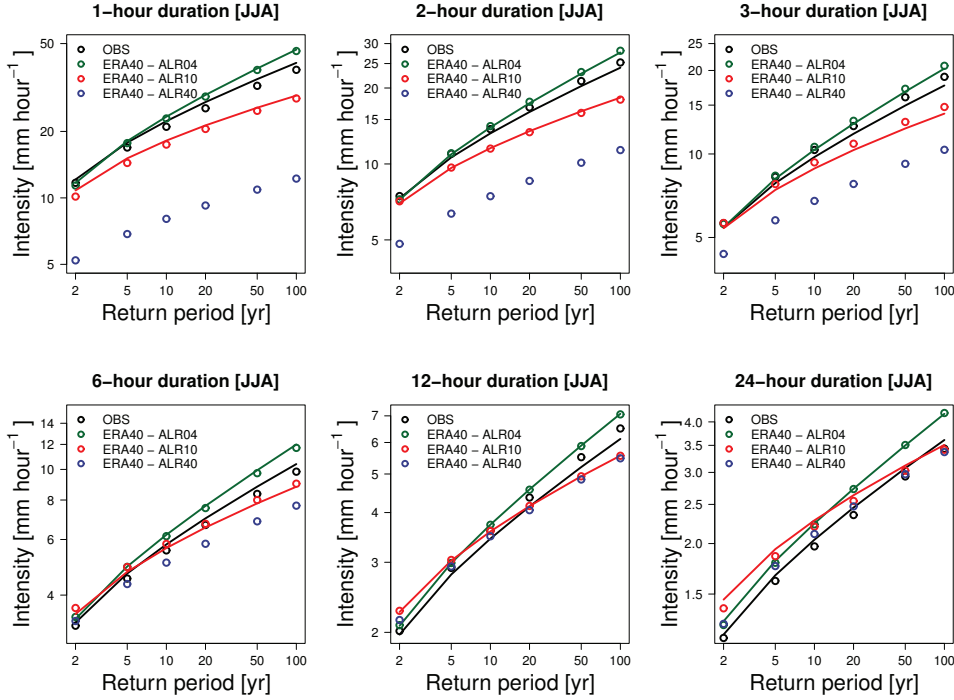
Figure 4.11: *JJA IDF-relationships with the intensities (return levels) calculated with (i) [Eq. (4.8)] using the GEV-estimated parameters (circles), and with (ii) [Eq. (4.23)] using the location- and scale parameter derived from the power law given by [Eqs. (4.21) to (4.22)] (solid lines). For both intensity calculations (i) and (ii) one and the same mean value over all durations for the shape parameter (γ) has been used. Intensities are given as a function of return period (T) for observations (black) and models (green for ERA4O-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40), plotted on log-log graphs, and computed for the station of Uccle and its nearest model grid box values. The different panels represent different durations (1, 2, 3, 6, 12, and 24 h).*

Furthermore, the estimated extreme precipitation intensities are for the highest-resolution simulation (ERA40-ALR04) and for all durations in close agreement to the observed estimated intensities. For the highest return periods (i.e. $T = 50$ years or $T = 100$ years) the modeled intensities deviate slightly from the observations. Since the intensity estimates are based upon summer annual maxima precipitation intensities for a period of only 30 years (1961-90), one should keep in mind that estimates for such large return periods are relatively uncertain. Similar to previous results, the 40-km low-resolution intensities strongly deviate from the observations, and consistently underestimate the observed intensities for all return periods. Again, from 6-hour durations onwards, ERA40-ALR40 also starts to approach very closely the observations.

The strong underestimation of the 40-km GEV location parameter and scale parameter for the lowest durations can be explained by the fact that the summer annual maxima of area-average precipitation, and consequently also the corresponding intensities, for a given dura-

tion tend to decrease if the area (or spatial resolution) becomes larger (Hanel and Buishand, 2010). To check the sensitivity of this reduction of the GEV parameters and consequently precipitation intensities with increasing area size on the results, the analysis has been repeated using the upscaled data for the 8×7, 40-km resolution subdomain. The results are shown in Appendix B, Fig. B.2 and Fig. B.3, and reveal similar conclusions: the estimated GEV scale and location parameters for the upscaled high-resolution 4-km and 10-km simulation show a linear dependency with duration, and the 4-km and 10-km IDF-curves show higher intensities than for the 40-km simulation, particularly for the shortest durations.

The increasing underestimation of precipitation amounts by low resolution RCM simulations with decreasing duration has also been found in other studies (e.g. Mailhot et al., 2007; Hanel and Buishand, 2010). It can be related to the average sizes of meteorological processes involved. We can assume that strong convective processes such as summer thunderstorms are responsible for the 1-h extremes. Such localized thunderstorms have spatial scales in the order of 10 km, which indeed are much smaller than the size of the average grid box of the low-resolution 40-km model. On the other hand, meteorological processes responsible for 24-h summer extremes are generally larger in scale, and can therefore possibly cover one or more 40-km model grid box (Mailhot et al., 2007). This is in agreement with our finding that from 6-hour durations onwards, ERA40-ALR40 starts to approach the observations, indeed suggesting that this "smoothing effect" increases as the duration of the summer extremes becomes longer.

To further investigate the relationship between return levels from the observation station point in Uccle and from the closest grid box estimates, ARF values have been calculated [Eq. (4.12)]. The return levels used for the ARF calculation are those obtained with the location- and scale parameter values derived from the power law [Eq. (4.23)]. Figure 4.12 shows how ARF values for Uccle evolve as a function of duration d (i.e. 1, 2, 3, 6, 12, and 24 hour) for different return periods T (i.e. 2, 5, 10, 20, 50, and 100 year). It can be seen that the ARFs vary with duration and size of the averaging area, showing for a fixed (grid box) area, lower ARF values for shorter durations (Mailhot et al., 2007). As mentioned previously, this is related to the fact that events involved in shorter duration extreme rainfalls are spatially smaller than those involved in longer duration extreme precipitation events. As this scale difference between duration and the respective size of the meteorological system involved, becomes larger when the area (resolution) becomes larger (lower), this variation is largest for the lower resolution ERA40-ALR10 and ERA40-ALR40 simulations.

Furthermore, in agreement with the ARF values obtained by Mailhot et al. (2007) for a region in southern Quebec, our ARF values of the high-resolution ERA40-ALR04 simulation suggest for all durations an increase as function of return period. In contrast, the lower resolution simulations ERA40-ALR10 and ERA40-ALR40 show an opposite response, i.e. ARFs are decreasing with increasing return period. This contrasting variation could be explained in a similar way as the variation of the ARF values with duration d. More specifically, longer return periods can generally be associated with high precipitation quantiles or more extreme precipitation amounts. As a consequence, it can be expected that the scale difference between extreme precipitation events associated with long return periods and the low resolution simulations will be large. Hence, for such long return periods (e.g. 100-yr return period), and in particular for the shortest durations, the ARF values of the lower res-

olution simulations (ERA40-ALR10 and ERA40-ALR40) deviate more from one than it is the case for ERA40-ALR04. Overall, since this variation with return period is dependent on the nature of the meteorological systems involved in the respective region during extreme events, differences in the variation of ARF values with return period are possible when other regions are considered (Mailhot et al., 2007).
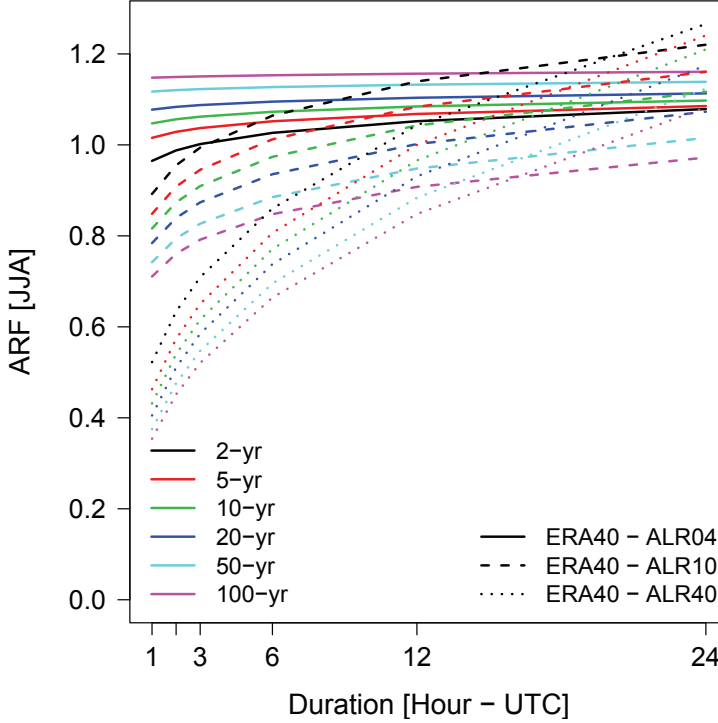
Figure 4.12: Areal reduction factor (ARF) between return levels for Uccle at the grid box and station scale as a function of duration d (i.e. 1, 2, 3, 6, 12, and 24 hour) for different return periods T (i.e. 2, 5, 10, 20, 50, and 100 year). The ARF is calculated with the return levels obtained with the location- and scale parameter values derived from the power law [Eq. (4.23)].

Finally, to compare in a consistent manner previous results from modeled grid box mean values with station point values, we have calculated so-called growth curves (see Section 4.3.3). The growth curves (Fig. 4.13) are in agreement with the findings from previous IDF-relation plots and ARF values. The differences between the low-resolution 40-km and 10-km growth curves and the observations are gradually reduced as the accumulation periods become longer, and the differences disappear for the 24-hour durations. Hence, changing the duration has little impact on the growth curves for the highest-resolution ERA40-ALR04 run, which is not valid for the lower-resolution ERA40-ALR40 and ERA40-ALR10 runs. Similar results were found by Chan et al. (2014), who used extreme value theory to compare subdaily extreme summer precipitation from two model simulations at 12-km and 1.5-km resolution over southern UK. The authors also found a gradually decrease in the difference between the 12-km RCM and the other datasets as the duration becomes longer.
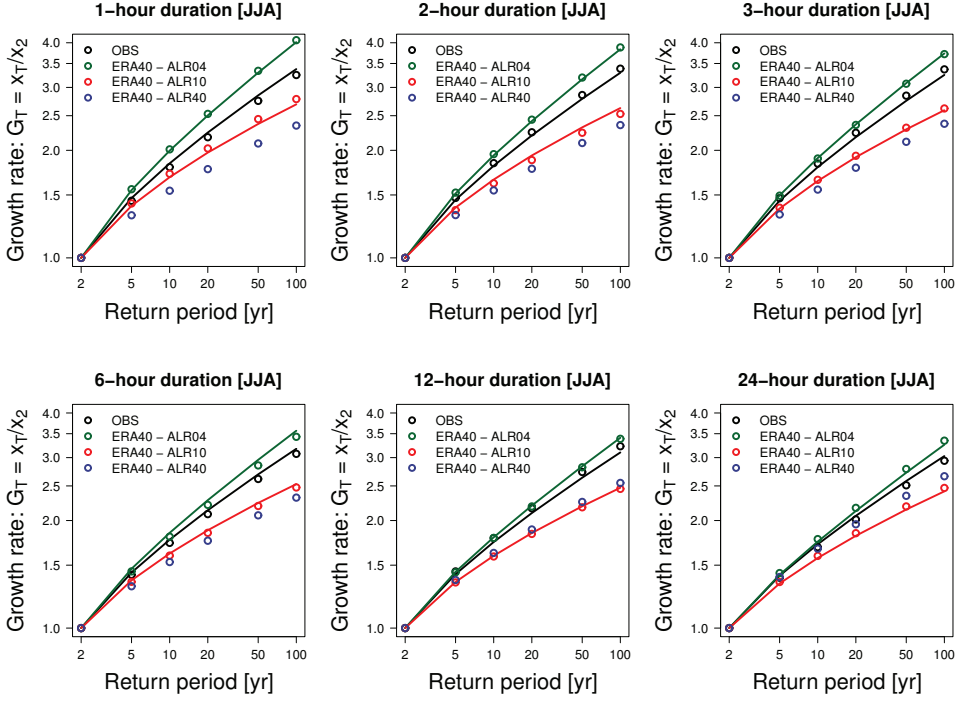
Figure 4.13: *JJA GEV-estimated growth curves [y axis $G_T = x_T/x_2$] as a function of return period (T). x_T and x_2 are the estimated intensities (return levels) for the Tth- and 2nd-return period, respectively. The intensities are again calculated in two ways (circles and solid lines), similar to the intensities given in Fig. 4.11. Colored lines correspond to the different datasets: green for ERA4O-ALR04, red for ERA40-ALR10, blue for ERA40-ALR40, and black for observations. Different durations (1, 2, 3, 6, 12, and 24 h) are shown in different panels and the axes are shown on a logarithmic scale.*

4.4.3.2 Clausius-Clapeyron (CC): Scaling of precipitation extremes with temperature

As described previously (Section 4.3.4), the CC relation gives the increase in the moisture-holding capacity of the atmosphere with temperature, approximately corresponding to a rate of 7% per degree temperature rise. Furthermore, it has been argued that the CC relation sets a scale for the increase in extreme precipitation extremes with global warming (Ban et al., 2014). We test this hypothesis, by deriving the relationships between local daily mean surface temperature and hourly precipitation from observational data for Uccle. However, RCMs do not necessarily reproduce the observed CC scaling, and the relationship between extreme precipitation and temperature can vary widely among models. Therefore, it is very relevant to assess how this relationship is represented in the ALARO-0 model w.r.t. other RCMs or CPMs, and to investigate whether the ALARO-0 model can reproduce the observed scaling. For this we use the closest grid box mean model output to the observation station of Uccle from the ERA-Interim downscaling at 4-km resolution (ERAINT-ALR04).

Figure 4.14 shows for the observations and ERAINT-ALR04 the aforementioned relation for the 90th, 95th, 99th, and 99.9^{th} percentiles of the 1-hour precipitation intensities (Fig. 4.14(a) and Fig. 4.14(b)) and daily maximum of hourly precipitation intensities

(Fig. 4.14(c) and Fig. 4.14(d)) as function of daily mean temperature for Uccle.
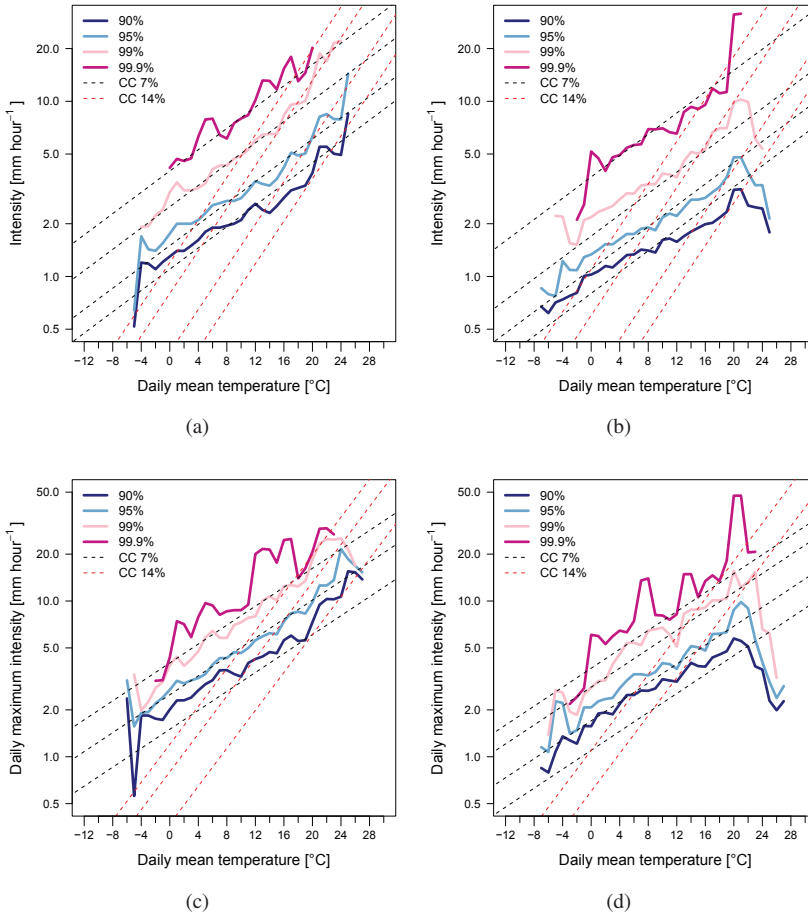
Figure 4.14: *Dependencies of different extreme percentiles (90th-99.9^{th}) of the distribution of (a, c) observed, and (b, d) modeled (ERAINT-ALR04) (a, b) hourly and (c, d) daily maximum of hourly precipitation on temperature in Uccle and its closest model grid box values. Exponential relations given by a 7% and a 14% increase per degree are given by the black and red dotted lines, respectively. Percentiles are displayed on a logarithmic y axis.*

To assess more in detail the observed and modeled scaling behavior, we have calculated the scaling exponents for different percentiles of hourly precipitation by fitting a linear regression between the logarithmic of the hourly precipitation amounts and the daily mean temperatures (Fig. 4.15(a)). Similarly, the scaling exponents are also obtained for the 99th percentile of different precipitation durations (Fig. 4.15(b)). It can be seen that the 90th, 95th, and 99th model percentiles of both the hourly precipitation as well as the daily maximum of hourly precipitation level off around temperatures of $22°C$ (Fig. 4.14(b) and Fig. 4.14(d)). To exclude the effect of this leveling off, the scaling exponents are calculated for daily mean temperatures to $22°C$. To quantify whether the observations and model

reproduce the super-CC scaling, as discussed previously in Section 4.3.4, the scaling expo-
nents are also calculated separately for temperatures below and equal to $12°C$, as well as for
temperatures above $12°C$.

Consistent with the CC hypothesis, observed extreme precipitation (Fig. 4.14(a) and
Fig. 4.14(c)) indeed increases at a rate of about 7% per degree (adiabatic scaling). As can
be seen from Fig. 4.15(a), the less extreme observed precipitation amounts given by the per-
centiles below the 90th percentile, show a scaling below the CC scaling. The 1-hour and
daily maximum hourly precipitation intensities generally exhibit a similar scaling behavior.
However, for temperatures above $\approx 12°C$, the 99th percentiles, and nearly all percentiles of
daily maximum hourly precipitation, clearly show a larger increase, close to the super-CC
rate. The observed scaling exponents for temperatures above $12°C$ are for almost all per-
centiles indeed larger than the CC scaling, and they show a maximum scaling of $\approx 12\%$ per
$°C$ for the 99th percentile (Fig. 4.15(a)). This is in agreement with results from previous
studies (e.g. Lenderink and van Meijgaard, 2008, 2010; Loriaux et al., 2013; Ban et al.,
2014). Nevertheless, the super-CC rate is not visible for the highest percentile of hourly pre-
cipitation, since no percentiles are calculated because of the high threshold for the minimum
number of data points that we have set to calculate the highest percentiles (i.e. 500).

The observed temperature dependency of hourly precipitation percentiles is relatively
well reproduced by the model (Fig. 4.14(b)). Although, the model exhibit a similar behavior
as the observations, it systematically underestimates the scaling exponents (Fig. 4.15(a)).

(a)   (b)
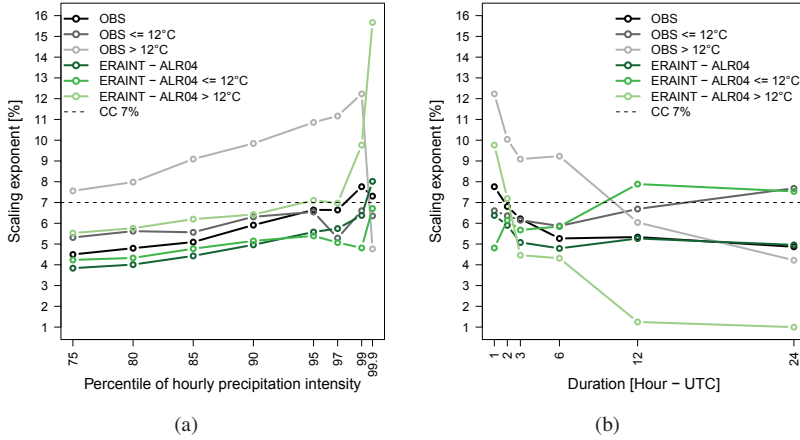
Figure 4.15: *(a) Variation of the observed (black) and modeled (ERAINT-ALR04, green) scaling ex-*
ponent with percentile of hourly precipitation, and (b) variation of the observed (black) and modeled
(ERAINT-ALR04, green) 99th percentile scaling exponent with precipitation duration. The scaling ex-
ponents are calculated for the whole temperature range up to $22°C$, as well as for temperatures below
and equal to $12°C$ and for temperatures above $12°C$. The horizontal black dotted line corresponds to
the theoretical 7% Clausius-Clapeyron (CC) increase per degree.

As shown by Fig. 4.15(a), for temperatures $\leq 12°C$ as well as for the whole temperature
range, all percentiles of hourly ERAINT-ALR04 precipitation increase with a rate lower
than the CC relation. Only for temperatures above $12°C$, the 90th, 95th and 97th percentiles

follow the CC scaling, and the highest percentiles show a dependency slightly below and according to the super-CC relation (Fig. 4.14(b) and Fig. 4.15(a)). However, as can be seen in Fig. 4.14(d) this super-CC behavior is not visible in the 99th percentiles calculated with the daily maximum of hourly precipitation.

Generally, the observations and model demonstrate an increasing scaling with increasing percentile, and a reduction in the 99th percentile scaling for longer precipitation durations (Fig. 4.15). This is in agreement with the results from Hardwick Jones et al. (2010), based upon station observations in Australia. Furthermore, the model only shows for the highest 99th and 99.9^{th} percentiles a scaling exponent close to or slightly larger than the 7%/°C CC scaling, while the observations approach the CC scaling already from the 95th percentile. As can also be seen from Fig. 4.15(b), ERAINT-ALR04 exhibits systematically lower scaling values than the CC scaling for all durations of the 99th percentile. Except for the longest (shortest) durations and for temperatures below (above) 12°C, the model shows a dependency close to or slightly above the CC scaling. The observations on the other hand, reproduce well the CC scaling for the 1-, and 2-hour durations, with scalings far above the 7%/°C scaling for temperatures above 12°C.

Overall, the modeled percentiles are in agreement with results from Lenderink and van Meijgaard (2008), who analyzed the CC relation for a simulation of the present-day climate (1971-2000) from the RCM RACMO2. The authors found for the highest percentiles (99th and 99.9^{th}) that the model reproduces well the observed CC (super-CC) relation for temperatures below (above) 12°C, but for temperatures above 20°C and for the lower percentiles, the model is not able to reproduce the two times CC relation.

To understand the relationship between frequency of extreme precipitation events and temperature, and the variations in intensity of precipitation, we estimated similar to Mishra et al. (2012), respectively the wet-time fraction (WTF) and fractional contribution of extreme precipitation events (above the 95th percentile) to total precipitation (PFRACT) in each temperature bin. The wet-time fraction is defined as the ratio of the total number of extreme precipitation events above the 95th percentile to the total number of wet hours (i.e. precipitation > 0.1 mm day^{-1}) in each bin. Figures 4.16(a) to 4.16(b) show the relationships between WTF and PFRACT and temperature for the observations and ERAINT-ALR04 in Uccle.

The observed WTF and PFRACT are relatively well captured by the model. However, the model appears to slightly overestimate the observed PFRACT. This is in line with previous result shown in Fig. 4.8(a), indeed showing an overestimation in hourly (summer) precipitation intensity for the 95th percentile. For both the observations and the model, WTF stays practically constant for the whole temperature range. On the other hand, PFRACT clearly increases more rapidly with temperature than WTF, suggesting that the regression slopes (i.e. scaling exponents) between extreme precipitation events and temperature are more influenced by changes in PFRACT (intensity) than in WTF (frequency) (Mishra et al., 2012).
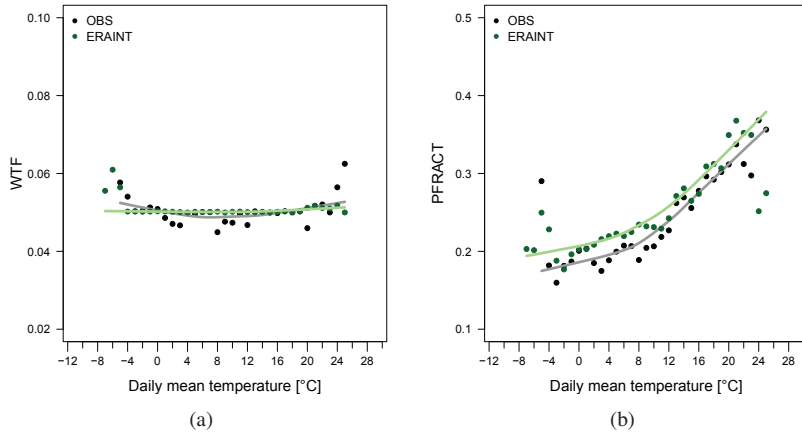
Figure 4.16: *(a) Relationship of observations (black) and model (ERAINT-ALR04, green) between wet-time fraction (WTF) and (b) fractional contribution of extreme precipitation to total precipitation (PFRACT) of precipitation above the 95th percentile and mean daily temperature in each bin. The thick solid lines are smoothed loess curves (local polynomial regression fitting).*

As discussed previously (Section 4.3.4), Loriaux et al. (2013) found that the super-CC scaling results from convective precipitation. Since convective precipitation is the dominant process during summer, the CC relation is also assessed for the summer (JJA) and winter (DJF) season separately (Fig. 4.17). In winter, our regions are typically affected by stratiform type precipitation associated with low intensity precipitation over a large area for several hours up to a day, while in summer short, intense convective precipitation events with a showery character, occurring on small spatial scales are more common (Berg et al., 2009). Hence, a distinction in the response between both seasons will possibly give us an indication of the different mechanisms contributing in summer and winter to the precipitation-temperature relationships.

Results for the winter period (DJF) show for the observations a dependency of the hourly precipitation extremes along the CC line for the whole temperature range (Fig. 4.17(b)). In contrast, no CC relationship is found in the hourly DJF precipitation extremes for ERAINT-ALR04 (Fig. 4.17(d)). Figure 4.17(a) shows that the observations during the summer season (JJA) are characterized by a pronounced super-CC scaling for temperatures above $\approx 18°$C for the 90th and 95th percentiles, and for almost the whole temperature range between 10 and $22°$C for the highest percentiles (99th and 99.9^{th}). Except for the 99th and 99.9^{th} percentiles, this super-CC scaling is not captured by the model, and even shows in the tail of the 90th-99th percentiles a dependency of hourly precipitation extremes below the CC relation. This decrease in slope above $\approx 22°$C is also visible in the modeled percentiles for all seasons (Fig. 4.14(b) and Fig. 4.14(d)), but appears even more pronounced for the summer percentiles.

Berg et al. (2013a) studied the CC relation for observations in Germany, and separated stratiform and convective precipitation events by cloud observations. For temperatures above $22°$C, their distributions for convective and total precipitation also display a decreas-

ing slope, and suggests it to be a stable high-temperature feature of convection (Berg et al., 2013a). Similarly, Lenderink and van Meijgaard (2010) found a strong reduction in modeled precipitation intensity for temperatures above 22°C. The authors attribute this fall off in intensity to model errors, but also note that for temperatures above ≈ 22°C rather anomalous atmospheric conditions (e.g. with severe soil drying and/or strong high pressure systems) could also suppress the occurrence and intensity of precipitation extremes (Lenderink and van Meijgaard, 2010).
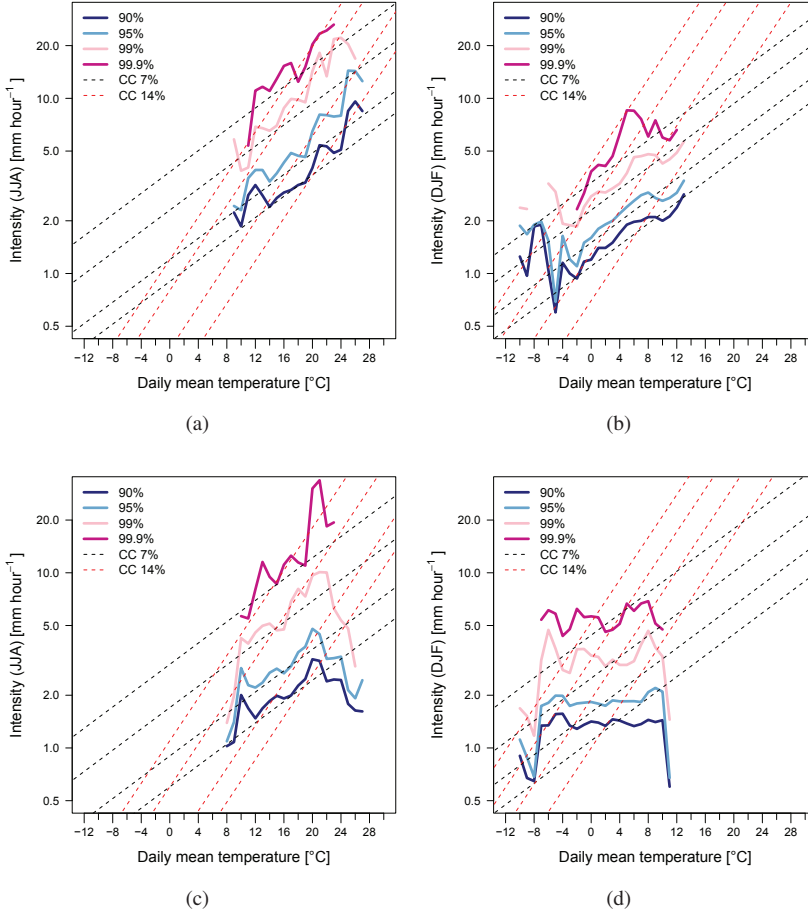
Figure 4.17: *Dependency of hourly precipitation extremes on temperature computed from observed data in Uccle of (a) the summer season JJA, and (b) the winter season DJF, and from model grid box data closest to Uccle of (c) the summer season JJA, and (d) the winter season DJF. Lines and axes are similar to Fig. 4.14.*

To explore more in detail the reason for this negative scaling in ERAINT-ALR04 that occurs at temperatures above approximately 20°C, we have plotted similar to Hardwick Jones et al. (2010), the daily mean relative humidity against daily mean temperature for the closest model grid point to Uccle. The relative humidity is plotted for each wet day (i.e.

daily precipitation > 0.1 mm day^{-1}), and separately for winter/spring and summer/autumn. Although there is a large variability in relative humidity values, in accordance to the results of Hardwick Jones et al. (2010), there seems to be an overall decrease in relative humidity with increasing temperature (Fig. 4.18). It also appears that the temperature at which the relative humidity values start to decline, coincides with the temperature value where the scaling of the hourly precipitation amounts become negative (i.e. $\approx 20°$C for June-July-August (JJA) + September-October-November (SON) and $\approx 10°$C for December-January-February (DJF) + March-April-May (MAM). Hence, as suggested by Hardwick Jones et al. (2010), this decrease in relative humidity can explain the negative scaling of extreme precipitation at high temperatures, and highlights the importance of both the atmosphere's capacity to hold moisture as well as the availability of moisture in the atmosphere.
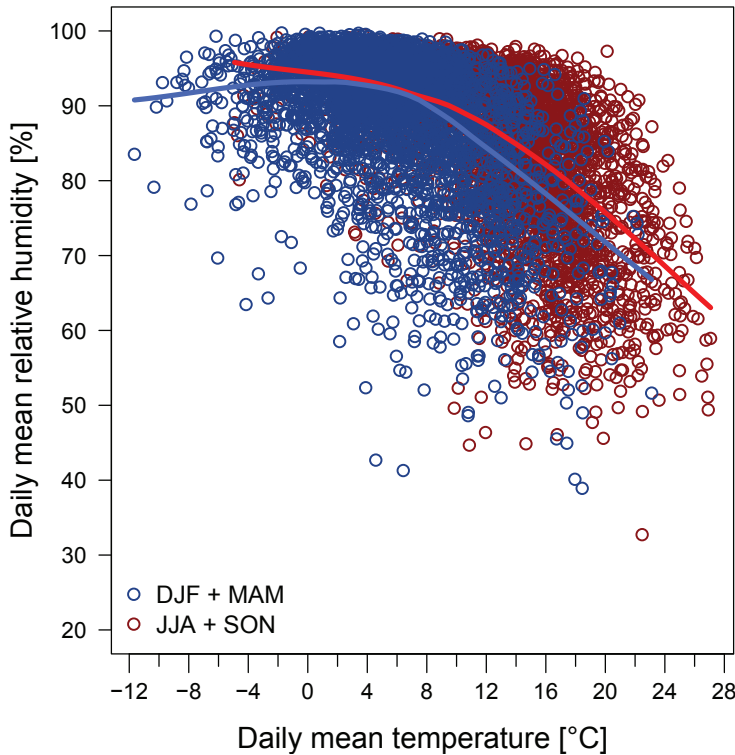
Figure 4.18: *Modeled daily mean relative humidity versus modeled daily mean temperature on wet days for the closest model grid box value to the observation station of Uccle. The values are plotted separately for winter/spring (December-January-February (DJF) + March-April-May (MAM), blue) and summer/autumn (June-July-August (JJA) + September-October-November (SON), red). The thick solid lines are smoothed loess curves (local polynomial regression fitting).*

Furthermore, this result indicates that dew point temperature instead of temperature could be used as an alternative measure to assess the dependency of extreme rainfall to temperature changes. This can be physically justified as follows. Dew point temperature is a direct measure of absolute specific humidity of the atmosphere. The difference between

temperature and dew point temperature (i.e. the dew point depression), is equivalent to the relative humidity. Thus assuming a constant relative humidity, an increase in temperature implies an equal increase in dew point temperature. This is reasonable for many areas and seasons, except for large continental areas in summer (Attema et al., 2014; Westra et al., 2014). Our finding showing that the relative humidity is a limiting factor for the highest temperature range, indeed suggests that the actual moisture (i.e. dew point temperature) in the atmosphere would be a better predictor of rainfall intensity than temperature itself. Lenderink and van Meijgaard (2010) for example, obtained from observations a more robust scaling, and a wider range where the CC scaling is valid when using local dew point temperature instead of temperature.

The lack of any clear relationship between the modeled hourly DJF precipitation and daily mean temperature is in contrast with the findings from Mishra et al. (2012). Based upon observational data for the period 1950-2009 across Contiguous United States, the authors found for both the summer and winter season a positive relationship between extreme precipitation intensity and temperature. On the other hand, Berg et al. (2009) studied for different regions in Europe the CC relationship between *daily* precipitation and daily temperature in three ERA-40 driven RCMs. The authors considered different seasons and months separately, and found different relationships between extreme precipitation and temperature during the summer and winter seasons, with a general increase in winter and a decrease in summer. Similar to Berg et al. (2009), we have calculated precipitation percentiles for observed and modeled *daily* precipitation. For both the observations and ERAINT-ALR04, the CC relation between extreme daily precipitation and temperature is less clear compared to the relationships obtained with the hourly intensities (Figs. 4.19(a) to 4.19(b)). In agreement with results from Lenderink and van Meijgaard (2008), the CC relation is approximately found for temperatures below 8-10°C, but a sub-CC scaling is visible for the higher temperatures. However, if the relationships between extreme daily precipitation and temperature are calculated for the months separately, we find a similar behavior than Berg et al. (2009): an increase in precipitation that nearly follows the CC scaling during the winter months (Figs. 4.19(c) to 4.19(d)), and a decrease during the summer months (Figs. 4.19(e) to 4.19(f)). This seasonal variability in the relationships may be attributable to the type and scale of precipitation, that is often influenced by the synoptic scale circulation pattern (Mishra et al., 2012).

Previous results are all obtained for the model grid box values closest to the observation station of Uccle. Finally, to check the sensitivity of the results when comparing point measurements and model grid averages from only one grid point, the CC scaling is also assessed for a region of 5 × 5 (25) model grid points surrounding the closest model grid point to Uccle. Based upon model data from this subregion of 25 grid points, all results remain the same and reveal similar findings as found for the closest grid box values to the observation station of Uccle (see Appendix B, Figs. B.4 to B.6).
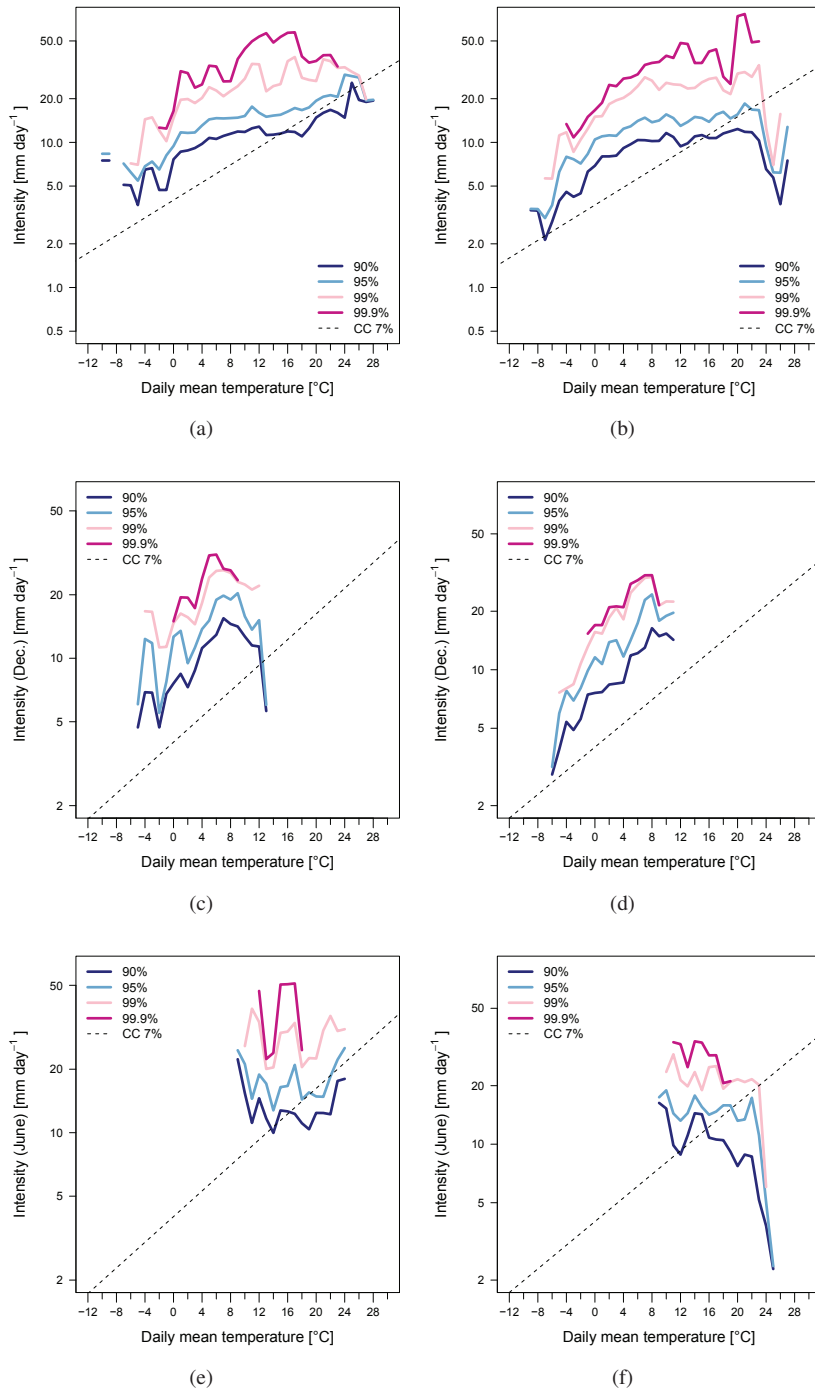
Figure 4.19: *Dependency of daily precipitation extremes on temperature computed from observed data in Uccle of (a) all months, (c) December, and (e) June, and from model grid box data closest to Uccle of (b) all months, (d) December, and (f) June. Lines and axes are similar to Fig. 4.14.*

4.5   Conclusion: the added value

This chapter assessed the added value of the ALARO-0 model in the simulation of subdaily precipitation at different spatial resolutions. Hence, we have addressed different aspects of the characteristics of (extreme) precipitation at the subdaily timescale, ranging between 1-hour aggregation levels up to the daily timescale (24-hour aggregation level).

The results for the diurnal cycle of both high-resolution simulations (ALR04 and ALR10) showed w.r.t. the observations an improvement in the onset and peak of convective activity. The 40-km run on the other hand, is not able to reproduce the observed diurnal cycle both in magnitude and phase. Nevertheless, the improvements for ALR04 and ALR10, there is still a time shift towards a too early onset of the convective build-up notable. Results from NWP demonstrated that recent changes in 3MT in ALARO-0, as well as in the new model version ALARO-1, significantly reduced (or even canceled) this time shift. These most recent changes are not implemented in the ALARO-0 version that is used for the simulations in this thesis. Hence, in the future with the use of the new ALARO-1 model version for climate applications, we could expect that this time shift in the diurnal cycle will be reduced or disappeared.

Furthermore, frequencies and intensities were calculated for observed and modeled hourly precipitation amounts. The results demonstrated that although the ALR04 and ALR10 models still suffer from common issues in climate models, we could conclude that with respect to the low-resolution ALR40 model, the simulation of high hourly precipitation amounts is consistently improved by the high-resolution ALARO-0 runs.

The rainfall features, where to a first approximation, the IDF curves display a power law dependence on averaging duration d and return period T, has been generally studied with observation data (e.g. Burlando and Rosso, 1996; Willems, 2000; Bougadis and Adamowski, 2006). We have demonstrated in this chapter, that the ALARO-0 model is also able to reproduce this rainfall feature, which is of particular interest for extremes. More specifically, the high-resolution ERA40-ALR04 simulation shows for all durations a power law dependency that approaches very closely the observed power law. In contrast, the low-resolution 40-km run does not reproduce the observed scaling properties for the lowest durations (1 to 3 hour aggregation times). Only for the 24-h durations both the high-resolution as well as the low, 40-km resolution runs display an IDF relationship based upon this power law which is in agreement with the observations.

Finally, another scaling property of extreme precipitation that expresses the dependency of hourly precipitation extremes on temperature by the CC relation, is also relatively well reproduced by the model at a spatial resolution of 4 km. In particular for the most extreme events (i.e. highest percentiles) and for a rather large temperature range, the model is able to reproduce the (super-) CC relation. However, the model seems to have difficulties in the representation of less extreme events (i.e. lower percentiles) and events in the highest temperature range as suggested by the fall off in precipitation intensity. In the light of a changing climate, often characterized by an increase in mean surface temperature, our results for the CC relation motivate further investigations to assess how the relation between precipitation intensity and temperature will change in a future climate (see Chapter 6).

All this suggest that the multiscale performance of ALARO-0 in the simulation of daily summer precipitation, as was found in previous chapter, does not hold for the simulation of

subdaily precipitation. Luca et al. (2012) determined by means of Potential Added Value (PAV) whether RCMs add value in the representation of climate statistics compared to the driving GCM data. Their results for precipitation indeed show that the PAV of RCMs is much higher for short temporal scales than for long temporal scales, due to the filtering resulting from the time averaging process. Our results have demonstrated that the highest-resolution simulations of ALARO-0 at 4 km indeed benefit from added value in the description of several characteristics of subdaily precipitation, such as the diurnal cycle, heavy precipitation amounts, and important scaling properties. This finding is in line with previous studies that explored w.r.t. observations the added value of low-resolution RCMs where deep convection is parameterized versus high-resolution CPMs where the convection parameterizations are (partly) switched off. Furthermore, the multiscale character of the physics package 3MT for clouds and deep convection allowed us to use consistently the same model physics at a range of different spatial resolutions, which is an important strength in our assessment w.r.t. to other studies where one is often obliged to use a different model physics or treatment of deep convection once the horizontal resolution is changed.

Hence, the results of the evaluation of extreme precipitation at the daily and subdaily timescale, as extensively discussed in the previous chapter (Chapter 3) and this chapter, demonstrate that the ALARO-0 model is able to consistently capture the relevant precipitation characteristics at a wide range of atmospheric and corresponding temporal scales, varying from the micro- to the mesoscales (Fig. 1.1). These highly promising and encouraging findings thus increase our confidence to use the ALARO-0 model at a spatial resolution of 4 km to compute Intergovernmental Panel on Climate Change (IPCC) scenarios for climate change studies. This will be the subject of the next chapters (Chapter 5 and Chapter 6).

5

FUTURE CHANGES OF EXTREME PRECIPITATION IN BELGIUM UNDER THE A1B SCENARIO DERIVED FROM THE ALARO-0 MODEL AT 4 KM RESOLUTION

"Even though European summers are anticipated to become drier on average, severe precipitation events are predicted by many models to occur more frequently and with higher intensities."
– Brockhaus et al. (2008)

5.1 Introduction

In the previous chapters (Chapter 3 and Chapter 4), extreme precipitation from the ALARO-0 model at several spatial and temporal scales has been extensively validated. In the context of climate change assessment, potential future changes in extreme precipitation are of great importance. In particular because such precipitation extremes are related to e.g. floods and landslides, which have a great impact on many aspects of human society: health, natural and urban environments, buildings and infrastructure, and economy (Beniston et al., 2007; Kyselý and Beranová, 2009; Nikulin et al., 2011). Hence, the future response of extreme precipitation over Belgium to increased greenhouse gas (GHG) concentrations, as obtained from the ALARO-0 model, will be the subject of the current chapter.

In general, the amount and intensity of precipitation is determined to a large extent by the available energy and moisture. Although global mean precipitation is primarily constrained by the energy budget, extreme rainfall events are likely to occur when effectively all the moisture (or a constant fraction thereof) in a volume of air is precipitated out (Allen and Ingram, 2002). This physical consideration suggests that the intensity of these events is determined by the moisture availability, which will generally rise in a warming climate. As discussed in the previous chapter (Chapter 4, Section 4.3.4), changes in atmospheric moisture roughly follow the saturation specific humidity, which in turn is governed by the Clausius-Clapeyron (hereafter CC) relation. Hence, it has been suggested that the moistening of the atmosphere could lead to an increase in the frequency and intensity of extreme precipitation events following the same CC rate of increase of 6-7% per degree (Allen and Ingram, 2002; Frei et al., 2006; Attema et al., 2014).

Consistent with these conceptual considerations, significant recent observational and modeling work also points to an increase of precipitation extremes in many regions in the world. Studies of observed data show an increase of extreme daily rainfall intensity and/or frequency over most continents, and for more than half of the areas where reliable data are available positive trends for annual maximum precipitation extremes over 1951-1999 are detected (Westra et al., 2014). Many climate change modeling studies have used coupled ocean-atmosphere Global Climate Model (GCM) simulations forced with projected GHG and aerosol emissions as a primary tool for studying possible future changes in mean climate, variability and extremes (Russo and Sterl, 2012). These global climate change experiments generally project an increase in frequency and intensity of precipitation extremes almost everywhere over the world, even for regions with a decrease in mean precipitation (Nikulin et al., 2011). At the regional scale, potential future changes in precipitation patterns and magnitudes are usually assessed by means of nested Regional Climate Models (RCMs) simulations, as these allow to calculate climate change scenarios consistent with the driving GCM but with more regional detail and a better representation of processes leading to heavy precipitation (Kyselý and Beranová, 2009). For Europe, numerous studies have analyzed

future changes in daily precipitation extremes from climate change simulations with RCMs and report a distinct seasonal response in extreme precipitation. In winter, models reveal an increase in extreme precipitation in central and northern Europe, and show smaller changes with a tendency to a decrease to the south. In contrast, in summer a gradual pattern is found with increases in northeastern Europe and Scandinavia and decreases in southern Europe and the Mediterranean region (e.g. Frei et al., 2006; Beniston et al., 2007; Buonomo et al., 2007; May, 2008; Fowler and Ekström, 2009; Kyselý and Beranová, 2009; Nikulin et al., 2011). Compared to the projections for summer, the wintertime projections of heavy precipitation based upon different RCMs driven by one GCM (e.g. Frei et al., 2006; Kyselý and Beranová, 2009) and based upon one RCM driven by different GCMs (e.g. Nikulin et al., 2011), are found to be much more robust. The projected changes of extreme summer precipitation show much more varying patterns and less agreement among the models, which may be explained by the strong dependency of the physical parameterization formulation during this season (Frei et al., 2006).

Furthermore, this contrasting seasonal pattern and consistency in European precipitation extremes is also reflected in changes in annual and seasonal mean precipitation. Jacob et al. (2014) for example, assessed changes in mean precipitation for Europe from the first set of EURO-CORDEX simulations at 12.5 km resolution for the new Representative Concentration Pathway scenarios (RCP) RCP4.5 and RCP8.5 (EURO-CORDEX, Coordinated Regional Climate Downscaling Experiment for Europe). The study also compares the EURO-CORDEX ensemble data to the Special Report on Emission Scenarios (SRES) A1B scenario results achieved within the ENSEMBLES project. Overall, for all three scenarios the ensemble mean projects a statistically significant increase in annual mean precipitation in large parts of central and northern Europe, and a decrease in southern Europe. For the transition zone located in central Europe where the climate change signals change in sign, climate models show small and mostly non-significant changes as well as large disagreements concerning the tendency of the change (Feldmann et al., 2012; Jacob et al., 2014). Seasonal changes of mean precipitation are similar to the annual mean pattern, with exception of this transition zone which shifts southwards in summer and northwards in winter (Jacob et al., 2014).

Before a RCM can be used to project future changes, it is important to evaluate the performance of the model in simulating past observed changes. Such an evaluation helps to decide whether or not we can trust a model response in a sensitivity experiment (Déqué, 2007). As it is often done in model evaluation studies, the regional model is driven by so-called perfect boundary conditions provided by reanalyses (Maraun et al., 2010). Previous chapters (Chapter 3 and Chapter 4) have indeed evaluated present-day daily and subdaily precipitation from the downscaling ALARO-0 model driven by reanalysis data. We have demonstrated that the ALARO-0 model, and in particular at the highest resolution of 4 km, is able to realistically simulate several characteristics of the daily and subdaily extreme precipitation climatology.

This chapter aims to derive potential future changes in extreme precipitation for Belgium from downscaled GCM data using the ALARO-0 model at 4 km resolution. In a first step, it is verified whether or not the observed climatology for the control period is correctly represented by the ALARO-0 model, which is important if one couples a RCM to a GCM. Whatever the model, it is a simplification of nature. Hence, it is not surprising that model

results do not fit exactly the observations. Such imperfections or model biases are usually postprocessed with bias correction methods (Déqué, 2007). However, we limit ourselves in this chapter to a sensitivity study (i.e. the impact of a warming on extreme precipitation), the use and relevance of different bias correction techniques will be briefly highlighted at the end of the chapter.

Future changes for Belgium are first assessed for mean precipitation both at daily and subdaily timescales. Climate model estimations of mean precipitation under several future climate scenarios for central Belgium have previously also been examined by Baguis et al. (2010). The authors based their assessment on GCM and RCM data from the 4th Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR4) and the European project PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects), respectively. In line with the general pattern found by many other modeling studies, their analyses of changes in monthly and seasonal means of RCM simulations demonstrated a clear shift in the precipitation pattern with an increase during winter and a decrease during summer. However, it has been found, that the signal may be less clear or different when another set of scenarios from the GCM simulations are taken into account (Baguis et al., 2010).

To extend the previous study, a great part of our analysis is attributed to changes in extreme precipitation at subdaily timescales. Furthermore, most of the previous modeling studies used a relatively coarse resolution (i.e., about 50-10 km). The ongoing EURO-CORDEX initiative provides regional climate projections for Europe up to resolutions of 12.5 km, and thereby complements the existing coarser resolution datasets of former coordinated activities like, e.g., PRUDENCE and ENSEMBLES (See also the General introduction Chapter 1 and Jacob et al., 2014). However, if one wants to study future climate changes at the regional scales, and provide relevant information for e.g. hydrological impact studies, higher spatial and temporal resolutions are needed (Willems et al., 2012). In this respect, the derived projected changes at hourly timescales and at high resolutions of 4 km offer a great benefit w.r.t. many other regional climate studies, as it may provide valuable information to hydrological applications and impact studies in Belgium.

The outline of this chapter is as follows. The next section (Section 5.2) provides a description of the model data and observations, as well as the experimental design. The methods and statistical procedures for significance testing are presented in detail in Section 5.3. Results on the evaluation and changes in mean and extreme precipitation are discussed in Section 5.4.

Finally, it should be kept in mind that model projections of future climate change encompass a wide range of uncertainties (Kendon et al., 2010). One source of uncertainty is associated with modeling deficiencies. For example, only a few studies have investigated changes in subdaily rainfall, mainly due to the difficulty of climate models to correctly simulate rainfall at such high temporal resolution (Kendon et al., 2014). Related to this, it is well-known that the uncertainties in the precipitation results of climate models are an order of magnitude higher in comparison with the climate model outputs of other variables such as e.g. temperature (Willems, 2011). To assess reliability in the context of modeling uncertainty, a mechanistic approach is commonly used (e.g. Rowell and Jones, 2006; Kendon et al., 2010). This approach aims to gain a better understanding of the underlying mecha-

nisms of the regional climate change. By attributing the changes to specific mechanisms, each of which is associated with a different level of confidence, a subjective judgement can be made of the overall reliability of the projected changes (Kendon et al., 2010). Furthermore, it should be taken into account that our climate projection represents just one possible realization of the future climate, solely based on downscaling results from one RCM, driven by the outputs of one GCM under one scenario of natural and anthropogenic forcing (Pan et al., 2011). Since natural variability of the climate limits the precision of future climate projections, it is thus expected that the use of a different RCM, GCM and/or emission scenario might give different results (Kendon et al., 2010; Pan et al., 2011). In addition, climate models are found to show large disagreements concerning the tendency of precipitation in the transition zone where our study region (i.e. Belgium) is located. Multi-model ensemble systems are commonly used to assess the uncertainties related to natural variability (Frei et al., 2006). Therefore, at the end of Section 5.4.4 results from a recent multi-model intercomparison study, including our ALARO-0 projection, are discussed. We end this chapter with the conclusions in Section 5.5.

5.2 Data and experimental design

5.2.1 Model data

To assess the possible changes in future precipitation, two climate projection simulations with the ALARO-0 model driven by the coupled GCM CNRM-CM3 (further in the text and figures also denoted as CNRM-CM3), are considered. Each of the simulations cover a 30-yr period: (i) 1961-1990 as period representative of the historical or control climate (hereafter denoted by CTL), and (ii) 2071-2100 as scenario period representing the future climate (hereafter denoted by SCN).

CNRM-CM3 is developed by the Centre National de Recherches Météorologi-ques (CNRM). The model was applied to carry out several climate simulation in the framework of the Coupled Model Intercomparison Project Phase 3 (CMIP3), which consists of a large ensemble of GCMs for which the outcomes have been synthetisized in IPCC AR4 (Salas-Mélia et al., 2005; Meehl et al., 2007). CNRM-CM3 is based on the coupling of the ARPEGE-Climate GCM version 3, which is the atmospheric part of the system, OPA8.1 is the ocean model, GELATO2 the sea ice model, and TRIP is used for the river routing scheme. ARPEGE-Climate uses a spectral representation for most variables, and was used on a horizontal grid corresponding to a T63 triangular truncation with 45 vertical layers (Salas-Mélia et al., 2005).

Similar to the reanalysis driven simulations (see Chapter 3 and Chapter 4), the experimental design consists of a dynamical downscaling of the GCM data with daily reinitializations using two intermediate nestings. First, the GCM data are downscaled using the ALARO-0 model at 40-km resolution corresponding to a 69×69 grid point domain that encompasses most of Western Europe. Subsequently, these 40-km output are used as input for a one-way nesting on a 181×181 4-km resolution domain centered on Belgium. More details on our approach for daily reinitializations are given in Chapter 2, Section 2.3.2 and Chapter 3, Section 3.2.1.

A realistic approach is to consider the transient climate response (TCR) associated

with a transient increase based on expected emission scenarios. IPCC has defined "emission scenarios" for future changes in the GHG concentrations which are dependent on different economic evolutions as well as on evolutions in decision making and policy (De Troch et al., 2014b).
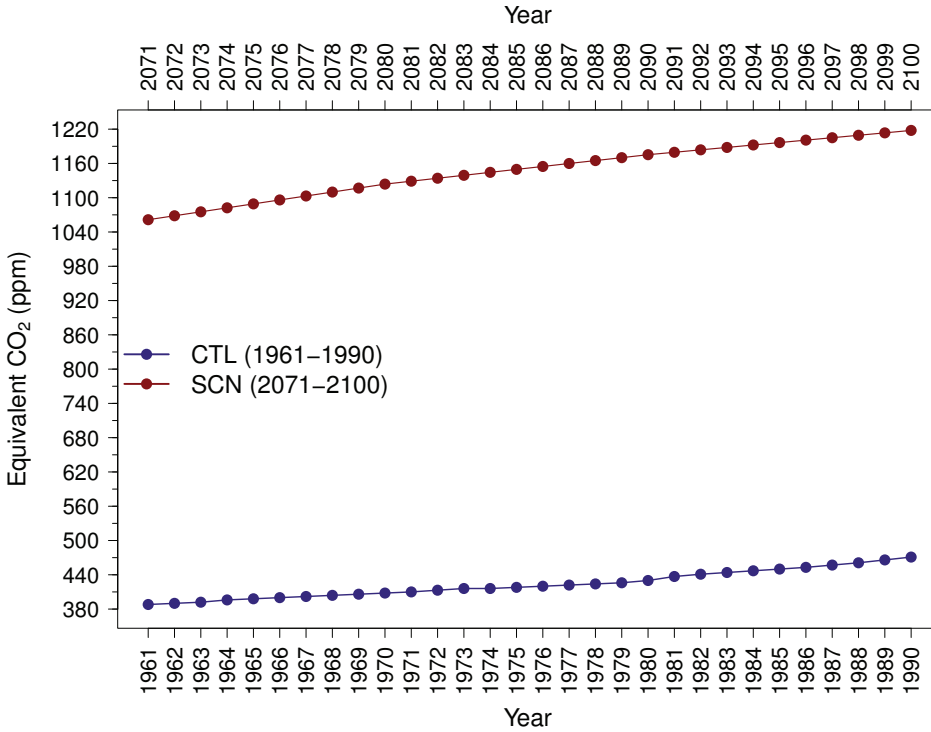
Figure 5.1: *Values for equivalent CO_2 (expressed in parts per million, ppm) corresponding to the Special Report on Emission Scenarios (SRES) A1B scenario, as used for the control (CTL, 1961-1990) and the scenario (SCN, 2071-2100) simulation, respectively.*

Therefore, in order to quantify the sensitivity of the climate to external forcings, the CTL and SCN model integrations are performed with an equivalent GHG forcing corresponding to an equivalent CO_2 increase according to the IPCC SRES A1B scenario (Nakićenović et al., 2000). As mentioned previously in the General introduction (Chapter 1), the A1B scenario is the scenario that was mostly used within the ENSEMBLES project (Niehörster et al., 2008), and it is still the current scenario used at the Royal Meteorological Institute of Belgium (RMI).

Figure 5.1 shows the equivalent CO_2 evolution as described by the SRES A1B scenario, used for the CTL and SCN simulation respectively. This scenario follows the storyline of the IPCC family of A1 scenarios. It assumes a future world of very rapid economic growth and the rapid introduction of new and more efficient technologies, with a worldwide population peaking in the middle of the 21st century, and a balanced use across all energy sources. This scenario leads to a rapid increase in fossil CO_2 emissions until 2050 and a decrease afterwards. Compared to other SRES scenarios, the CO_2 emissions in the A1B scenario lie in the middle of the scenario range (Jacob et al., 2014).

In addition to the GCM driven model data, the ERA40-reanalysis driven model results (see Chapter 3) are used as a reference for validation of the hourly and daily precipitation model output for the control period (1961-1990).

5.2.2   Observations

The simulated results for daily precipitation are validated against the same observation dataset as used in Chapter 3, comprising 93 climatological stations with daily accumulated precipitation, selected from the climatological network of the RMI for the control period 1961-1990 (Fig. 5.2, right).
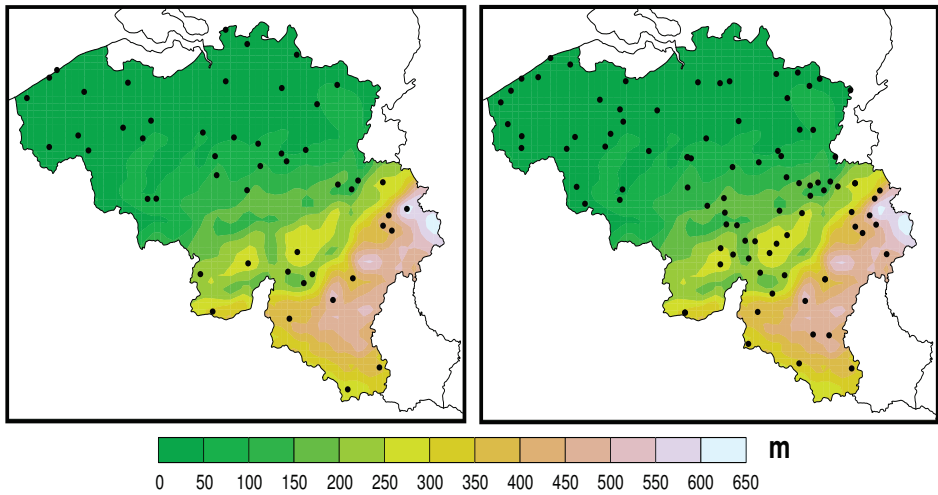
Figure 5.2: *Model topography (m) of Belgium showing the location of the 50 (left) and 93 (right) selected climatological stations (black dots) which are used for the validation of modeled daily temperature and precipitation, respectively.*

Furthermore, observations of daily mean temperature for 50 station locations in Belgium (Fig. 5.2, left) are selected from the climatological observation network of the RMI. These temperature observations are used to validate the modeled daily mean temperature for the control period, before the projected warming under the A1B scenario is analyzed. These daily mean temperatures are also used to assess the dependency of hourly precipitation with temperature change (i.e. CC relation). Finally, hourly observation data for Uccle are used, for which a more detailed description can be found in the previous Chapter 4.

5.3   Methods

5.3.1   Data processing and analysis

Validation of the model data (precipitation and temperature) against the observations is done by selecting the nearest grid box to the station points. The average lapse rate of 6.5 K km^{-1}

under standard atmospheric conditions is added to the modeled temperature values, according to the altitude difference between the model grid point and the observational station. In this way the modeled temperature values have the same altitude as the nearest observation series. For the assessment of the CC relation, observed and modeled daily mean temperature values are simply derived by averaging daily maximum- and minimum 2-meter temperature values. Observations of daily maximum (minimum) temperatures are measured between 8 and 8 Local Time (LT) of the day after (before). The modeled daily maximum- and minimum temperatures are calculated accordingly the definition of the observations.

As previous modeling studies have demonstrated fundamental different responses in precipitation changes for the winter and summer season, we limit our analysis to these seasons. Note that through the selection of the winter months December-January-February (DJF), the length of our CTL and SCN study periods is reduced to 29 winter seasons.

The analysis at the subdaily timescale is restricted to 1-hour durations only. In the previous chapter (Chapter 4), it has been demonstrated that observed hourly precipitation is fairly well reproduced by the 4-km ALARO-0 model, driven by the so-called "perfect boundary conditions" from ERA-40. Hence, to extend the validation of the hourly precipitation from the CTL simulation to more than only the model grid box value closest to the Uccle station, hourly model values from the ERA-40 downscaling are used as a gridded reference or "pseudo-observation" dataset at 4-km resolution. Furthermore, since the ERA-40 driven simulations were only done for the summer period 1961-90, the analysis of the 1-hour precipitation is limited to the June-July-August (JJA) summer months.

Two types of methodologies that are generally used for the assessment of changes in extreme precipitation are applied: (i) Extreme Value Analysis (EVA) which allows to estimate the change in terms of return level values giving an extreme value of a variable that will statistically be exceeded exactly once within a certain period (i.e. the return period) (Knote et al., 2010), and (ii) direct summaries of extremal behavior described by seasonal or annual maxima, quantiles, and threshold-based indices (Beniston et al., 2007). In the first method the return levels of daily and 1-hour precipitation are computed by fitting the parametric Generalized Extreme Value (GEV) distribution to a sample of summer- and winter annual maxima. More details on the fitting of the GEV distribution and calculation of return levels are given in the methodology section of the previous chapter (Chapter 4, Section 4.3.2). In addition, it is verified to what extent future changes in extreme hourly precipitation as simulated by our scenario are consistent with the expectations from the CC relation.

Finally, for the spatial analysis with gridded model data, an analysis grid of 81×81 grid points that coincide with Belgium is selected.

5.3.2   Significance testing

Two approaches are used to test the statistical significance of the differences between the CTL simulation and the observations or ERA-40 driven simulation, as well as between the SCN and CTL simulation: (i) significance testing based on the Kolmogorov-Smirnov (K-S) test, and (ii) significance testing based on bootstrap resampling. In both approaches the statistical significance is assessed at the 5% significance level.

When the significance in the biases and projected changes are assessed for all 81×81 grid points over Belgium, the K-S test is used because of computational efficiency. Here

it is tested whether or not the differences are drawn from the same distribution at a chosen level of significance (i.e. 5%). The K-S test is a nonparametric test (see also Chapter 3 and Chapter 4), having the advantage of making no assumptions about the distribution of the underlying data. For each model grid point where the null hypothesis, that both distributions of CTL and the ERA-40 driven reference data or SCN and CTL data are drawn from the same distribution, can be rejected at the 5% significance level, significant biases or changes are quantified (Russo and Sterl, 2012).

In case that only station points are considered, the significance of model biases and future changes w.r.t. year-to-year variability is assessed through bootstrap resampling (Kendon et al., 2014). This approach consists of producing 1000 bootstrap samples by selecting randomly with replacement 30 (29) JJA (DJF) seasons from the full dataset. For each station point, 1000 estimates of the mean differences between either CTL model data and observations, or either CTL and SCN model data, are calculated. These 1000 differences are then used to calculate for each station point lower and upper limits of the 95% confidence intervals for the difference or future change as respectively the 2.5^{th} and 97.5^{th} percentile. Differences or changes lying inside the respectively confidence interval are considered to be significant at the 5% significance level compared to year-to-year variability.

The significance of the EVA results is also tested through bootstrap resampling applied over each model grid point or station point, although, the approach is slightly different. Similar as in the previous chapter (Chapter 4), 95% confidence intervals around the GEV parameter estimates (i.e. location, scale, and shape parameter) are obtained from 1000 parametric bootstrap iterations. Differences between the GEV estimates of the CTL and the observations/ERA-40 simulation, as well as changes between the GEV estimates of the SCN and CTL simulation, are considered to be significant at the 5% significance level if the respectively 95% confidence intervals do not overlap. According to Eq. (4.5), the estimated GEV parameters are then used to calculate return levels x_T for different return periods T. Likewise, statistically significance of differences and changes in return level estimates are considered to be statistically significant at significance level 5%, if the 95% confidence intervals derived from 1000 bootstrap samples of the two separate datasets are not overlapping. The 1000 bootstrap samples of return level estimates are derived from 1000 GEV parameter estimates, which are obtained by fitting the GEV to 1000 bootstrap samples of size 30, randomly generated from the originally fitted GEV. Here, the 2.5^{th} and 97.5^{th} percentile of the resulting collection of 1000 return level estimates are also used as lower and upper 95% confidence bounds for the true T-yr return level. Again, the difference or future change between two return level estimates is said to be statistically significant if their 95% confidence intervals do not overlap, which corresponds to the 5% significance level (Kharin and Zwiers, 2000).

5.4   Results and discussion

5.4.1   Simulated mean precipitation and temperature in the control period

As mentioned in the Introduction (Section 5.1), if one couples a RCM to a GCM, it is important to first verify whether or not the climatology for the control period is correctly

represented by the regional model. Differences between model and observations help to decide whether one can trust a model response or not in a sensitivity experiment. One is more confident in a model response greater than the model bias than in a small model response (Déqué, 2007). However, validation of a GCM driven RCM is fundamentally different from a validation of a RCM driven by reanalysis data. In the former case, simulated and observed weather are independent. Therefore, as suggested by Maraun et al. (2010), the validation of our CTL simulation is limited to an evaluation of the spatial structure of the climatological fields and the frequency distributions.

Evaluation of the simulated daily temperature and (sub)daily precipitation for the summer (JJA, 1961-1990) and winter seasons (DJF, 1962-1990) is shown in the spatial distributions in Figs. 5.3 to 5.5 and frequency distributions in Figs. 5.6 to 5.8.
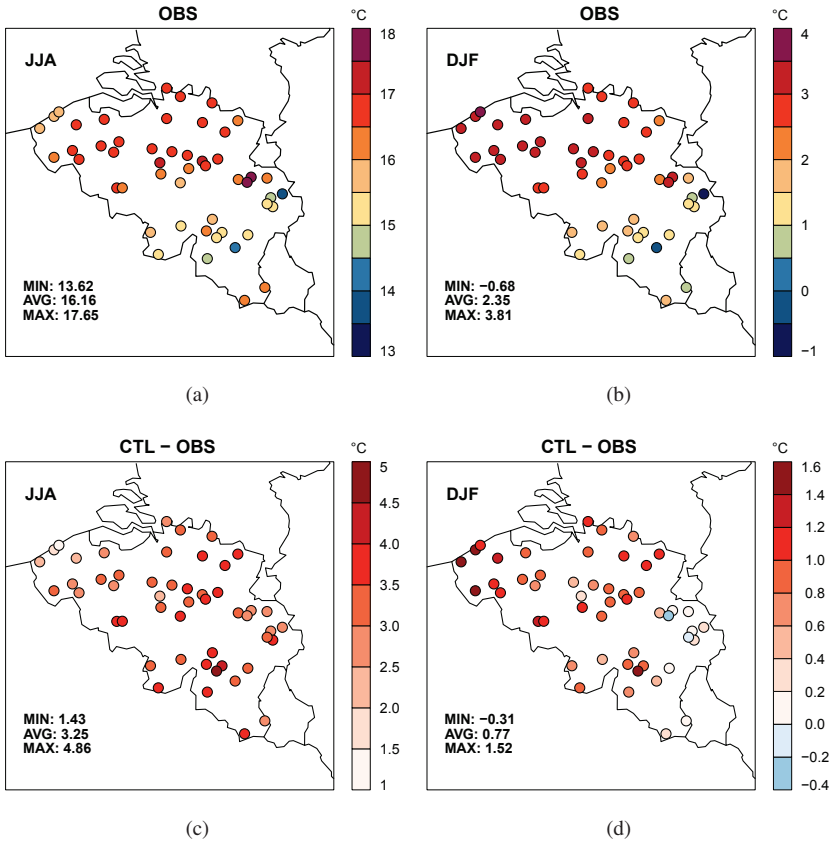
Figure 5.3: *30-yr mean daily summer (JJA, 1961-1990) and winter (DJF, 1962-1990) temperature (°C) estimated from (a-b) observation data (OBS) and (c-d) biases of daily summer and winter temperature (°C) estimated from the control (CTL) simulation with respect to the observations (OBS). For each station the significance of the biases between CTL and OBS is tested through bootstrap resampling. All biases are found to be significant at the 5% significance level compared to year-to-year variability. The numbers correspond to the minimum (MIN), average (AVG) and maximum (MAX) of the significant values over all station points.*

The observed daily temperature and precipitation fields show a topographical dependency, with a gradual decrease (increase) in temperature (precipitation) going from the northwest (low altitudes) to the southeast (high altitudes) of the country (Fig. 5.3(a), Fig. 5.3(b), Fig. 5.4(a)and Fig. 5.4(b)). A common feature, evident for both summer and winter daily temperatures simulated during the control period, is a systematic and significant warm bias over the whole country (Fig. 5.3(c) and Fig. 5.3(d)). During summer the overestimation of daily temperature by CTL is largest (up to 4°C), while, for winter the overestimation is smaller.

In contrast to daily temperature, significant biases of daily precipitation show an opposite pattern for summer and winter. Simulated precipitation during summer are systematically underestimated (up to more $\approx$ 50%), while daily winter precipitation is strongly overestimated (up to $\approx$ 80%) during the control period (Figs. 5.4(c) to 5.4(d)).
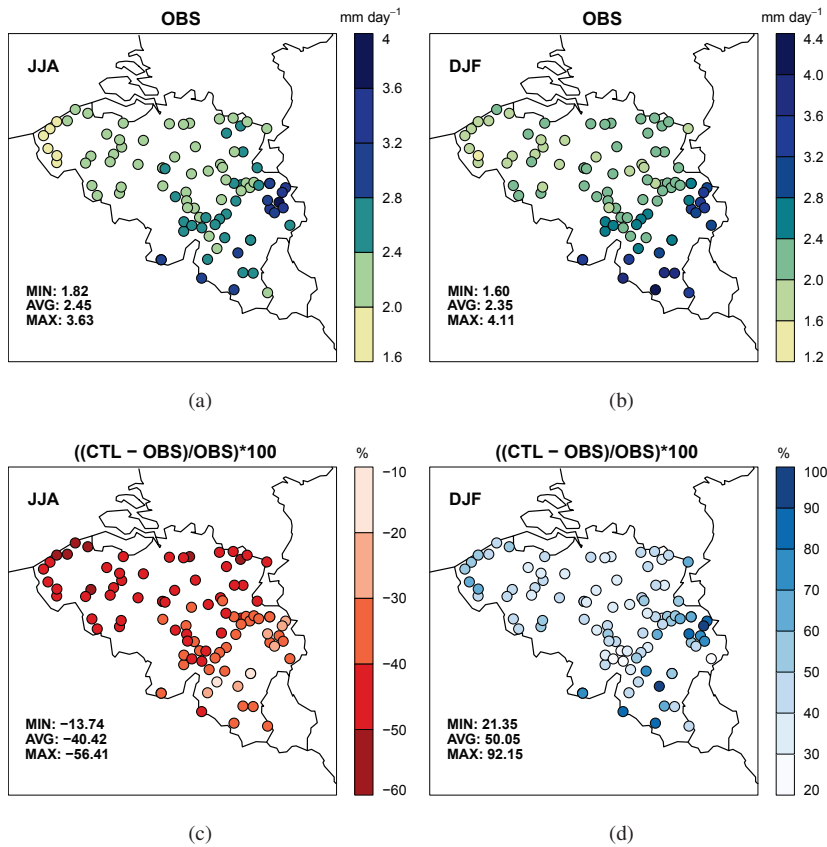
Figure 5.4: *As in Fig. 5.3, but for daily precipitation (mm day^{-1}).*

The spatial distribution of the 30-yr mean hourly precipitation during summer of the ERA-40 driven ALARO-0 simulation is displayed in Fig. 5.5(a). The subdaily precipitation shows also an orographical dependency, with higher precipitation rainfall amounts at the

higher altitudes in Belgium. Similar to the evaluation of daily precipitation, the hourly precipitation amounts are significantly underestimated in the control simulation w.r.t. the ERA-40 driven ALARO-0 simulation.

(a)   (b)
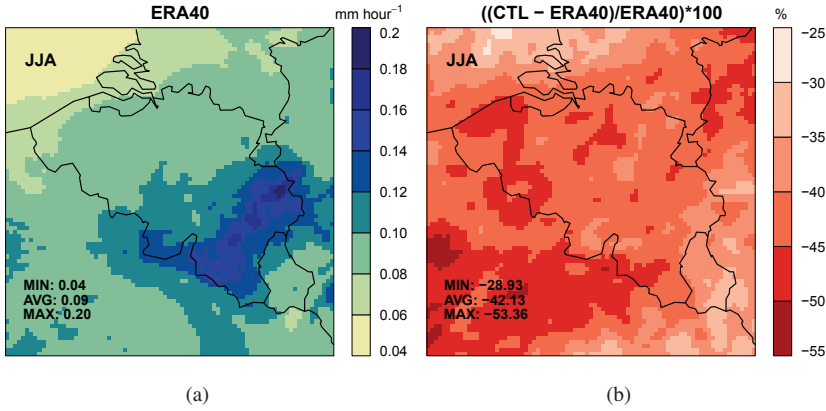
Figure 5.5: *30-yr mean hourly summer precipitation (JJA) for 1961-1990 (mm hour^{-1}) estimated from (a) the ERA40-driven simulation (ERA40) and (b) relative biases of hourly summer precipitation (JJA) for 1961-1990 (%) estimated from the control (CTL) simulation with respect to ERA40. For each grid point the significance of the biases between CTL and OBS is tested with the Kolmogorov-Smirnov (K-S) test. All biases are found to be significant at the 5% significance level. The numbers correspond to the minimum (MIN), average (AVG) and maximum (MAX) of the significant values over all 81×81 grid points.*

The differences between modeled and observed temperature and precipitation as have been identified in the spatial distributions, are also apparent in the frequency distributions. Frequencies are obtained by binning daily temperature, daily precipitation, and 1-hour precipitation values into bins of 1°C, 1 mm day^{-1}, and 1 mm hour^{-1}, respectively. Frequencies of 1-hour summer precipitation values in Uccle are calculated from 30-yr values which are binned into bins of 0.1 mm hour^{-1}. For the precipitation frequency distributions, a logarithmic scale has been used for a better representation of the extreme values.

Figure 5.6 shows the frequency distributions of daily summer- and winter 2-meter temperature for the observations and CTL simulation. Compared to the observations, both the summer and winter frequency distribution for the CTL simulation is systematically shifted towards higher temperatures. Although, the shift is clearly smaller for the winter than for the summer. As can be seen from the relative differences in frequencies between the model and observations, CTL has the tendency of producing too few (much) temperatures in the lower (higher) end of the temperature range. The frequency distribution of daily precipitation (Fig. 5.7) is clearly more variable than the one of temperature (Fig. 5.6). Precipitation is known to be a much more variable field than temperature. This is reflected by the presence of outliers in the frequency distributions, i.e. rainfall events with very high precipitation amounts. Given the rareness of these events, we limit the discussion of the results on the precipitation frequencies to the precipitation range corresponding to the more frequent events (i.e. ≈ 60 mm day^{-1}).

(a)   (b)
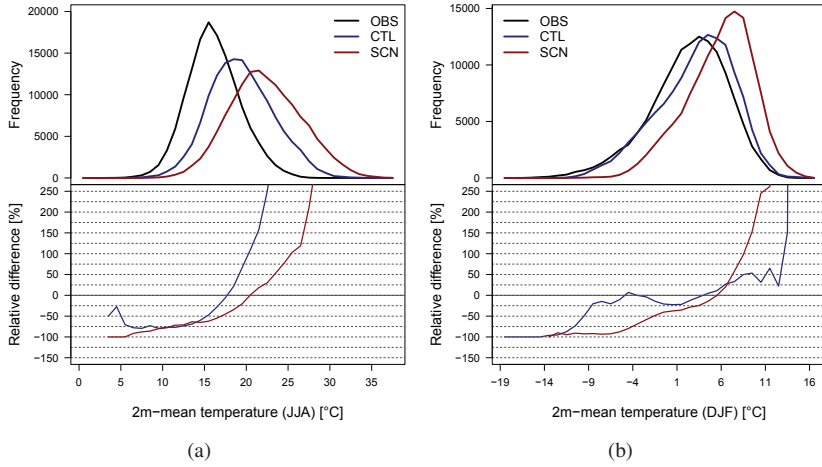
Figure 5.6: *Frequencies of the observations (OBS, black), the control simulation (CTL, blue), and the scenario simulation (SCN, red). Frequencies are computed with the daily temperature (°C) in (a) summer (JJA, 1961-1990 and 2071-2100) and (b) winter (DJF, 1962-1990 and 2072-2100) given for each of the 50 climatological stations and their closest model grid points separately (Fig. 5.2). Relative biases and projected changes between CTL and OBS (blue), and SCN and CTL (red) are shown in the bottom of the figures.*

(a)   (b)
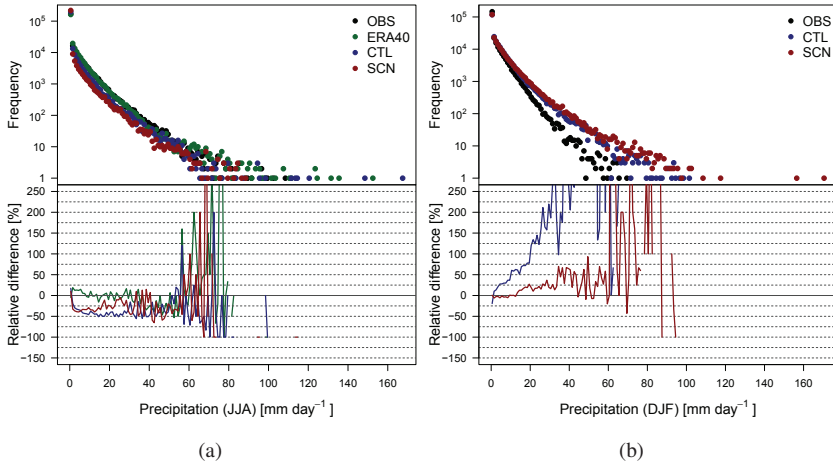
Figure 5.7: *As in Fig. 5.6, but for daily precipitation (mm day^{-1}) given for each of the 93 climatological stations and their closest model grid points separately (Fig. 5.2). The frequency distribution in summer for the ERA-40 driven simulation (ERA40, green) and its relative biases with respect to observations (OBS) are also shown. Frequencies are displayed on a logarithmic scale.*

The relative differences in frequencies of daily precipitation again reveal an opposite bias for summer and winter, characterized by a systematic underestimation of frequencies

during summer and a strong overestimation during winter. As a reference, the frequencies of daily precipitation in summer from the ERA-40 driven simulation during the control period are also calculated (Fig. 5.7(a)). Two features are apparent: (i) the number of dry days and low precipitation amounts are slightly overestimated by ERA40, and (ii) the highest precipitation values between 40 and 60 mm day^{-1} values are not very well captured by the model. However, these are known model deficiencies and compared to CTL, the reanalysis driven rainfall frequencies are much better represented and approach satisfactory well the observations.

The frequency distribution of 1-hour summer precipitation is given in Fig. 5.8. For the calculation of the relative bias in frequencies of 1-hour summer precipitation from the control simulation for all 81 × 81 grid points coinciding with Belgium (Fig. 5.8(a)) we have used ERA40 as a reference. The relative biases for the closest model grid point to Uccle on the other hand are calculated w.r.t. the observations (Fig. 5.8(b)). The control simulation shows, both w.r.t. ERA40 as well as w.r.t. the observations, a consistent underestimation of 1-hour precipitation amounts, suggesting that the underestimation is a robust feature present in CTL. Only for the lowest precipitation values (i.e. between 0-1 mm hour^{-1}), the underestimation by CTL could be attributed to a wet bias in ERA40, which produces too much dry hours w.r.t. the observations in Uccle.

(a)   (b)
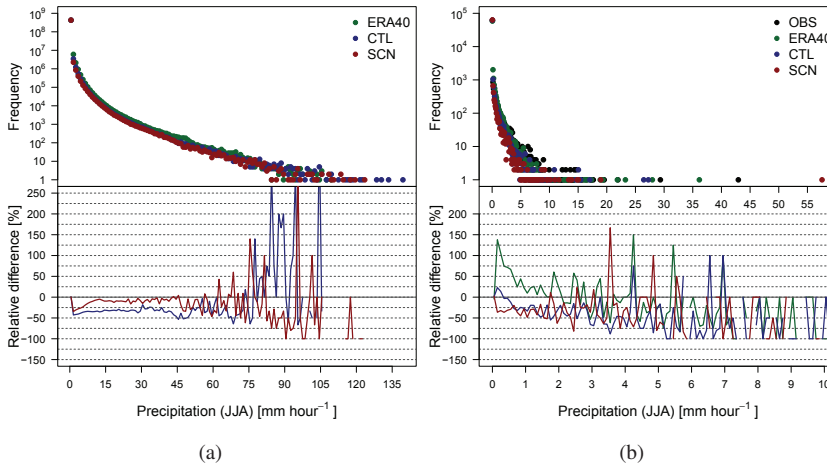
Figure 5.8: *Frequency distribution of hourly precipitation (mm hour^{-1}) in summer (JJA, 1961-1990 and 2071-2100) for (a) all 81 × 81 subgrid points for the ERA-40 driven simulation (ERA40, green), control simulation (CTL, blue), and scenario simulation (SCN, red), and (b) the closest model grid point to Uccle for observations (OBS, black), the ERA-40 driven simulation (ERA40, green), control simulation (CTL, blue), and scenario simulation (SCN, red). Relative biases and projected changes between CTL and reference (which is ERA40 in (a) and OBS in (b)), ERA40 and OBS (green), and SCN and CTL (red) are shown in the bottom of the figures. Note that frequencies are displayed on a logarithmic scale and the different limits of the x axes in (b).*

In a recent model intercomparison study of Tabari et al. (2016), the downscaled daily precipitation from both the ERA-40 driven as well as GCM CNRM-CM3 driven 4-km

ALARO-0 model simulations are compared with the daily precipitation values from the driving models. Figure 5.9 shows for the CTL period 1961-90 the daily summer precipitation intensities in Uccle as a function of return period. In line with our previous results (e.g. Fig. 5.7(a)), the daily summer precipitation intensities from the ALARO-0 model driven by ERA-40 (ALARO$_{ERA40}$, blue dots) closely approach the observed intensities (Obs., black dots). Furthermore, Fig. 5.9 demonstrates that the previously suggested underestimation in daily (and hourly) summer precipitation by the CTL simulation (ALARO$_{CNRM-CM3}$, red dots), can be attributed to a pronounced dry bias which is apparent in CNRM-CM3 (yellow dots). It can be seen that this large bias in the CNRM-CM3 GCM, is significantly reduced when the GCM is downscaled to a high resolution of 4 km using the ALARO-0 model (ALARO$_{CNRM-CM3}$, red dots).
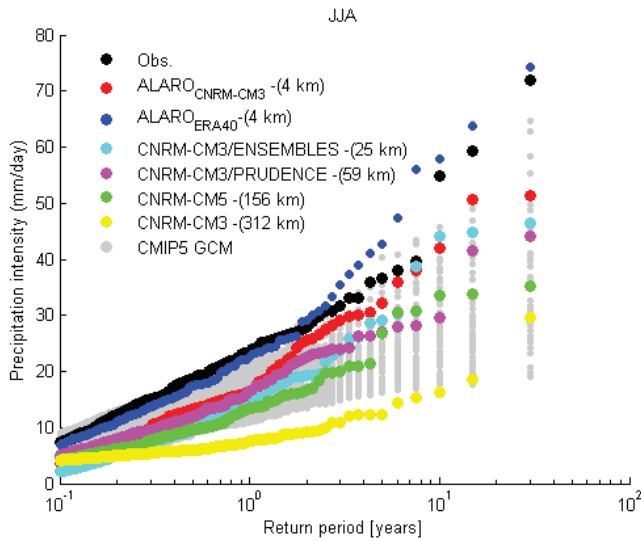
Figure 5.9: *Daily summer (JJA) precipitation intensities in Uccle as a function of return period for the control period 1961-90 obtained from the ALARO-0 model at 4 km resolution driven by the ERA-40 reanalysis (ALARO$_{ERA40}$, blue dots) and the GCM CNRM-CM3 (ALARO$_{CNRM-CM3}$, red dots), together with intensities from different versions of the driving CNRM-CM3 model and the latest Coupled Model Intercomparison Project Phase 5 (CMIP5) GCM ensemble (adopted from Tabari et al., 2016).*

It is unclear to what extent these biases for Belgium are representative for other regions in Europe. However, published evaluations suggest biases of a similar magnitude in some regions of central and northern Europe. Baguis et al. (2010) for example, evaluated for the same control period 1961-1990 monthly mean precipitation amounts for central Belgium from an ensemble of GCM driven RCM data from the European project PRUDENCE. Compared to the observations of Uccle, the RCMs show a general overestimation in precipitation for the winter months and an underestimation during the summer months. In line with the findings of Baguis et al. (2010) and with the tendencies identified from previous spatial- and frequency distributions, the 30-yr average of monthly precipitation values as simulated by our control simulation in the closest grid point to Uccle reveals a similar response w.r.t. the observations (Fig. 5.10).
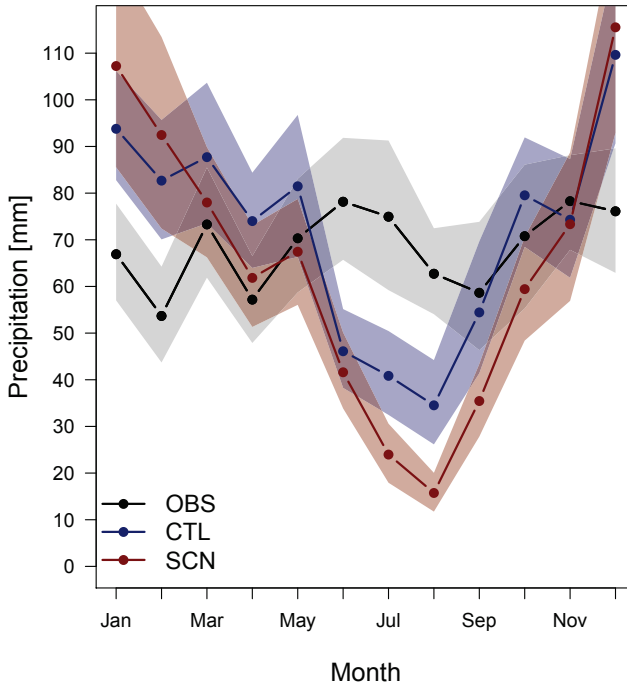
Figure 5.10: *30-yr (control period 1961-90, scenario period 2071-2100) average monthly mean cumulated precipitation in Uccle from observations (OBS, black), the control simulation (CTL, blue), and the scenario simulation (SCN, red). Shaded areas indicate the 95% confidence intervals as inferred from the 2.5th and 97.5th percentile of 1000 bootstrap samples which are obtained by randomly selecting 30 monthly mean values from the respectively datasets.*

Fowler et al. (2005) assessed the performance of the 50-km resolution RCM HadRM3H, driven by the GCM HadCM3, in the simulation of UK mean and extreme rainfall. Similar to our results, the authors find an overestimation by the model in mean rainfall during winter, particularly at high elevations, but underestimates rainfall in summer. These seasonal anomalies lead to significant underestimations (up to 170%) of annual mean rainfall in some parts of the UK (Fowler et al., 2005). In another validation study for Germany, Berg et al. (2013b) analyzed the performance of a five-member ensemble of high-resolution RCM simulations downscaled with two different GCMs at a 7 km nested domain over Germany. Although, in contrast to our evaluation results, the authors find a consistent cold bias for their study region, the magnitude of the deviations of individual members of the ensemble reach similar values up to 4.3 K for summer and 1.9 K for winter. It is suggested by the authors that the temperature biases are a direct consequence of biases in shortwave radiation due to deficiencies in cloud cover.

Furthermore, for precipitation, the authors find for all seasons for the GCM driven RCM simulations an overestimation in comparison to the observations. For summer, this response is in contrast to our results, showing a systematic underestimation of summer precipitation. However, the magnitudes of the precipitation biases as found by Berg et al. (2013b) lie in the same range as ours. From this, we could conclude that although the biases are sometimes

large, the high-resolution control simulation with ALARO-0 is similar to other state-of-the-art simulations efforts within Europe.

5.4.2   Mean changes in daily precipitation and temperature

The previous section has investigated how well ALARO-0 reproduces the control climate w.r.t. observed mean temperature and precipitation. It has become clear that the good model performance obtained from the reanalysis driven simulations does not persist if the model is coupled to a GCM, leading to large model-observation differences up to 4°C for daily temperature and more than 50% for daily and hourly precipitation.
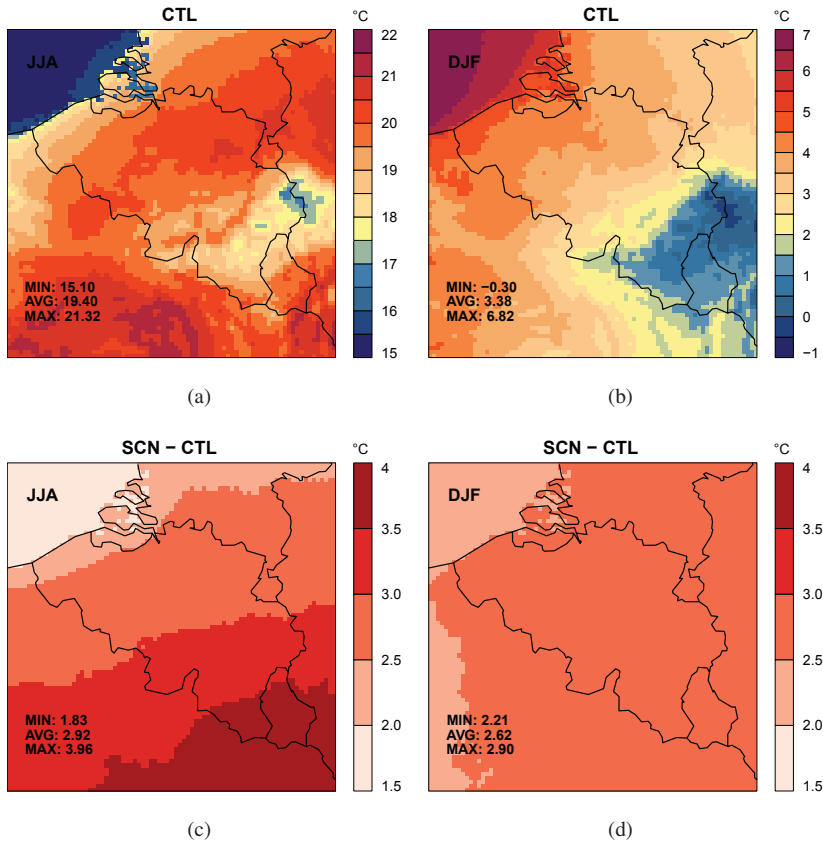
Figure 5.11: *30-yr mean daily summer (JJA, 1961-1990) and winter (DJF, 1962-1990) temperature (°C) estimated from (a-b) the control simulation (CTL) and (c-d) projected changes of daily summer and winter temperature in the scenario period (SCN) (JJA, 2071-2100; DJF, 2072-2100) relative to the control period (CTL) (JJA, 1961-1990; 1962-1990) (°C). For each grid point the significance of the changes between SCN and CTL is tested with the Kolmogorov-Smirnov (K-S) test. All changes are found to be significant at the 5% significance level. The numbers correspond to the minimum (MIN), average (AVG) and maximum (MAX) of the significant values over all 81 × 81 grid points.*

The actual effect of biases which are present in the GCM and its consequences on the

uncertainty of the future climate change results, will be further explored in a later section of this chapter (Section 5.4.4). Provided that we keep in mind the above mentioned limitations present in the control model simulation, we investigate in this section the potential changes in mean temperature and precipitation.

Projections of the future changes in summer and winter daily mean temperature and precipitation are shown in the spatial distributions displayed in Figures 5.11 to 5.12 and the frequency distributions are given in Figures 5.6 to 5.7.
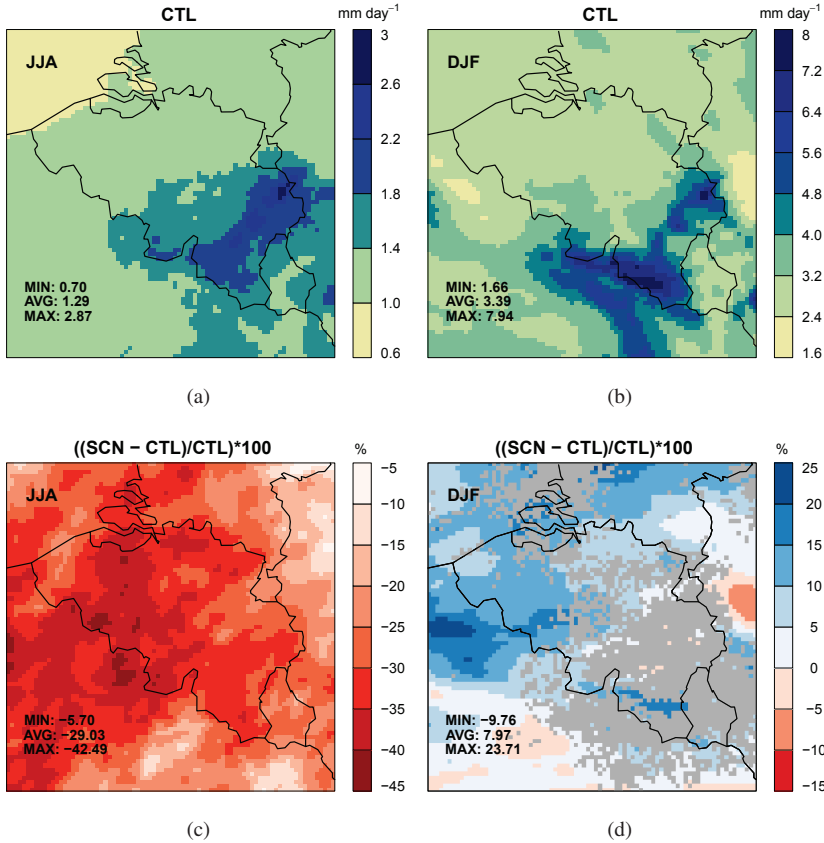
Figure 5.12: *As in Fig. 5.11, but for daily precipitation (mm day^{-1}). For each grid point the significance of the changes between SCN and CTL is tested with the Kolmogorov-Smirnov (K-S) test. Gray indicates changes that are not statistically significant at the 5% significance level according to the results of the K-S test.*

We find a significant warming in 2-meter mean temperature by the end of the 21st century throughout the summer and winter season (Fig. 5.11). For both seasons the differences are smaller in the coastal region than in the rest of the country. This gradual distribution is most pronounced during the summer showing a clear increase in anomalies going from the northwest to the south of the country. Furthermore, the increase in daily mean temperature during summer corresponds on average to $\approx 3°C$. The precipitation field given in

Fig. 5.12(c), shows a significant decrease up to $\approx 45\%$ in summer precipitation for whole Belgium. For the winter, only the coastal and central region of the country displays significant and positive changes in precipitation (Fig. 5.12(d)).

Changes in frequencies of the SCN simulation w.r.t. the CTL simulation, give a first indication of changes in extreme temperature and precipitation (Figs. 5.6 to 5.8). In line with the spatial distributions, the frequencies for 2-meter mean temperature show a clear shift to higher temperature values for SCN, with an increase in the number of the highest summer- and winter temperature values w.r.t. CTL. For precipitation, summer frequencies in SCN are consistently lower compared to CTL, while the SCN frequency distribution for winter displays w.r.t. CTL a slight decrease in dry days and an increase in frequencies for the higher precipitation amounts ranging between 20 and 60 mm day^{-1}.

Tabari et al. (2016) assessed by means of change factors for the model grid point of Uccle, the projected changes in daily precipitation from the CNRM-CM3 driven ALARO-0 simulation in the context of the latest Coupled Model Intercomparison Project Phase 5 (CMIP5) GCM ensemble. The authors found for the summer season that the change factors of ALARO-0 are rather situated on the lower side of the CMIP5 ensemble, while for the winter season the models' change factors were found to lie above the mean of the CMIP5 ensemble. This is in agreement with our findings of a substantial decrease (increase) in daily summer (winter) precipitation amounts and frequencies in the SCN simulation compared to the CTL simulation. In addition, the authors compared the change factors of daily winter and summer precipitation from the CNRM-CM3 driven ALARO-0 simulation with the change factors obtained from the driving GCM CNRM-CM3 model. The comparison revealed for both seasons no remarkable differences in the change factors for the downscaled ALARO-0 simulation and those obtained from the driving CNRM-CM3 model. This demonstrates that the projected changes of the CNRM-CM3 driven ALARO-0 simulation are mainly a reflection of the climate change signal which is present in the driving GCM (Tabari et al., 2016).

The projected changes in 1-hour precipitation for summer are found to be similar to the changes in daily summer precipitation, i.e. a systematic and significant decrease in 1-hour summer precipitation throughout Belgium (Fig. 5.13). As can be seen from Fig. 5.8, the same response is also visible from the frequency distributions obtained for the subgrid domain of 81×81 grid points as well as for Uccle. However, the projected relative decrease as shown by the 81×81 grid points appears to be lower than the projected decrease for Uccle.

These results are in general agreement with other regional modeling studies that assessed the potential future changes in mean temperature and precipitation. Jacob et al. (2014) found for the EURO-CORDEX ensemble seasonal positive changes of mean temperature in Europe in the same order of magnitude as we have found for our A1B scenario. When it comes to precipitation, the projected changes are more uncertain. As discussed in the Introduction (Section 5.1), Belgium is located in the transition zone where climate models are found to show large disagreements concerning the tendency of future precipitation changes. The sign of the changes in seasonal daily and 1-hour mean precipitation is similar to what has been found by Baguis et al. (2010) for central Belgium; an increase in seasonal precipitation during winter and a decrease during summer.
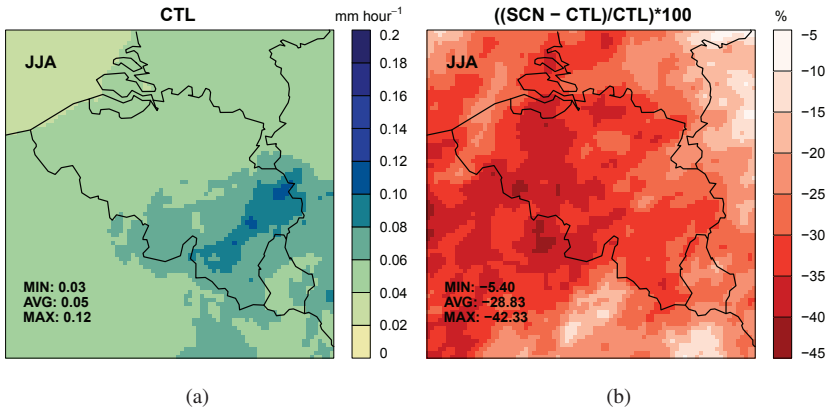
Figure 5.13: *30-yr mean hourly summer (JJA, 1961-1990) precipitation (mm hour^{-1}) estimated from (a) the control simulation (CTL) and (b) projected relative changes of hourly summer precipitation in the scenario period (SCN) (JJA, 2071-2100) relative to the control period (CTL) (JJA, 1961-1990) (%). For each grid point the significance of the changes between SCN and CTL is tested with the Kolmogorov-Smirnov (K-S) test. All changes are found to be significant at the 5% significance level. The numbers correspond to the minimum (MIN), average (AVG) and maximum (MAX) of the significant values over all 81×81 grid points.*

However, as stated before, one is more confident in a model response which is greater than the model bias than in a model response which is smaller than the bias. As can be seen from the 30-yr mean monthly precipitation values in Uccle given for the observations, and the control and scenario simulation, future changes only appear to be significantly different from the control simulation during the summer months (Fig. 5.10). Furthermore, relative differences in the frequencies of daily winter precipitation show a future positive response which is much lower than the strong wet biases present in the control simulation. Hence, making the positive change in mean winter daily precipitation relatively doubtful. Tendencies in summer daily and 1-hour precipitation are generally also smaller than the model biases. The only feature which has similar or larger magnitudes than the model biases, appears to be the strong and consistent decrease in the number of dry days and hours, or a lowering in the number of days and hours with low precipitation amounts during summer. In the next sections, we will explore more in detail if these mean precipitation changes during summer are also reflected in the extremes.

5.4.3   Changes in extreme precipitation

As previously mentioned in the Introduction (Section 5.1), increases in the frequency or intensity of heavy precipitation events under global warming, are potentially one of the most important impacts on society (Muller et al., 2011). In accordance to the methods used for the evaluation of extreme daily- and subdaily precipitation, future changes in extreme (sub)daily precipitation as simulated by SCN are assessed by means of the CC relation, Extreme Value Analysis, and extreme indices. Furthermore, for reliable projections it is important that processes leading to long-term changes in local precipitation such as e.g. relationships of

precipitation with temperature are captured (Maraun et al., 2010).

5.4.3.1 Clausius-Clapeyron relation in a warming climate

If the relative humidity in the future climate remains approximately the same as in the present-day climate - which is generally expected based on model results and physical arguments - the moistening of the atmosphere could lead to an increase in the frequency and intensity of extreme precipitation events following the same CC rate of increase of 6-7% per degree (Allen and Ingram, 2002; Frei et al., 2006; Lenderink and van Meijgaard, 2010; Attema et al., 2014).

Figures 5.14(a) to 5.14(c) show the dependencies of different extreme percentiles (90th-99.9^{th}) of the distribution of hourly precipitation on temperature in Uccle and its closest model grid box values for the observations, and CTL and SCN simulation. Scaling exponents are also calculated for different percentiles of hourly precipitation by fitting a linear regression between the logarithmic of the hourly precipitation amounts and the daily mean temperatures (Fig. 5.14(d)). Similar to the results from the ERA-Interim driven simulations (see Chapter 4, Section 4.4.3.2), CTL also shows a strong reduction in modeled precipitation intensity for temperatures above $\approx 22°C$, and in particular for the 90th and 95th percentiles. Hence, likewise as in previous Chapter 4, the scaling exponents are calculated for daily mean temperatures to $22°C$. To quantify whether the observations and model reproduce the super-CC scaling, the scaling exponents are also computed separately for temperatures below and equal to $12°C$, as well as for temperatures above $12°C$.

As can be seen from Fig. 5.14(b) and Fig. 5.14(d), the CTL simulation reproduces for Uccle relatively well the CC scaling for the highest 99th and 99.9^{th} percentiles. However, the super-CC scaling as visible in the highest observed percentiles for temperatures above $12°C$ is not captured by CTL (Fig. 5.14(d)). As can be seen, CTL systematically underestimates the observed scaling exponents. Although, a similar underestimation has been identified in the ERA-Interim driven simulation (Fig. 4.15(a)). The hourly precipitation dependency with daily mean temperature in the SCN simulation (Fig. 5.14(c) and Fig. 5.14(d)) are consistent with the expectations from the CC relation, i.e., the increase of the most extreme events appears to converge to a value of about 7%/°C. This implies that on the regional scale hourly precipitation extremes are constrained by the CC relation, that is, by moisture availability (Ban et al., 2015). Furthermore, the scaling exponents obtained for the SCN simulation approach very closely the ones from CTL. Only for the lower percentiles up to the 90th percentile the scaling exponents for SCN are consistently higher than those of CTL. Although, the differences between the SCN and CTL scaling exponents are still smaller than the differences between the model and observations for the control period.

To further asses the climate change signal of subdaily precipitation, percentiles of observations, and the control and scenario simulation in Uccle are calculated for 1-hour summer and winter precipitation (Fig. 5.15). For summer, only from the 99.95^{th} percentile onwards SCN percentiles are higher than CTL. However, the changes in the summer percentiles of SCN w.r.t. CTL is as large as the bias in CTL w.r.t. the observations. Hence, the future response cannot be considered to be robust. On the other hand, one can be more confident in the changes of the percentiles, and in particular the highest percentiles for winter hourly precipitation. The highest winter percentiles show a positive future change of more

than 50% which is much larger than the model bias in the control simulation. Although, it was previously found that we could not grant much confidence to the positive changes in daily mean winter precipitation, the latter suggests a more certain intensification of extreme 1-hour winter precipitation in Uccle under the A1B scenario as calculated with the ALARO-0 model.
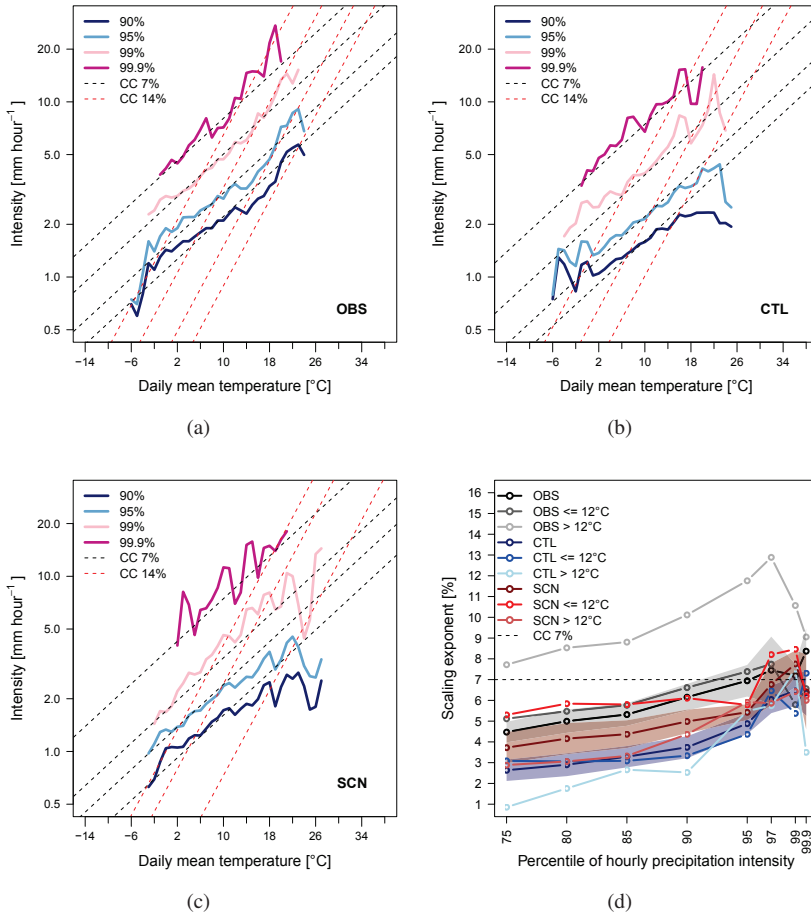
Figure 5.14: *(a-c) Dependencies of different extreme percentiles (90th-99.9^{th}) of the distribution of (a) observed (OBS), (b) control (CTL), and scenario (SCN) hourly precipitation on temperature in Uccle and its closest model grid box values, for the CTL period 1961-1990 and the SCN period 2071-2100. Exponential relations given by a 7% and a 14% increase per degree are given by the black and red dotted lines, respectively. Percentiles are displayed on a logarithmic y axis. (d) Variation of the observed (OBS, black), CTL (blue) and SCN (red) scaling exponent with percentile of hourly precipitation. The scaling exponents are calculated for the whole temperature range up to $22°C$, as well as for temperatures below and equal to $12°C$ and for temperatures above $12°C$. The horizontal black dotted line corresponds to the theoretical 7% Clausius-Clapeyron (CC) increase per degree. Shaded areas correspond to 95% confidence intervals given by $\pm 1.96 \times$ the standard deviation of the respectively scaling exponents which are calculated for the whole temperature range up to $22°C$.*
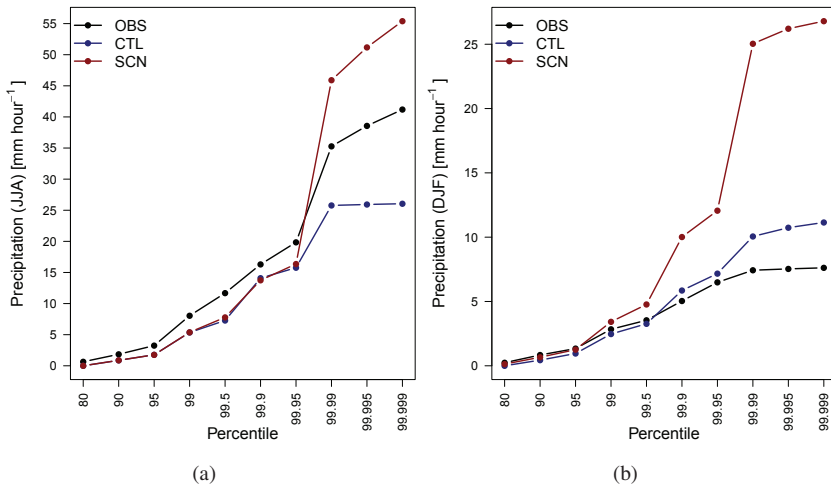
Figure 5.15: *Percentiles of 1-hour precipitation in (a) summer (JJA, 1961-1990 and 2071-2100), and (b) winter (DJF, 1962-1990 and 2072-2100) for observations (OBS), control simulation (CTL), and scenario simulation (SCN) in Uccle.*

For summer, these results are in contrast to what has been found in some previous studies, suggesting a super-CC scaling for extreme summer precipitation events in future. Lenderink and van Meijgaard (2008) for example, analyzed the relative change of a 30-yr present and future climate integration with a 25-km RCM, and found for central Europe changes in 1-hour summer precipitation extremes that exceed 10% per degree. Furthermore, Kendon et al. (2014) performed GCM driven climate change experiments with a 12-km RCM and a very high resolution 1.5-km model, nested into the 12-km model. The 12-km RCM and the 1.5-km convection permitting model have similar model physics, except that at 1.5 km resolution the convection scheme has been switched off. The authors found for southern UK a robust future increase in extreme hourly winter precipitation for the 12-km model as well as for the 1.5 km model. However, in summer the convection-permitting model detected an intensification of extreme summertime hourly rainfall which was not seen in the coarser 12-km resolution model. As suggested by the authors, in summer, deficiencies in the convective parameterization scheme in coarse resolution models thus have a serious impact on projections of changes in precipitation (Kendon et al., 2014). The positive response in extreme winter hourly precipitation for southern UK, as found by Kendon et al. (2014), is in agreement to our findings for Uccle (Fig. 5.15). Although the future increase in extreme hourly summer precipitation as found by Lenderink and van Meijgaard (2008) and Kendon et al. (2014) is not present in our results. As suggested by Ban et al. (2015), these discrepancies with our results might be due to different models and regions considered. As mentioned previously, Belgium is indeed located in the transition zone where climate models are found to show large disagreements concerning the tendency of future precipitation changes.

Furthermore, up to now there is no complete and consistent picture of how subdaily extreme rainfall patterns might change in a future climate (Westra et al., 2014). There is

generally low confidence in the projections for subdaily precipitation, in particular because of the difficulties of the model to correctly simulate precipitation at these high temporal scales (Kendon et al., 2014).

5.4.3.2 Extreme Value Analysis

The GEV distribution is fitted to the summer annual maxima of 1-hour precipitation from the ERA-40 driven ALARO-0 simulation, as well as the from the GCM driven control and scenario simulation. The analyses are done for each grid point over the 81×81 subdomain separately. In a first step, the spatial distributions of the biases and future changes of the estimated parameters describing the GEV distribution (i.e. location μ, scale σ, and shape γ) are assessed.
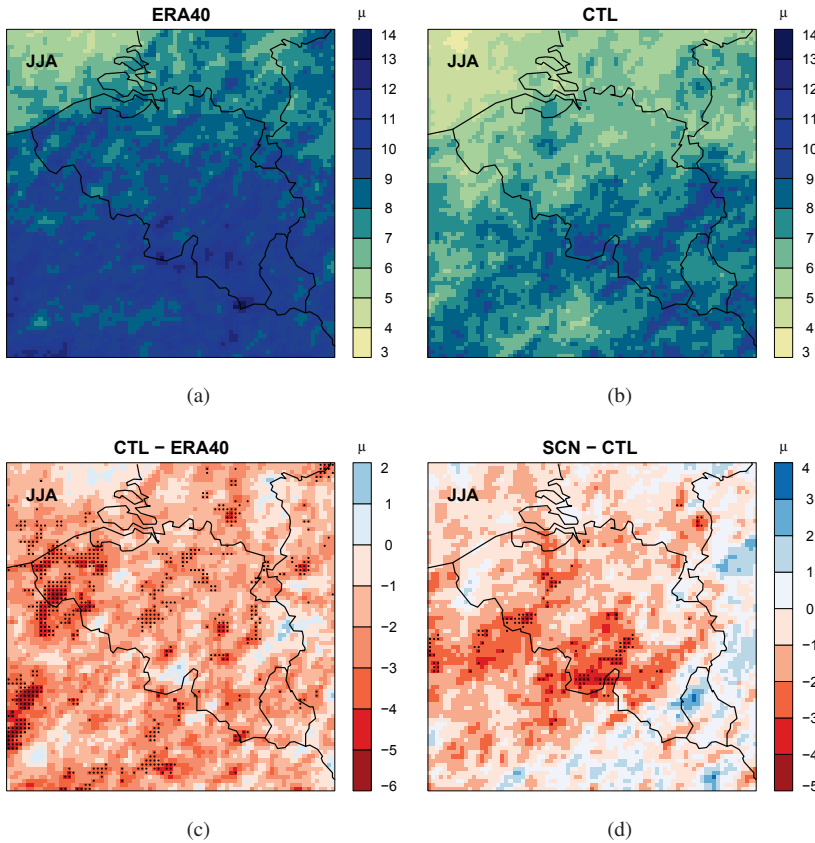
Figure 5.16: *Spatial distributions of the estimated location parameter μ from the GEV fit of 1-hour precipitation in summer (JJA, 1961-1990 and 2071-2100) for (a) the ERA-40 driven simulation (ERA40), and (b) the control simulation (CTL), and (c-d) absolute differences between CTL and ERA40, and SCN and CTL, giving the bias and projected change, respectively. Dotted areas indicate regions where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and ERA40, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*

The spatial distributions of the location parameter as estimated from CTL shows systematic negative biases w.r.t. ERA40 (Fig. 5.16(c)). Also the estimated future change in the location parameter appears to be mostly negative throughout Belgium.
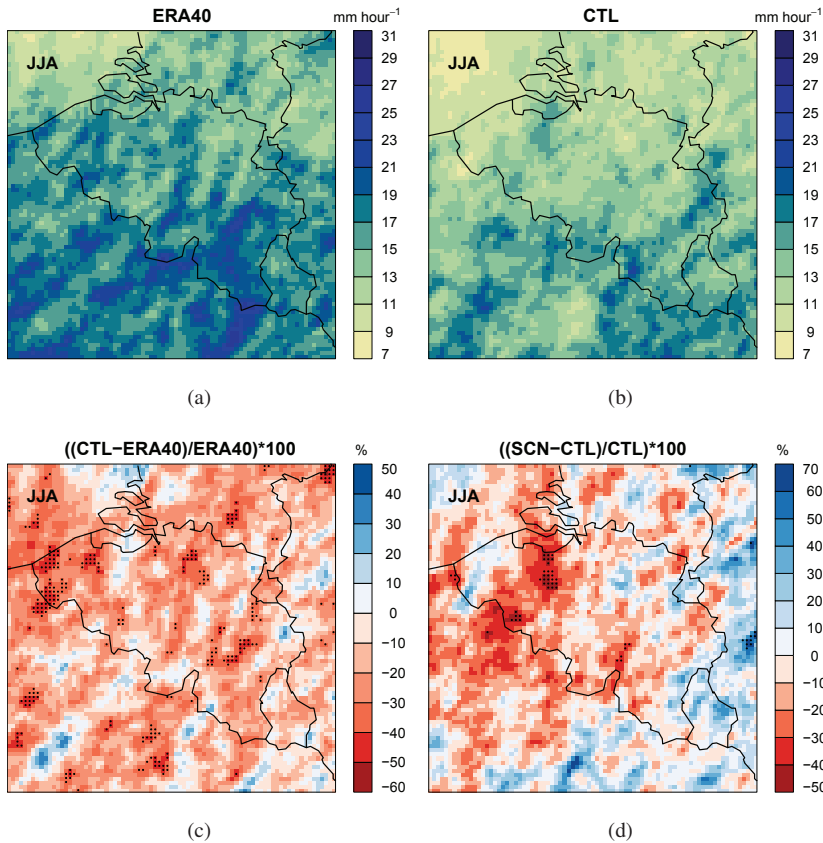
Figure 5.17: *Spatial distributions of the 5-yr return level of 1-hour precipitation in summer (JJA, 1961-1990 and 2071-2100) for (a) the ERA-40 driven simulation (ERA40), (b) the control simulation (CTL), and (c-d) absolute differences between CTL and ERA40, and SCN and CTL, giving the bias and projected change, respectively. Dotted areas indicate regions where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and ERA40, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*

However, both the biases as well as the projected changes are significant for only a limited number of grid points. Taking into account that μ is closely related to the mean of the GEV distribution, these negative responses in the location parameter describe a shift of the whole distribution towards lower values. The spatial distributions of the biases and projected changes for the scale and shape parameter display a mostly non-significant and highly spatially variable distribution (see Appendix C, Figs. C.1 to C.2). Hanel and Buishand (2010), analyzed hourly and daily precipitation extremes in the Netherlands in 13 RCM simulations, where eight of the simulations were driven by transient runs of GCMs forced by the SRES

A1B emission scenario and five of them by perfect boundary conditions of ERA-40 reanalysis. The authors fitted the GEV distribution to the annual maximum amounts and evaluated and assessed the GEV parameter estimates for present and future climate conditions. In agreement with our results, the authors find a negative bias in the location parameter of 1-hour precipitation extremes in the majority of the RCM simulations considered. However, their evaluation reveals for most RCM simulations a significant overestimation of the scale and shape parameter. Although, the large differences between the projected GEV parameters for the different RCM simulations, all three GEV parameters show a general increase leading to very large increases in large quantiles (or return levels), i.e., a 45-60% increase at return periods from 50 to 200 years.
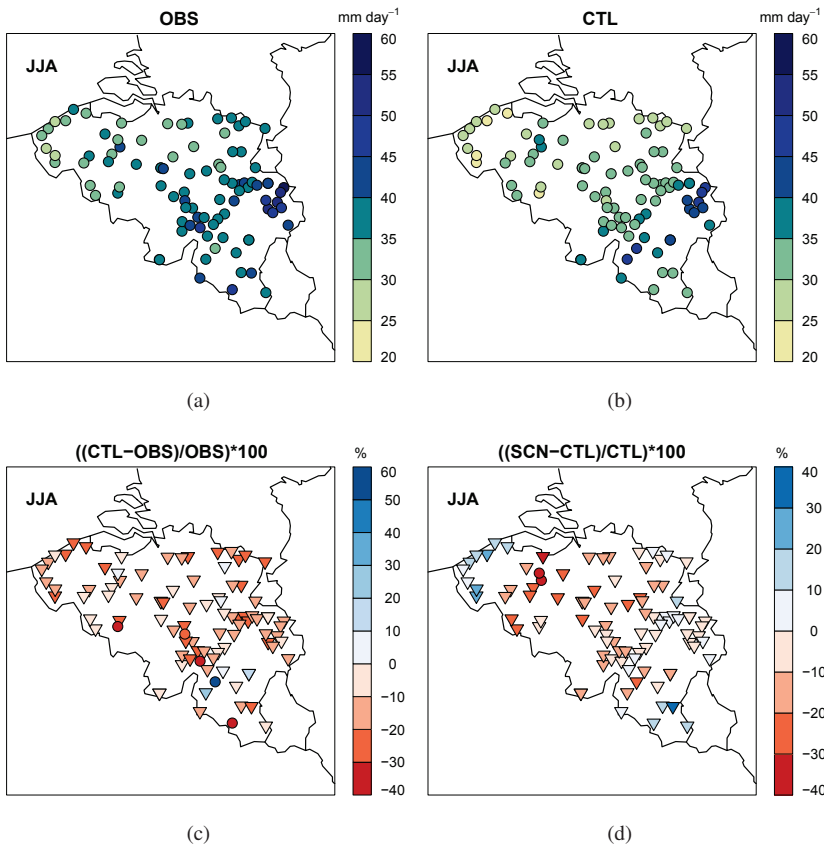
Figure 5.18: *Spatial distributions of the 5-yr return level of daily precipitation in summer (JJA, 1961-1990 and 2071-2100) for (a) the observations (OBS), (b) the control simulation (CTL), and (c-d) absolute differences between CTL and OBS, and SCN and CTL, giving the bias and projected change, respectively. Circled symbols indicate stations where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and OBS, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*

Similar to the GEV analyses in previous chapter (Chapter 4), we have also applied a K-

S goodness-of-fit test to verify the accuracy of the GEV fits. At the 5% significance level, the null hypothesis that the annual extremes are drawn from the GEV distribution, is accepted for all 81×81 grid points in the ERA40 simulation, and the control and scenario simulation. This thus justifies the use of the GEV distribution as a model for the precipitation extremes.

The spatial distributions of the bias and future change in 5-yr return levels are presented in Fig. 5.17. Overall, negative biases and negative relative changes in 5-yr return levels are visible, which may result from the negative tendencies in the location parameter. However, the decrease in the return levels is only significant for a few grid points and smaller or equal in magnitude than the biases in the control simulation.

Return level values are also calculated for daily precipitation extremes, both for the summer and winter season (Fig. 5.18 and Fig. 5.19).
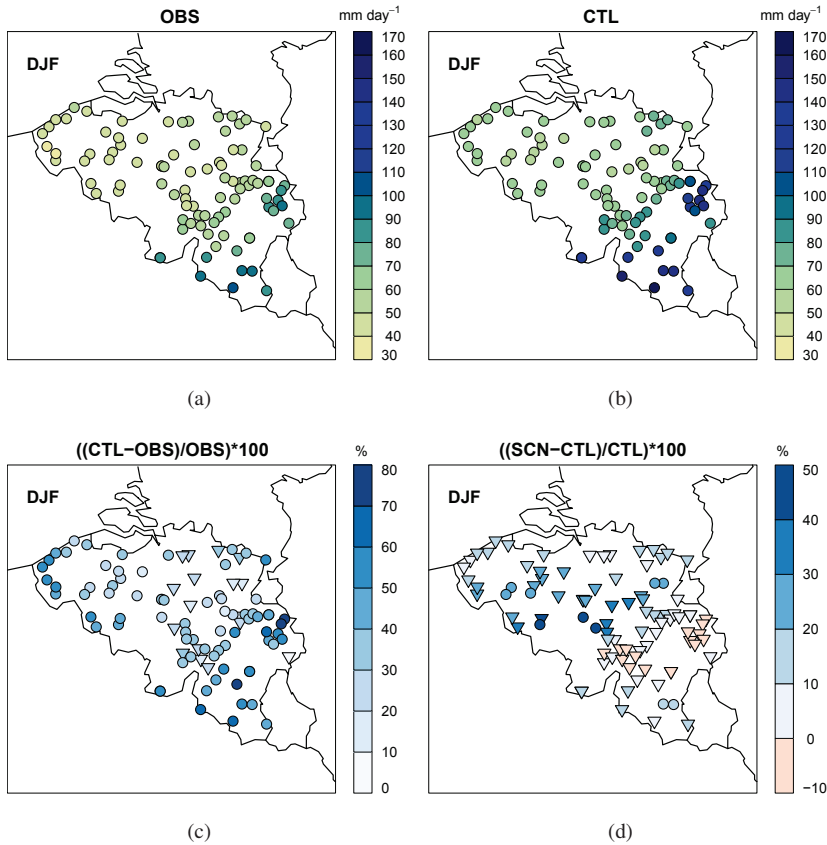
Figure 5.19: *Spatial distributions of the 5-yr return level of daily precipitation in winter (DJF, 1962-1990 and 2072-2100) for (a) the observations (OBS), (b) the control simulation (CTL), and (c-d) absolute differences between CTL and OBS, and SCN and CTL, giving the bias and projected change, respectively. Circled symbols indicate stations where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and OBS, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*

Here, the analysis is done for the observation stations, and the closest model grid point values to the station locations from the CTL and SCN simulation. Although the magnitudes are smaller, the results of 5-yr return level values for daily summer precipitation (Fig. 5.18) show negative biases and future changes, similar to the 1-hour precipitation results. Both the biases and changes are again significant for a few station locations. Similar with previous results for daily mean precipitation, the results for winter daily precipitation display an opposite pattern to the results for summer, with positive biases and changes in 5-yr return level estimates which are significant for a relatively high number of stations (Fig. 5.19). Our results are in agreement with the findings of Kyselý and Beranová (2009). The authors applied a peaks-over-threshold analysis to an ensemble of RCM outputs from the PRU-DENCE project to estimate the effects of climate change of extreme daily precipitation in Czech Republic (i.e. central Europe) by means of changes in return level values. Their results demonstrate that in a warmer late-21st century climate, extreme precipitation events are likely to increase in severity in winter and, with less agreement among the models, also in summer. With some level of confidence, we could conclude from previous results that extreme daily precipitation during winter is expected to increase.

5.4.3.3   Extreme statistics

Finally, similar to Beniston et al. (2007), we have calculated relative changes in several extreme statistics of daily precipitation in summer and winter averaged over all 93 station points: fre: Frequency of wet days (daily amount larger than 1 mm); me: mean seasonal precipitation; int: precipitation intensity (average amount on wet days); $q95$: 95% quantile of wet days; dsl: dry spell length (annual maximum of consecutive periods of dry days, i.e. daily amount smaller or equal than 1 mm); wsl: wet spell length (annual maximum of consecutive periods of wet days, i.e. daily amount larger than 1 mm); $R1d$: 30-yr means of summer maximum 1-day precipitation totals; $R5d$: 30-yr means of winter maximum 5-day precipitation totals. These different aggregation times of 1- and 5-day aggregations account for the different character and impact of extreme precipitation in the two seasons; extreme winter precipitation is generally due to persistent large-scale precipitation whereas extreme summer precipitation is more often due to short-term localized convective activity (Beniston et al., 2007). To assess the uncertainty of the relative changes, the relative biases of the extreme statistics as inferred from the control simulation w.r.t. the observations have also been calculated.

In line with previous results, the relative changes of extreme statistics for summer show a decrease in the frequency of wet days (fre) and mean precipitation (me). Furthermore, the annual maximum of both consecutive periods of dry as well as wet days appears to increase under future-climate conditions. This suggests that when dry or wet periods occur, they will last longer. However, it can be seen that the level of confidence in the changes is relatively low, since most changes (and its confidence intervals) are not or only slightly larger (smaller) than the biases. For winter, most extreme statistics vary around one, and are smaller than their respective biases. Although, in agreement with the mostly significant increase in 5-yr return levels for daily winter precipitation (Fig. 5.19), the statistics for which a change is visible, also show a positive change. The mean precipitation (me), the mean precipitation on wet days (int), and the 95% quantile of wet days show slight positive changes of around 10%

under future climate conditions. The 30-yr means of winter maximum 5-day precipitation totals ($R5d$) also show a positive response, which is opposite to the negative change of the corresponding $R1d$ statistic as obtained from daily summer precipitation.
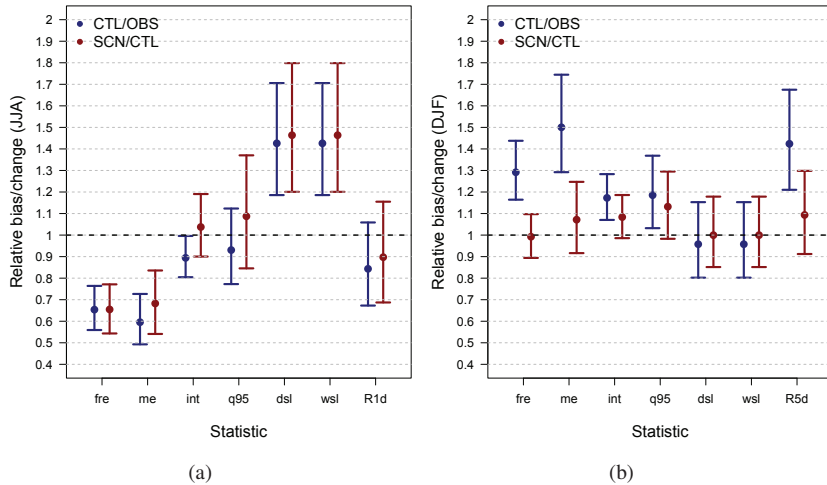
Figure 5.20: *Relative bias (CTL/OBS, blue) and relative change (SCN/CTL, red) in several statistics of daily precipitation in (a) summer (JJA, 1961-1990 and 2071-2100), and (b) winter (DJF, 1962-1990 and 2072-2100) averaged over all 93 station points. Vertical bars represent the 95% confidence interval of the estimated bias / change as inferred from 1000 bootstrap samples.*

5.4.4 Uncertainty assessment and bias correction

Déqué et al. (2007) attribute uncertainties in projected climate change to four different sources: (i) radiative uncertainty, related to the fact that the radiative forcing as described by the amplitude of anthropogenic emissions and the resulting GHG concentration is merely one hypothesis; (ii) model uncertainty, caused by the formulation and accuracy of the atmosphere-ocean GCM (AOGCM) driven by this radiative forcing; (iii) boundary uncertainty, introduced due to the fact that projections of local climate change at high spatial resolutions needed for the impacts community and policy makers, require a further downscaling of the AOGCM with RCMs; and (iv) sampling uncertainty, since the climate statistics are estimated from a finite sample or number of years (usually 30 years) (Déqué et al., 2007). Previous results of the projected climate change can to some extent be assessed by means of these sources of uncertainty.

Generally, the uncertainty of scenarios of anthropogenic climate changes is evaluated by two approaches. One approach is to gain a better understanding of the underlying mechanisms of the regional climate change. By attributing the change to particular mechanisms, and using a qualitative knowledge of the uncertainty of each mechanism, a subjective judgement can be made of the reliability of the change (Rowell and Jones, 2006).

Another common approach, are multi-model ensemble studies, as these allow to estimate an uncertainty range in the projected regional model results. There has been consider-

able international effort to quantify uncertainty in regional climate change through the inter-comparison of multiple RCMs, for example the PRUDENCE and ENSEMBLES projects for Europe, and the NARCCAP (North American Regional Climate Change Assessment Program) project for North America. Furthermore, the recent CORDEX initiative from the World Climate Research Program promotes running multiple RCM simulations at 50 km and 12.5 km resolution for multiple regions (Maraun et al., 2010). The use of such large ensembles of models results in more accurate estimates of future changes, since the effect of internal variability is reduced (May, 2008). In order to provide a range of potential expected future changes in (extreme) precipitation over Belgium, ideally one should have done the above analyses using different future climate scenarios. Unfortunately, due to limited com-puter power it is impossible for a small institute like the RMI to perform 30-year model integrations using different regional models and GHG scenarios at such high spatial resolu-tions. Hence, it should be kept in mind, that the results as discussed above are *one* possible outcome of future changes at the end of the 21st century.

The previous chapters (Chapter 3 and Chapter 4) have demonstrated that the ALARO-0 model, when driven by so-called perfect boundary conditions from reanalyses, is able to correctly simulate daily and subdaily extreme precipitation over Belgium. Furthermore, a validation study by Hamdi et al. (2012), has shown that summer maximum surface air tem-perature over Belgium from the ALARO-0 model at 4 km resolution is correctly represented by the model. Nevertheless, results from previous sections suggest that once ALARO-0 is coupled to the CNRM-CM3 GCM, the model is subject to significant biases which are not present in the control reanalysis driven simulations (i.e. boundary uncertainty). This is related to the second source of uncertainty, as proposed by Déqué et al. (2007). Hence, it is clear that the GCM as driving data introduces additional biases in the downscaling simulations. Therefore, it is important to explore to which extent these GCM biases, are propagated through the downscaling chain. It should however be noted, that both precipita-tion and 2-meter temperature are products of the GCM which are not used in the coupling process to the RCM. The biases as found in these variables from the ALARO-0 control sim-ulations could thus have various sources, and are not necessarily good proxies for biases in the driving fields of atmospheric temperature, winds and humidity provided at the RCM lateral boundaries (Berg et al., 2013b).

As previously discussed, the results from the recent study of Tabari et al. (2016), in-deed suggest that the underestimation (overestimation) in summer (winter) daily extreme precipitation amounts in the CNRM-CM3 driven ALARO-0 simulation for the control pe-riod are explained by underestimations (overestimations) in the global CNRM-CM3 model, rather than in the ALARO-0 model itself. Furthermore, it has been demonstrated by the authors that compared to the driving low-resolution CNRM-CM3 GCM, the downscaling of the GCM data towards higher spatial resolutions of 4 km results in more extreme summer precipitation, significantly reducing the bias w.r.t. the observations in the control simulation. Similarly, Berg et al. (2013b) found that biases from the driving GCMs are generally trans-ferred unchanged to the high-resolution RCMs, but that compared to the GCMs, the RCMs add value to the intensity distributions of precipitation, when compared to observations at the fine nest resolution, and especially for extreme events.

Although the projections of hourly precipitation extremes are often needed for climate change impact assessment, the large deviations in e.g. return level values or extreme statis-

tics obtained from the control simulation w.r.t. the observations, limit our confidence in the results under future climate conditions. If one considers the future impact as the difference between the model output under future climate conditions from the scenario simulation minus the output under the control climate conditions, these deviations or biases mostly due to model imperfections will indeed be included in the impact term. Therefore, to reduce the GCM/RCM biases, bias correction or postprocessing methods can be applied. In recent years, a wide range of bias correction methods have been developed and applied by many users of GCM/RCM output. However, it is beyond the scope to extensively discuss all the existing bias correction methods. We rather aim to provide a taste of the range and approaches of bias correction methods, by briefly highlighting the different methods and indicating for which purposes one or another method might be more useful. Déqué (2007) distinguishes five categories of bias correction methods. The *confident methods*, in which no correction is applied, as used in this chapter, assume that most of the bias is canceled out when studying the climate change response as the difference between the scenario and control simulation. Secondly, the *delta methods*, are often used, and simple methods which assume that both the control and scenario model output can be represented by observations, with the scenario model output being corrected by the mean difference (or ratio in case of precipitation) between the scenario and control simulation. These methods are very robust, but it requires that time series of observational data are available for the study region. Furthermore, they assume that the climate variability in the scenario projection is unchanged and inherited from the observed variability. Therefore, if one is interested in changes in the frequency of extreme events, these methods which assume a simple shift of the observed mean and variance may not be convenient. A third family of methods are the *unbiasing methods*, in which both the control and scenario model output are corrected for the mean model bias (i.e. observations minus control model output). One advantage of these methods compared to the delta methods is that it might be easier to get a climatology (i.e. mean) of the observations than a (daily) times series of the observations at some locations. In the *variable correction methods* a particular function f is build with the observation dataset and applied on the control dataset as well as on the future scenario dataset. These methods are a generalization of the previous unbiasing methods, in such a way that the function $f(x)$ corresponds to $x - B$ where x is the model field to be corrected and B is the mean bias. The functions used in these methods cover many statistical techniques and can for example be based on a linear regression, a quantile-quantile function, etc. The last family of methods, as described by Déqué (2007), are the *regime methods*. These methods cover the clustering and analogue techniques, which are based on statistical models that generate for the control and scenario model output random weather sequences that resemble the statistical properties of observed weather (Déqué, 2007; Maraun et al., 2010).

Although bias correction methods are important tools to address the model biases, this brief overview yet indicates that the large variety of the existing bias correction methods and their corresponding assumptions do not always facilitate it to use the right method in the right way or context. Ehret et al. (2012) argue that bias correction as used to correct GCM or RCM output in climate change impact studies is often used in an invalid way. In general, the biases corrected for are a function of time, space, and meteorological variable and spread in a non-uniform way through the entire distribution of the variable. As suggested by the authors, these complex interactions of the biases are often neglected in the application of the bias

correction. In this way bias correction methods often alter the spatiotemporal consistency of the model fields or the relations among variables. Furthermore, it is unclear whether bias correction methods are time-invariant under climate change conditions. The physical justification of the correction method is often not provided or not transparent to the end user, and hence, the authors conclude that the use of a bias correction rather hides than reduces the uncertainty range of simulations and projections. This is an important motivation for the approach used in this chapter, where the future changes in temperature and (extreme) precipitation are quantified as the difference between the scenario and control model output, without any bias correction.

5.5 Conclusion

This chapter assessed the potential future changes in extreme precipitation for Belgium under the A1B scenario as inferred from the ALARO-0 model at 4 km resolution. The comparison of daily precipitation and temperature, as well as hourly mean precipitation from the 1961-90 CTL simulation with the observations revealed significant biases, which are found to be mainly related to model errors present in the driving GCM CNRM-CM3. Furthermore, it is found that the downscaling of the GCM data towards higher spatial resolutions of 4 km results in more extreme summer precipitation, and significantly reduces the bias w.r.t. the observations in the control simulation.

We have limited ourselves in this chapter to a sensitivity study, where future changes are analyzed through quantification of the differences between the SCN and CTL simulation, without any bias correction. The results from the analysis on the future changes in mean temperature show a significant warming in 2-meter mean temperature by the end of the 21st century throughout the summer and winter season. For mean precipitation our results demonstrate for summer an overall significant decrease for the whole country, and a significant increase during winter for the coastal and central region of the country. The future changes in extreme precipitation were explored through analysis of the CC relation, EVA, and extreme statistics. The most extreme percentiles in the hourly precipitation dependency with daily mean temperature in the SCN simulation converge closely to the CC scaling, i.e. 7%/°C. Although, the scaling exponents for the SCN simulation are only for the lower percentiles up to the 90th percentile slightly higher than those of the CTL simulation. Finally, the results from the EVA show for summer little significant but negative changes in the GEV location parameter, which are reflected in the negative response of 5-yr return level values of hourly and daily precipitation. For winter, 5-yr return level values of daily precipitation show mostly positive and significant changes between the SCN and CTL simulation.

However, we did not attempt to *quantify* the uncertainties of these future changes, the comparison of the changes w.r.t. the model biases as well as w.r.t. other GCM/RCM results, allowed us to provide a qualitative notion of the uncertainty and confidence of our results. Except for a few responses, most of the changes are as large as, or not much larger than the biases. For example, the positive changes in daily mean winter precipitation are of the same order of magnitude as the biases. However, the positive future change in the highest percentiles of hourly winter precipitation in Uccle is more than 50% larger than the model bias. This suggests that with some level of confidence, one can expect an intensification of

extreme hourly winter precipitation in Uccle, which is also reflected in the positive changes in extreme daily winter precipitation as given by the 5-yr return level estimates for Belgium.

The future changes as derived from the ALARO-0 model are to some extend in agreement with other GCM/RCM results. The positive changes in mean temperature as inferred from our ALARO-0 CTL and SCN simulations, are in line with other modeling studies. The distinct seasonal response in mean and extreme precipitation as apparent in our results, is also reported by other modeling studies. Although, the actual projected precipitation tendencies are more uncertain, as they do not show a general agreement with the tendencies as obtained from other regional climate change experiments in Europe. Our finding in the decrease of future summer mean precipitation is also reflected in the negative responses in extreme and hourly summer precipitation. However, these negative changes in extreme summer precipitation are relatively uncertain since (i) most of the changes are non-significant and smaller than the respectively biases, and (ii) other modeling studies for our regions project an increase in extreme summer precipitation, which is in disagreement with our results. In general, the disagreements and highly varying patterns of projected changes in extreme summer precipitation can be explained by the transition zone in which Belgium is located. Furthermore, the strong dependency of the parameterizations, and in particular the deep convective parameterizations, are also a key source for the uncertainty in future climate projections of extreme summer precipitation. In contrast, wintertime projections of heavy precipitation are found to be much more robust. Our results for winter extreme precipitation are indeed in line with other RCM and high-resolution model results for our region, showing a future increase in extreme winter precipitation amounts. Finally, our CC results are in disagreement with the limited number of other high-resolution modeling studies that assessed the CC scaling relationships in the context of climate change. This points to the need for more extensive observational and modeling studies to investigate the dominant physical processes that are effectively responsible for changes of (sub)daily extreme rainfall in a future climate.

As mentioned previously, the CTL climate simulation demonstrates that the downscaling of the driving GCM CNRM-CM3 data with ALARO-0 towards high spatial resolutions of 4 km, results in more extreme summer precipitation. However, compared to the change factors derived from the driving GCM, the change factors as obtained from the CTL and SCN simulation with ALARO-0 at 4 km resolution do not show a significant additional change. Nevertheless, our high-resolution scenario simulations provide valuable information to hydrological applications, as well as useful spatial detail for impact studies in Belgium. For climate change studies at for example the urban scale, this high-resolution information may still remain relatively coarse in space and time. Hence, to tackle this scale difference between the climate model scales and the local urban drainage scale, statistical downscaling or bias correction methods are commonly used. As a future outlook, the ongoing project CORDEX.be (COmbining the Regional Downscaling EXpertise in BElgium: CORDEX and beyond), which aims to produce a set of comparable simulations by the Belgian regional climate modeling groups, will apply several statistical downscaling techniques, to indeed overcome the resolution differences between the RCM simulations and the local impact model simulations, as well as to infer the climate uncertainties so that the CORDEX.be micro-ensemble can be properly situated w.r.t. the other runs in the EURO-CORDEX archive.

6

PRACTICAL ASSESSMENT OF CLIMATE CHANGE IMPACT ON WINTER SMOG EPISODES: A CASE STUDY FOR BRUSSELS

"There is consensus among Global Climate Models (GCMs) that 21st-century climate change will increase the frequency of stagnation episodes over northern mid-latitudes continents."

– Jacob and Winner (2009)

6.1   Introduction

In the previous chapter we have assessed the impact of climate change on extreme precipitation. This chapter focuses on another major area where climate change may cause adverse effects, i.e. air quality. Pollution peaks have a great impact on health and environment, in particular in large cities and urban environments. Hence, reducing the impact of air pollution is a major issue in environmental policy making. Policy makers express a growing interest in quantifying the effect of climate change on air pollution to make an effort to meet the air quality targets in the next years and decennia (Lauwaet et al., 2014). The effects of climate change on surface air quality is often described in the broader context of chemistry-climate interactions. Generally, the concentrations of air pollutants are determined by three factors: (i) anthropogenic and natural emissions, (ii) atmospheric chemistry and (iii) meteorological conditions (Jacob and Winner, 2009; Giorgi and Meleux, 2007). As can be seen from Fig. 6.1, emissions may be affected by an external forcing from a perturbation to anthropogenic emissions resulting from socio-economic factors external to the chemistry-climate system.
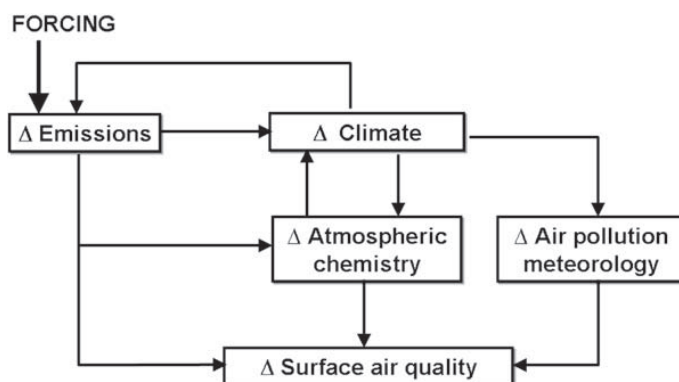
Figure 6.1: *Schematic overview of the effect of climate change on surface air quality in the context of chemistry-climate interactions. An external forcing from change in anthropogenic emissions triggers interactive changes within the chemistry-climate-emissions system, resulting in perturbation to surface air quality (adopted from Jacob and Winner, 2009).*

These changes in emissions in turn cause changes in the other determining factors (i.e. atmospheric chemistry and meteorology), possibly driven by interactive changes in climate. Examples of forcings include anthropogenic emissions of carbon dioxide (CO_2) (driving change in climate), nitrogen oxides (NO_x) (driving atmospheric chemistry), or elemental carbon (driving change in climate as well as direct change in air quality) (Jacob and Win-

ner, 2009). Anthropogenic emissions of pollutants thus change the chemical composition of the atmosphere, which in turn has a feedback effect on the regional and global climate (Juda-Rezler et al., 2012). Furthermore, change in atmospheric chemistry affects air quality (ozone O_3 and Particulate Matter PM) and climate (O_3, PM, methane CH_4), and change in climate affects natural emissions (biosphere, dust, fires, lightning) which cause changes in air quality (Jacob and Winner, 2009). Finally, global, regional and local weather conditions, determined by temperature, precipitation, clouds, atmospheric water vapor, wind speed, and wind direction, also influence the surface air quality through atmospheric chemical reactions and by affecting atmospheric transport and deposition processes as well as the rate of pollutant export from urban and regional environments to global scale environments and vice-versa (Juda-Rezler et al., 2012).

Climate impact on air quality is often assessed by computing scenarios of future greenhouse gas (GHG) emissions using Global Climate Models (GCMs). These global meteorological data together with the changes in the global anthropogenic precursor emissions consistent with the greenhouse scenario may serve as input to a Chemical Transport Model (CTM) that calculates the atmospheric composition on a global scale. If one is interested in air quality changes at finer scales and for a specific region of interest, dynamically downscaled meteorology from a Regional Climate Model (RCM), chemical boundary conditions from the global CTM and (if desired) future pollutant emissions provide then the forcing conditions for a regional CTM (Jacob and Winner, 2009). The GCM/RCM and CTM are usually runned in off-line mode, with no feedback between the input meteorological fields and the tracer concentrations calculated by the CTM, but in some studies also the on-line mode is used (e.g. Giorgi and Meleux, 2007) (Juda-Rezler et al., 2012). Each of the models address a different aspect; the GCMs and RCMs take mostly care of the transport and the dispersion of the pollutants, while the CTMs mostly address the problem of the emissions and the chemical reactions.

This GCM/RCM-CTM approach has been used in many previous studies assessing the relative effects on air pollution (most of them dealt solely with O_3 and PM) from changes in emissions vs. changes in meteorology associated with climate change (e.g. Kelly et al., 2012; Hedegaard et al., 2013; Lauwaet et al., 2014). Kelly et al. (2012) found a potential improvement in ambient air quality for the USA with reducing anthropogenic precursor emissions according to the Intergovernmental Panel on Climate Change (IPCC) Representative Concentration Pathway (RCP) RCP6. On the other hand the impact of climate change alone, when emissions remain fixed at their present values, shows an increase in O_3 and $PM_{2.5}$ concentrations and thus a worsening of the air quality. However, it was found that the magnitude of the potential improvements from changes in emissions is larger than the deteriorated signal from climate change. Other studies for respectively the Benelux and Belgium from Hedegaard et al. (2013) and Lauwaet et al. (2014) reported similar findings, i.e. future changes in air pollutants and more specific O_3 concentrations, are dominated by the expected emission reductions over the expected climatic changes.

Other studies have investigated solely the effect of climate change. In that case the changes in the anthropogenic precursor emissions are not taken into account and the CTMs consider a constant emission rate under the greenhouse scenario (Jacob and Winner, 2009). Several studies demonstrated for Europe that due to changes in meteorological variables favoring the production of O_3 (e.g. temperature, precipitation, solar radiation, ...) a general

degradation of (summer) ozone air quality can be expected (e.g. Giorgi and Meleux, 2007; Katragkou et al., 2011; Juda-Rezler et al., 2012). However, for PM the response to climate change is found to be more complicated than for O_3, because of the diversity of PM components, compensating effects, and general uncertainty in GCM projections of the future hydrological cycle (Jacob and Winner, 2009). Although the effect of climate change on PM could also be significant, the little consensus between studies on the sign of the effect, makes it very uncertain (Jacob and Winner, 2009). Furthermore, Jacob and Winner (2009) reported in their review paper that there is a general agreement among GCMs that 21st-century climate change will increase the frequency of stagnation episodes over the northern mid-latitudes. So even if the emission rates and background concentrations of specific air pollutants would stay constant, climate mitigation measures following climate scenarios will have an impact on air-pollution events due to changes in the meteorological conditions that are unfavorable for the dispersion of air pollutants.

This chapter addresses the impact of climate change on winter smog events in Brussels. These winter smog peaks occur during anticyclonic, calm and stable meteorological conditions, characterized by low wind speeds, an absence of horizontal transport and a temperature inversion with a stable boundary layer that prevents the mixing of the air pollutants into the higher atmospheric layers (European Environment Agency, 1998). If these conditions last for several days, the trapped air pollutant emissions will accumulate until the concentrations exceed the prescribed European threshold values (i.e. European Directive 2008/50/EC European Community, 2008). The magnitude of the resulting concentrations are then proportional to the emissions. The main pollutants involved in such events are a mixture of PM, NO_x (or nitrogen dioxide NO_2) and sulfur dioxide SO_2, and mainly originate from anthropogenic activities such as transport, energy consumption or industrial and power generation facilities (European Environment Agency, 1998). Hence, such events during winter are particularly relevant for urbanized and industrialized regions with higher than average and constant emission rates. The emissions of the winter smog pollutants can be considered to be relatively constant on a day-to-day basis, since they are weakly reactive and during wintertime when photochemistry is less important, they do not act significantly as precursor for the production of ozone for example (Sillman, 1999). This fact allows to determine the peaks to a high degree of reliability by the meteorological conditions, which also has been explored in the past with the application of atmospheric stability classification schemes, such as the Pasquill-type of systems (Pasquill, 1961), where pollutant events are detected or forecasted based on meteorological data only.

Furthermore, the most important anthropogenic driver of climate change is the radiative forcing of the GHGs. However, the most important GHGs (e.g. CO_2 and CH_4) are not considered to be direct pollutants of human health concern such as O_3, PM_{10}, $PM_{2.5}$, SO_2, NO_2, carbon monoxide CO, and lead Pb (Ebi and McGregor, 2008). We asses the future climate change impact on winter smog episodes under the Special Report on Emission Scenarios (SRES) A1B. The background values of the relevant elements of winter smog peaks (i.e. SO_2 and NO_x) also contribute to the overall radiative climate forcings and the changes in their background concentrations and emissions are taken into account in the SRES (Nakićenović et al., 2000), as well as in the more recent RCPs (Lamarque et al., 2011; van Vuuren et al., 2011) used in the latest Assessment Report of the IPCC (Fifth Assessment Report, AR5).

However, as can be seen from Fig. 6.2, the radiative forcing (or climate sensitivity) of these short-lived compounds SO_2 and NO_x is of an order of magnitude smaller than the main anthropogenic GHGs such as CO_2 and CH_4.
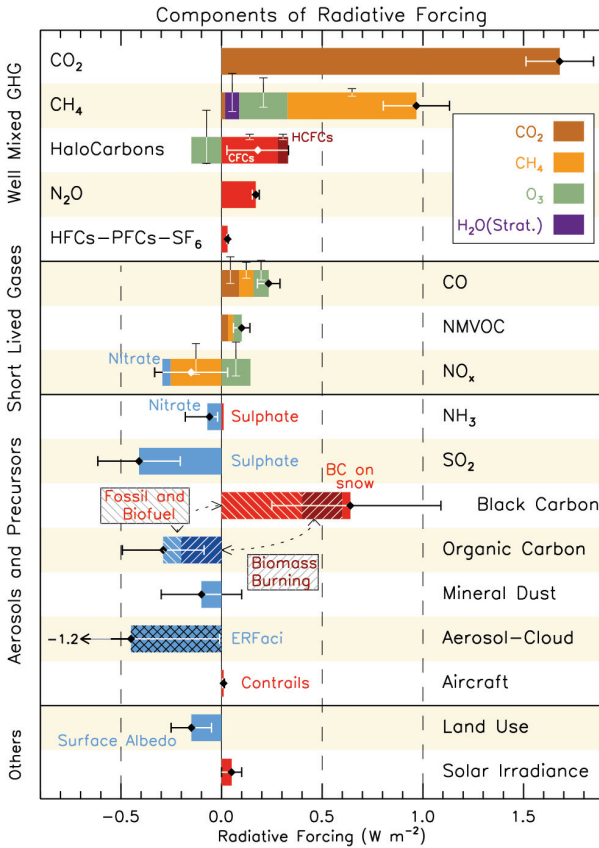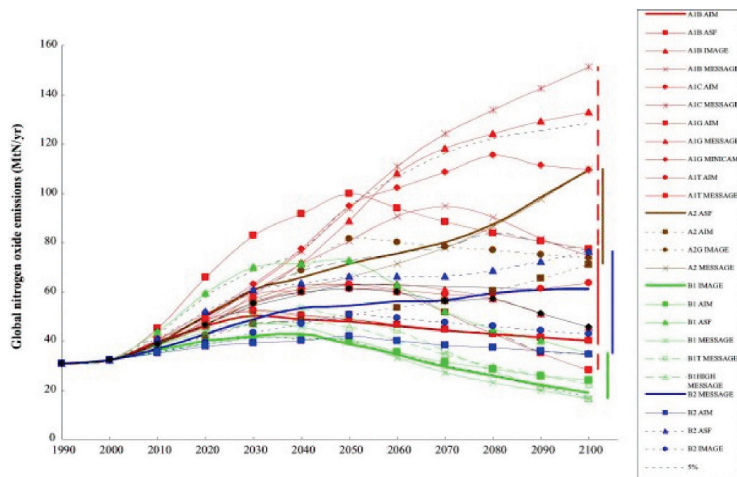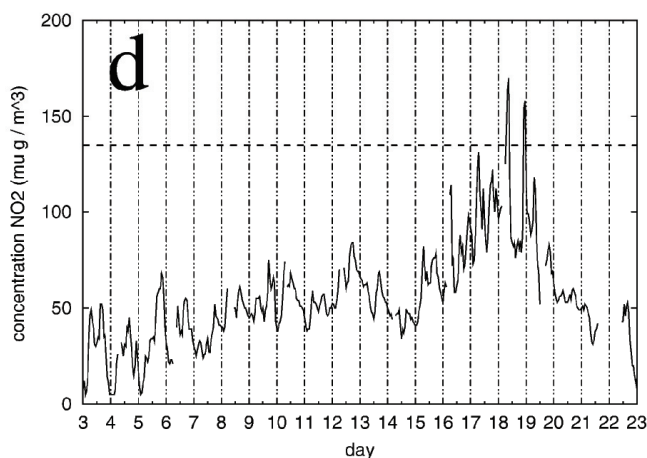
Figure 6.2: *Radiative forcing bar chart for the period 1750-2011 based on emitted compounds (gases, aerosols or aerosol precursors) or other changes. Red (positive radiative forcing) and blue (negative forcing) are used for emitted components which affect few forcing agents, whereas for emitted components affecting many compounds several colors are used as indicated in the inset at the upper part the figure. The vertical bars indicate the relative uncertainty of the radiative forcing induced by each component. Their length is proportional to the thickness of the bar, that is, the full length is equal to the bar thickness for a ± 50% uncertainty. The net impact of the individual contributions is shown by a diamond symbol and its uncertainty (5 to 95% confidence range) is given by the horizontal error bar (adopted from Myhre et al., 2013).*

Furthermore, the projected emissions in the background concentrations of the relevant pollutants of winter smog peaks as given by the SRES and RCPs are of a completely different order of magnitude than the actual peak concentrations occurring during such winter smog episodes when the European prescribed thresholds are exceeded. This important difference between the background and peak concentrations is illustrated in Fig. 6.3. The top figure (Fig. 6.3(a)) shows for the SRES A1B an overall decrease or stabilization by 2100 in the

projected global NO_x emissions.

(a)

(b)

Figure 6.3: (a) *Standardized global NO_x emissions in the SRES scenarios, classified into four scenario families (each denoted by a different color code - A1, red; A2, brown; B1, green; B2, blue). Marker scenarios are shown with thick lines without ticks, globally harmonized scenarios with thin lines, and non-harmonized scenarios with thin, dotted lines. Black lines show percentiles, means, and medians for SRES scenarios (adopted from Nakićenović et al., 2000). (b) The hourly values of the NO_2 concentrations (in $\mu g \, m^{-3}$) in January 2001 for the station of St.-Jans-Molenbeek as measured by the Brussels Institute for Management of the Environment (BIME). The EU guide value of 135 $\mu g \, m^{-3}$ for the 98th percentile of the hourly values measured during the calendar year (Directive 85/580/EEC) is indicated by the dashed line (adopted from Termonia and Quinet, 2004).*

In contrast, the observed NO_2 concentrations, as illustrated in Fig. 6.3(b) for a station in Brussels (St.-Jans-Molenbeek) for January 2001, show that the peak concentrations during winter smog events correspond almost to a threefold of its background concentrations (i.e.

$\approx 50~\mu\mathrm{g~m^{-3}}$ versus $135~\mu\mathrm{g~m^{-3}}$) (Termonia and Quinet, 2004). Hence, these extreme and localized concentrations, as measured during winter smog events, can be considered as *superimposed* peak concentrations on the everyday background concentrations.

In this way, our study that focuses on winter smog peaks allows to isolate the impact of climate change on unfavorable meteorological conditions for the dispersion of air pollutants from the impact of the anthropogenic emissions and production of air pollutants, since both impacts are driven by different forcings and thus potentially by different economic activities. Two stability indices are calculated for a 9-yr winter period using present (1990/91-1998/99) and future (2046/47-2054/55) climate data that has been obtained from a dynamically downscaling of GCM data from the ERA-Interim reanalysis as well as from the coupled CNRM-CM3 model using the limited area model ALARO-0 at 4 km spatial resolution (De Troch et al., 2013; Hamdi et al., 2014, 2015). The simulations are performed within the framework of the ACCEPTED project (Assessment of Changing Conditions, Environmental Policies, Time-activities, Exposure and Disease, Delcloo et al., 2014). The ACCEPTED project aims to set up an observational and modeling approach, accounting for the effects of a changing urban climate, in order to improve our understanding of future exposure situations and their impact on health in a mid-century horizon (2050s) (Hamdi et al., 2015).

The first index has been proposed and tested by Termonia and Quinet (2004) and combines horizontal transport and vertical stability into one single and unique transport index. The index can be easily computed from meteorological output of an atmospheric model such as a Numerical Weather Prediction (NWP) model. In Belgium the transport index has been used for launching smog alerts and has been part of policy measures over the past years. Based upon the index the authors introduced a simple criterion to determine whether the meteorological conditions will lead to a winter smog alert. Furthermore, this criterion can be easily adapted to *detect* events that would lead to smog alerts using output of climate models without the need to run a CTM. This transport index is similar to an index used in an old Pasquill-type scheme (i.e. Bultynck-Malet scheme, Bultynck and Malet, 1972). Although, Termonia and Quinet (2004) demonstrated that the transport index is more convenient to detect the strongest pollution peaks and more easy to interpret physically, the well known Pasquill classes are also calculated in order to provide a reference.

Each model component in the GCM-CTM model chain (i.e. GCMs, RCMs, CTMs, or inline chemical modules) requires additional computing resources and the coupling between the models may imply specific challenges such as lateral boundary issues, two-way nesting issues, consistency of the underlying scientific hypotheses and of the used numerical algorithms. Furthermore, all models have model errors and the resulting model error of the coupled system is the superposition of the errors of the different individual components. It should be avoided to correct for biases in one downstream model if they are originating from model errors in the driving model (Ehret et al., 2012). Assessing and diagnosing the behavior of the GCMs and RCMs for the conditions that are unfavorable for the dispersion of air pollution, independently of the outputs of the CTMs, is therefore important.

In climate modeling, model biases are tackled by carrying out bias corrections, such as for example a correction for the distribution where the mean and standard deviation are readjusted to match better with the observations. Identifying the biases and applying correction methods can also be seen as a way to quantify model uncertainties (Ehret et al., 2012). For this study, where the climate impact on air pollution is studied by means of stability

indices, applying a bias correction directly on the vertical profiles of the meteorological variables used as input to calculate the indices, would break the physical consistency of the profiles or not properly retain the climate change signal of the stability. Applying the bias correction directly on a simple index that incorporates both the stability information and the horizontal transport, we assure a more physically coherent bias correction and facilitate the estimation of the uncertainties in the impacts. Therefore, in the present paper we propose a direct bias correction on the transport index of Termonia and Quinet (2004) based on the quantile mapping approach.

The approach proposed here will be evaluated and applied for Brussels, the capital of Belgium. This is motivated by the fact that Belgium is one of the countries in Europe where air quality levels of different pollutants such as O_3, NO_x, and PM still exceed the prescribed European norms multiple times a year (European Environment Agency, 2014). Furthermore, winter smog pollutants mainly originate from urban emission sources and its impacts are often strongly felt in urban areas. The main known effect on Brussels due to climate change is the so-called Urban Heat Island effect (UHI). Recently, Hamdi et al. (2014, 2015) proposed a new high-resolution dynamical downscaling strategy to examine the Brussels's UHI under present and future climate conditions. The regional climate simulations in Hamdi et al. (2014, 2015) were performed with ALARO-0 coupled to the single layer urban canopy parameterization scheme, Town Energy Balance (TEB). Our study extends the work of Hamdi et al. (2014, 2015), in the sense that we assess an additional impact of climate change on Brussels (i.e. winter smog episodes) by applying our presented method on the same high-resolution downscaled A1B scenario data that accounts for the urban scale aspects in the meteorological conditions. As an application of the presented method we will show that under the studied scenario in this chapter, we can expect an increase of the unfavorable conditions for the dispersion of air pollutants up to 60 - 70% in Brussels by the middle of the 21st century.

Finally, we emphasize that the methodological approach proposed for this study does not replace the use of CTMs or on-line chemistry modeling. Instead, our method can provide a *complement* to the GCM-CTM approach in three ways by: (i) giving an additional estimate of the impact of climate change on winter smog events isolated from changes in emissions, (ii) facilitating a physically coherent bias correction that gives an uncertainty estimation of the atmospheric part in the modeling chain, and (iii) providing a computationally inexpensive indicator (no extra computing time needed) which is physically easy to interpret.

This chapter is organized as follows. The data used for the calculation of the indices are described in Section 6.2. A detailed description of the transport index and the Pasquill indices is given in Section 6.3. The results are discussed in Section 6.4 and conclusions are given in Section 6.5.

6.2   Data

The different datasets used for the calculation of the transport- and Pasquill indices are described in Table 6.1.

Acronym in the text and figures	Description	DJF Period
Present climate		
OBS	Station observations. Used as reference for the Pasquill indices.	1990/91-1998/99 1981/82-2009/10
ERAINT	ALARO-0 model data driven by ERA-Interim reanalysis. Used as reference for the transport index.	1990/91-1998/99 1981/82-2009/10
CTL	ALARO-0 model data driven by CNRM-CM3.	1990/91-1998/99
Future climate		
SCN	ALARO-0 model data driven by CNRM-CM3.	2046/47-2054/55

Table 6.1: *Description of the different datasets.*

High-resolution model data from the ALARO-0 model at 4 km resolution is obtained by dynamically downscaling the ERA-Interim reanalysis (Dee et al., 2011) as well as GCM data from CNRM-CM3 (Salas-Mélia et al., 2005) to a 20-km model grid that encompasses most of western Europe. Our study periods are limited to the December-January-February (DJF) winter months. The reason for this is that transport index only applies for stable atmospheric conditions which mainly occur during the winter season. For the validation of the stability indices (see further Section 6.3), the 29-yr DJF present-day ERA-Interim study period is considered (1981/82-2009/10) as well as a 9-yr DJF subperiod hereof (1990/91-1998/99). The study period for the assessment of the future climate impact covers two times nine winter seasons: 1990/91-1998/99 and 2046/47-2054/55, representing the present-day control (CTL) and future scenario (SCN) climate conditions. In contrast to the previous chapter (Chapter 5) which used the last 30 years of the 21st century as study period, the future climate period in this chapter corresponds to a near future time period. Considering that the time horizon in the decision making process of the air quality targets rarely exceed the next decades, a near future study period is highly relevant.

In contrast to the previous chapters, the land-surface parameterization used for the ALARO-0 model simulations is not based upon the ISBA scheme (Interactions between Soil, Biosphere and Atmosphere) (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996). Instead, the ALARO-0 model is coupled to a new externalized land and ocean surface platform called SURFEX (SURface Externalisée; Masson et al., 2013). The coupling strategy relies on a simple interface to allow implicit coupling between the atmosphere and the tiled surface proposed by Best et al. (2004). To allow a further downscaling to an urban scale resolution of 1 km, the SURFEX land surface modeling system is employed in offline mode using the forcing coming from the lowest model level of the 4-km simulations. To account for the simulation of the interactions with urban areas, the TEB scheme, was switched on. Further details on SURFEX and the downscaling approach are described in Chapter 2 and

Hamdi et al. (2015).

The ERA-Interim driven runs have used the ACRANEB radiation parameterization scheme. To ensure a consistent use of the radiation scheme within the driving GCM CNRM-CM3 and the ALARO-0 model, the CTL and SCN model integrations for this chapter were performed with another radiation scheme than in previous chapter (i.e. the European Centre for Medium-Range Weather Forecasts (ECMWF) Fouquart-Morcrette Radiation (FMR) scheme). Similar to previous chapter, the CTL and SCN model integrations are performed according to the IPCC SRES A1B scenario. However, in previous chapter the sensitivity of the climate to external forcings was quantified with an equivalent GHG forcing value given by one value for equivalent CO_2, the use of the FMR scheme allowed us here to specify the radiative forcing by the concentration evolution of the GHG components separately. More details on the SRES A1B scenario and the radiation schemes are given in the previous chapters, Chapter 5 and Chapter 2, respectively.

The observation dataset which is used as a reference for the validation of the Pasquill indices (see further Section 6.3.2) consists of hourly SYNOP observations of cloudiness and wind speed as well as global solar radiation measurements performed at the station in Uccle (longitude: $4.358°E$, latitude: $50.798°N$). The observation data have undergone a thorough quality control.

6.3   Methodology

6.3.1   Transport index

The transport index gives a measure of the horizontal and vertical transport of nonreactive pollutants in stable atmospheric conditions and has been proposed and tested by Termonia and Quinet (2004). It gives a characteristic length scale l which is the ratio of the mean horizontal wind speed $\overline{u}$ and the Brunt-Väisälä frequency ν:

$$l = \frac{\overline{u}}{\nu} \tag{6.1}$$

where

$$\nu = \sqrt{\frac{g}{\theta}\frac{\partial \theta}{\partial Z}} \tag{6.2}$$

is a measure for the stability of the atmosphere, where Z is the geopotential height, g is the gravity, and θ is the potential temperature. This physical relation given by Eq. (6.1) shows that low values for l correspond to weak horizontal transport (i.e. small $\overline{u}$) and weak vertical transport (i.e. high ν). Hence when l reaches its lowest values in the lower part of the boundary layer during an extended period of several hours, one can be sure that these calm situations with a weak horizontal wind and a very stable atmosphere indicate conditions that are unfavorable for the dispersion of air pollution (Termonia and Quinet, 2004).

6.3.2 Pasquill stability indices

The stability scheme that is used to determine the Pasquill indices has been adopted from Van Der Auwera (1991a,b) (Table 6.2).

Wind speed (m s^{-1})	METEOROLOGICAL DAY Global radiation (W m^{-2})				METEOROLOGICAL NIGHT Cloudiness N (okta)	
	Strong]600-...]	Moderate]300-600]	Slight]150-300]	Weak [0-150]	$N \geq 4/8$	$N < 4/8$
≤ 1.5 m s^{-1}	A	A	B	C	F	F
≤ 2.5 m s^{-1}	A	B	C	C	E	F
≤ 3.5 m s^{-1}	B	B	C	D$_d$	E	E
≤ 4.5 m s^{-1}	B	C	C	D$_d$	D$_n$	E
≤ 5.5 m s^{-1}	C	C	D$_d$	D$_d$	D$_n$	D$_n$
≤ 8.0 m s^{-1}	C	D$_d$	D$_d$	D$_d$	D$_n$	D$_n$
> 8.0 m s^{-1}	D$_d$	D$_d$	D$_d$	D$_d$	D$_n$	D$_n$

Table 6.2: *Pasquill stability scheme with definition of the stability classes ranging from the very unstable atmospheric conditions (A) to the most stable atmospheric conditions (F) (adopted from Van Der Auwera (1991a,b)).*

Pasquill (1961) defined six stability classes ranging from the very unstable atmospheric conditions (A) to the most stable atmospheric conditions (F). The classification depends on the global solar radiation G during the day or the cloudiness N during the night, combined with the wind speed w at 10 m. The D stability class is defined for the meteorological night (i.e. D$_n$) and meteorological day (i.e. D$_d$) separately. The thresholds for the different categories of radiation (strong, moderate, slight and weak) as defined by Van Der Auwera (1991b) are based on other quantitative definitions found in the scientific literature. Furthermore, the meteorological day has been defined from one hour after sunrise to one hour before sunset and thus varies throughout the year. Here we have used the timings for sunrise and sunset of the year 2013.

6.3.3 Data processing and analysis

All meteorological variables that are used for the calculation of both indices are listed in Table 6.3. For each variable it is also indicated whether the variable is given as direct output from the model (i.e. base variable) or whether other quantities have been used to derive the respectively meteorological variable.

In the ALARO-0 model the vertical variation of the base variables T_a, u, v and q serving as input for the transport index, is represented by dividing the atmosphere into 46 layers n. These layers are defined by the pressures at the interfaces between them (so-called "half-pressure-levels") (European Centre for Medium-Range Weather Forecasts, 2010):

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2}p_s \tag{6.3}$$

for $0 \leq k \leq n$, with $n = 46$ corresponding to the lowest model level. The $A_{k+1/2}$ and $B_{k+1/2}$ are constants whose values define the vertical coordinate and p_s is the surface pressure field. The "full-level" pressure p_k associated with each model level (middle of layer) can then be reconstructed from:

$$p_k = \frac{1}{2}(p_{k-1/2} + p_{k+1/2}) \tag{6.4}$$

Index	Variable (Unit)	Symbol	Derived from
	Air temperature (K)	T_a	Base variable
	Zonal wind speed (m s^{-1})	u	Base variable
	Meridional wind speed (m s^{-1})	v	Base variable
Transport index	Specific humidity	q	Base variable
	Surface pressure (Pa)	p_s	Base variable
	Pressure (Pa)	p	p_s
	Potential temperature (K)	θ	p, T_a
	Geopotential height (m)	Z	p, q, T_a
	Brunt-Väisälä frequency (s^{-1})	ν	θ, Z
	Horizontal wind speed (m s^{-1})	$\overline{u}$	u, v
Pasquill indices	Cloudiness (okta)	N	Base variable
	Global solar radiation (W m^{-2})	G	Base variable
	Wind speed at 10 m (m s^{-1})	w	Base variable

Table 6.3: *Overview of the meteorological variables used as input for the transport- and Pasquill indices.*

with $1 \leq k \leq n$ by using Eq. (6.3). The "half-level" geopotential is then calculated using the discrete analogue of the hydrostatic equation:

$$\phi_{k+1/2} = \phi_s + \sum_{j=k+1}^{n} R_d(T_v)_j \ln\left(\frac{p_{j+1/2}}{p_{j-1/2}}\right) \tag{6.5}$$

where ϕ_s is the geopotential at the surface, R_d is the gas constant of dry air (i.e. 287.058 J kg^{-1} K^{-1}) and T_v is the virtual temperature defined for each vertical level k by:

$$(T_v)_k \approx T_{a_k}[1 + 0.61q_k] \tag{6.6}$$

where T_a is the air temperature, q is specific humidity and R_v is the gas constant of water vapor. Full-level values of the geopotential height Z are then calculated by:

$$Z_k = \frac{\phi_{k+1/2} + \alpha_k R_d(T_v)_k}{g} \tag{6.7}$$

where $\alpha_1 = \ln 2$ and for $k > 1$:

$$\alpha_k = 1 - \frac{p_{k-1/2}}{\Delta p_k} \ln\left(\frac{p_{k+1/2}}{p_{k-1/2}}\right). \tag{6.8}$$

The potential temperature θ is obtained for each vertical level k by:

$$\theta_k = T_{a_k}\left(\frac{p_0}{p_k}\right)^{R/c_p} \tag{6.9}$$

with $p_0 = 100\,000$ Pa being the standard reference pressure, R is the universal gas constant of air (i.e. 8.314 J mol^{-1} K^{-1}), and c_p is the specific heat capacity at constant pressure (i.e. 29.070 J mol^{-1} K^{-1}). Finally, the horizontal wind speed $\overline{u}$ is calculated for each vertical level k as:

$$\overline{u}_k = \sqrt{u_k{}^2 + v_k{}^2} \tag{6.10}$$

The transport index is calculated up to the model level corresponding to a maximum height of $\approx$ 2000 m, which roughly coincides with the height of the boundary layer. This layer of air near the ground that is affected by diurnal heat, moisture, and momentum transfer to/from the surface (Ebi and McGregor, 2008), can thus be considered as most relevant for the assessment of the dispersion of air pollutants.

As mentioned earlier, the impacts from air pollution are strongest felt in urban environments. Therefore, both indices are calculated for the closest model grid point to the station of Uccle, located some 6 km from the city centre of Brussels. As the transport index only applies for stable conditions which mostly occur during the winter season, our analyses focuses on nine DJF winter months. This 9-yr winter period, hence allows to assess the *climate* impact on meteorological conditions that are unfavorable for the dispersion of air pollution by means of frequencies of index values corresponding to stable atmospheric conditions.

6.4 Results and discussion

The transport index has previously been validated in Termonia and Quinet (2004) for three consecutive winter seasons (DJF) 2000/01-2002/03 using the output of 0-48-h forecasts from 0000 UTC runs of the ALADIN model. There are several differences in the experimental setup of the study of Termonia and Quinet (2004) and the experimental setup as used for our model simulations in this chapter. First, the model simulations used here are obtained with the ALARO-0 model at 4 km spatial resolution, which uses the 3MT physics package specifically developed for the convection permitting scales (see Chapter 2). The ALADIN model as used in in Termonia and Quinet (2004) at 7 km spatial resolution, is based on different physical parameterizations. Secondly, the coupling data used for the simulations in this chapter are the recent reanalysis from ERA-Interim, while Termonia and Quinet (2004) used an "old" forecast as initial state for the 48-h forecasts. This points to a last and third difference, which is related to the fact that our simulations are performed through a dynamical downscaling with daily reinitializations, instead of a continuous forecast mode which was applied in Termonia and Quinet (2004).

We can expect that the above differences in the experimental setup will affect the simulation of the stable boundary layers. Therefore, both the transport index as well as the Pasquill indices are in a first step validated for the present-day winter (DJF) periods 1990/91-1998/99 and 1981/82-2009/10 (Section 6.4.1). In a next step, the future climate impact on both stability indices is assessed for the near-future DJF period 2046/47-2054/55 (Section 6.4.3).

6.4.1 Validation for present-day climate

6.4.1.1 Transport index

For the validation of the transport index we consider the profiles of l obtained from the ERAINT dataset as a reference. The reason why we do not use observed profiles of l calculated from observations of radio-sounding profiles is simply because such profiles are usually only available with time intervals of 12 or 24 h, which does not allow to validate l for consecutive periods up to e.g. 12 hours (Termonia and Quinet, 2004).

Similar to Termonia and Quinet (2004), we validate our transport index with the measured concentrations of NO_2 for the three consecutive DJF seasons 2000/01-2002/03.
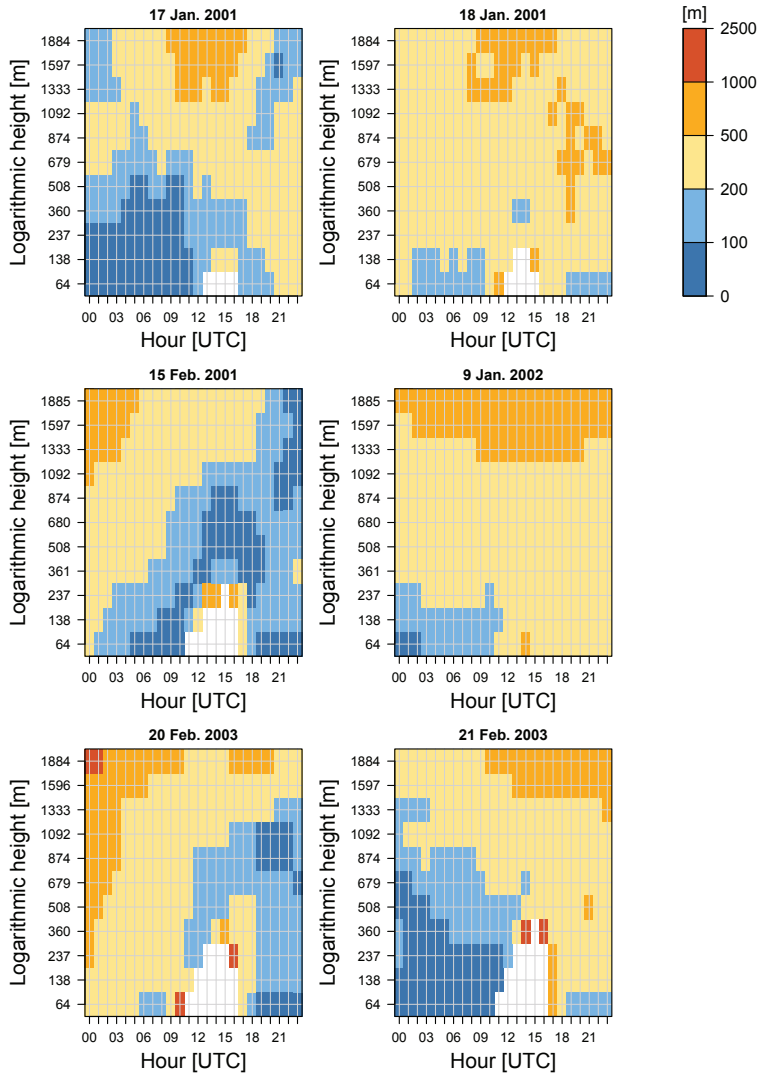
Figure 6.4: *Transport length values on 5 days when an extreme pollution peak has been recorded during the winter season DJF 2000/01-2002/03 (i.e. 17 January 2001, 18 January 2001, 15 February 2001, 9 January 2002, and 20-21 February 2003). The transport index is calculated from the ERA-Interim (ERAINT) dataset for the closest model grid point to Uccle. The white areas indicate unstable parts of the atmosphere where the Brunt-Väisälä frequency is not defined.*

Termonia and Quinet (2004) defined a pollution peak as extreme when the hourly value of observed NO_2 concentrations exceeded the value of 135 μg m^{-3} in at least two-third of the measurement stations in the Brussels Capital Region. For improving health protection and for longterm protection of the environment, this guide value of 135 μg m^{-3} has been

prescribed by the European Directive 85/580/EEC as the 98th percentile of the hourly mean values of the concentrations of NO_2, measured during the respectively calender year (European Community, 1985). In the more recent Daughter Directive 1999/30/EC (European Community, 1999) or Directive 2008/50/EC on ambient air quality and cleaner air for Europe (European Community, 2008), this guide value is not included anymore. Both directives solely provide a limit value of 200 $\mu g\ m^{-3}$ for the hourly values that can be exceeded only 18 times per year. As stated by Termonia and Quinet (2004), this limit value is exceeded only rarely in the Brussels Capital Region. Therefore, we keep similar to Termonia and Quinet (2004) the same definition for extreme pollution peaks based upon the old guide value of 135 $\mu g\ m^{-3}$. With this definition, the authors found an exceedance of the guide value for observed NO_2 concentration on 5 days: (i) 17 January 2001, (ii) 18 January 2001, (iii) 15 February 2001, (iv) 9 January 2002, and (v) 20-21 February 2003.

For these 5 days, the calculated profiles of l obtained from the ERAINT dataset are shown in Fig. 6.4. For most days, there appear instabilities (i.e. white areas) in the stable layers near the surface where the Brunt-Väisälä frequency is not defined. These instabilities are the result of heating during the day, and can be ignored if they are covered from above with a layer of small transport length values (Termonia and Quinet, 2004). The profiles of l, which correspond to days with extreme pollution peaks of observed NO_2 concentrations, allow us to define a criterion for the conditions that are *unfavorable* for the dispersion of nonreactive pollutants.
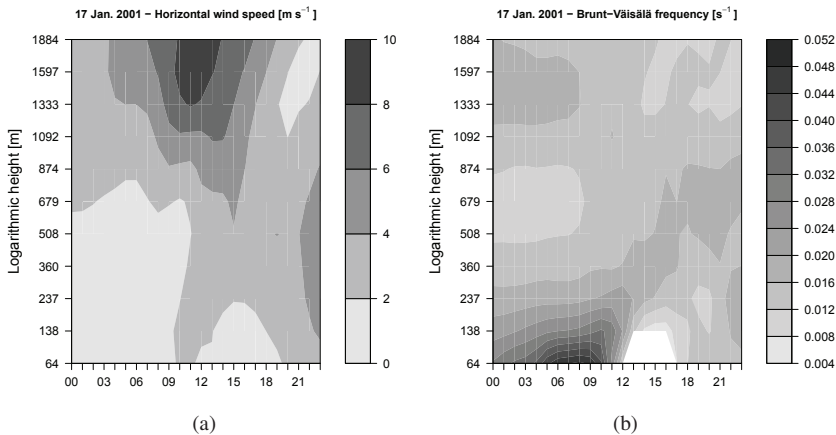
Figure 6.5: *Vertical profile of the horizontal wind speed ($m\ s^{-1}$) and Brunt-Väisälä frequency (s^{-1}) on 17 January 2001, for the closest model grid point to Uccle from the ERA-Interim (ERAINT) dataset. The white areas indicate unstable parts of the atmosphere where the Brunt-Väisälä frequency is not defined.*

These conditions are generally determined by (i) the value of the transport length l, (ii) the height of the stable layer, and (iii) the time span or duration. Hence, we consider conditions as unfavorable for the dispersion of air pollutants, when for a minimum duration of 9 h, a layer at the surface with transport length values $l < 200$ m, reaches a minimum height of 100 m. In this way, 4 out of the 5 days with observed extreme pollution peaks,

fulfill the criterion for unfavorable conditions for the dispersion of air pollutants.

As an illustration, Fig. 6.5 shows the horizontal wind speed $(\mathrm{m\,s^{-1}})$ and Brunt-Väisälä frequency $(\mathrm{s^{-1}})$ for 17 January 2001, as obtained from the ERAINT dataset. As expected, the low l values in the lower surface layers are mainly determined by low horizontal wind speed values and stable atmospheric conditions (i.e. high Brunt-Väisälä frequencies).

Figure 6.6 shows the absolute frequencies of the transport index for values lower than 200 m from ERAINT for the 9-yr winter period 1990/91-1998/99.
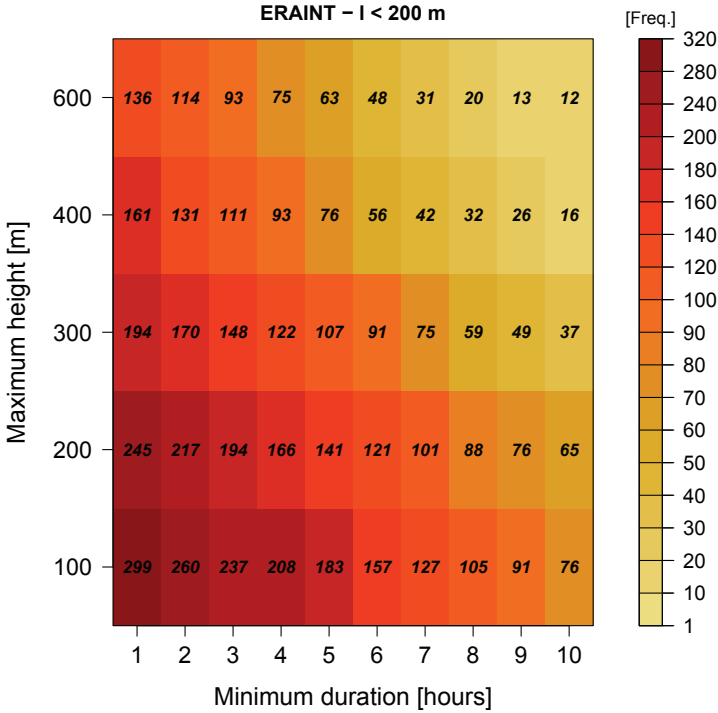
Figure 6.6: *Absolute frequencies of transport length values $l < 200$ m from the ERA-Interim driven ALARO-0 simulation (ERAINT) for the closest model grid point to Uccle. The frequencies are calculated for the 9-yr DJF present climate period 1990/91-1998/99 and are shown for maximum heights ranging between 0 and 600 m, and minimum durations ranging between 1 and 10 h.*

These frequencies, corresponding to cases with calm wind situations and a very stable atmosphere, are shown accordingly to our defined criterion, i.e. for different durations up to 10 hour and vertical heights between 100 and 600 m. It can be seen that the number of cases per year in ERAINT that meets the criterion for the 9-yr study period (i.e. 76 cases) corresponds to 8-9 cases per winter season. This number gives an indication of the "order of magnitude" of the number of cases of very stable atmospheric conditions or extreme pollution peaks, that can be expected according to our criterion under present climate winter conditions. However, this number of 8-9 cases per winter season is more than double of the number of observed extreme pollution peaks (i.e. 1-2 cases per winter season) as had been selected in Termonia and Quinet (2004). The reason for this difference can be ex-

plained by the experimental setup used for our model simulations, which differs from the one used in Termonia and Quinet (2004). As mentioned previously, Termonia and Quinet (2004) performed forecasts up to 48 h from 0000 UTC runs to assess the evolutions of the transport length. On the other hand, our procedure of dynamically downscaling for climate integrations uses daily reinitializations, where a 36-h run at a high spatial resolution of 4 km is performed. The first 12 h of this 36-h run are dismissed for spinup, to end up with 24 h of output (see also Chapter 2, Section 2.3.2 for more details on our procedure of daily reinitialization).

Hence, because of this daily reinitialization procedure, our algorithm cannot be applied over a continuous time span of 48 h. However, when it comes to the assessment of the evolution of transport length values throughout a time span of 48 h, it is clear that this procedure will not allow a build up of a stable layer for a sufficiently long period. Hence, over a separate duration of two times 24 h, our algorithm will detect two separate cases while in reality these correspond to one case which has been build up over a continuous time span of 48 h. To illustrate this issue, the profile of l for 16 January 2001 is shown in Fig. 6.7.
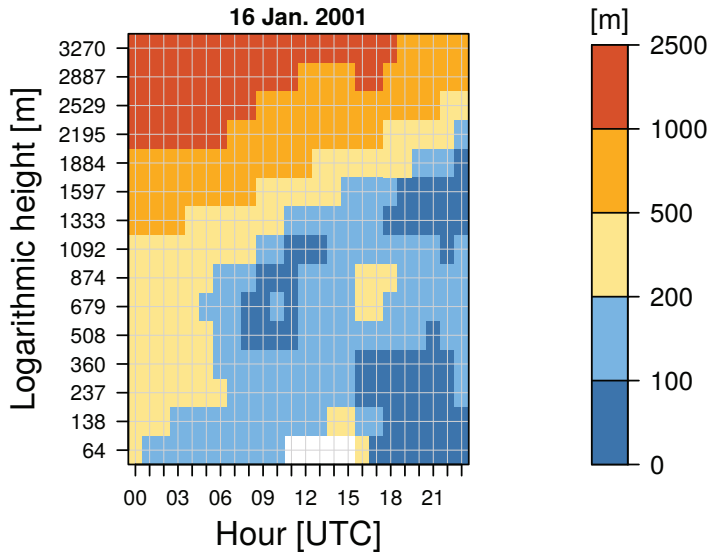
Figure 6.7: Transport length values for 16 January 2001. The transport index is calculated from the ERA-Interim (ERAINT) dataset for the closest model grid point to Uccle. The white areas indicate unstable parts of the atmosphere where the Brunt-Väisälä frequency is not defined.

As can be seen, over the course of the day a stable layer characterized by low transport length values is build up. These stable conditions then persist throughout the night and morning of the following day (i.e. 17 January 2001), for which a extreme pollution peak had been observed. Overall, our results demonstrate that also with the new model version ALARO-0, the transport index can be used to detect peaks of extreme concentrations of pollutants such as NO_2.

6.4.1.2 Pasquill indices

Before validating the frequencies of Pasquill indices, it is first assessed how well the model reproduces the observed hourly frequencies of the meteorological variables that are used to obtain the indices.

Figure 6.8 presents barplots for 29-yr DJF (1981/82-2009/10) observed (OBS) and simulated (ERAINT) wind speed, global solar radiation and cloudiness.

(a)

(b)

(c)
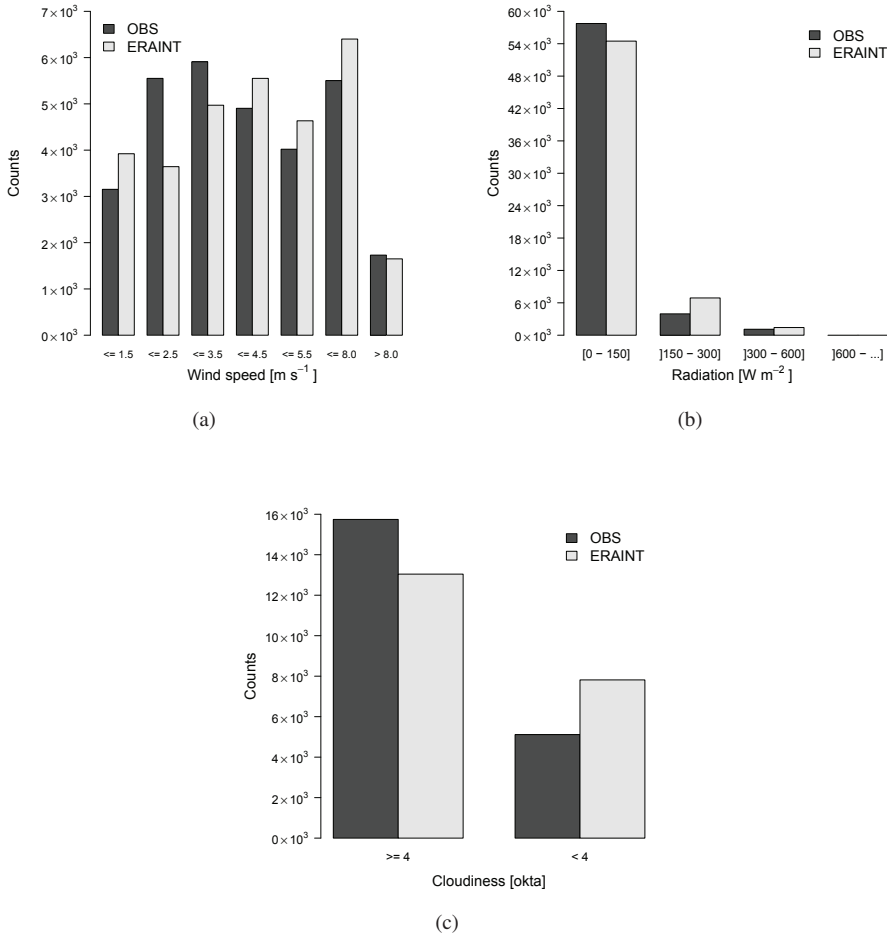
Figure 6.8: *Frequency distribution of (a) wind speed at 21 m ($m\,s^{-1}$), (b) global solar radiation ($W\,m^{-2}$), and (c) cloudiness (okta) for the 29-yr DJF period 1981/82-2009/10. Frequencies are calculated for observed (OBS) values in Uccle and modeled (ERAINT) values for the closest model grid point to Uccle, according to the thresholds given by the stability scheme that is used to determine the Pasquill indices (Table 6.2).*

The absolute frequencies are calculated according to the thresholds given by the stability scheme that is used to determine the Pasquill indices (Table 6.2). To ensure a fair comparison between ERAINT and OBS frequencies, missing values within the observed

time series are also set on missing in the ERAINT dataset. Furthermore, it is important to note that wind speed measurements at the station of Uccle are performed at a height of the mast that corresponds to 21 m. Therefore, the simulated ERAINT wind speed values are also obtained for the model level corresponding to a height of 21 m, and thus the frequency distribution for wind speed as shown here are obtained with wind speed values at 21 m instead of 10 m. The observed wind speed frequencies are relatively well reproduced by ERAINT. Only for the wind speed (w) classes]1.5-2.5] $\mathrm{m\,s^{-1}}$ (]5.5-8.0] $\mathrm{m\,s^{-1}}$) the model strongly underestimates (overestimates) the number of observed wind speed hours (Fig. 6.8(a)). Since we are only considering the winter season, both the observed and modeled distributions of global solar radiation (G) are positively skewed towards the lowest radiation values (i.e. [0-150] $\mathrm{W\,m^{-2}}$). The number of hours which are underestimated by the model for the lowest radiation values are mostly shifted towards the next radiation class, for which the model displays an overestimation. A similar feature is apparent for the frequency distribution of cloudiness (N); the model underestimates the cloudy hours (i.e. ≥ 4 okta) and shows a tendency for too many hours with cloudiness values < 4 okta. However, if frequencies are obtained by binning the hourly cloudiness into bins of 1 okta, it is found that the underestimation by ERAINT for the cloudiness values ≥ 4 okta, is mainly due to a strong underestimation by the model for the overcast cases (see Appendix D, Fig. D.1). This underestimation by ALARO-0 in the occurrence of overcast conditions has also been detected in a previous study of Hamdi et al. (2012). As indicated by Hamdi et al. (2012), the diagnostic of total and partial cloud cover (low, medium, high, and convective) is computed in the model with two options; (i) random overlap of adjacent clouds assumption, and (ii) maximum overlap of adjacent clouds. Hence, when using the maximum overlap assumption, the occurrence of cloud covers near 100% is underestimated with respect to the observed frequencies. Taking into account that differences between observed and modeled frequencies are strongly dependent on the threshold values, we can say that the model reproduces the observed frequencies of w, G, and N fairly well.

The frequencies of Pasquill indices as obtained from ERAINT and station observations for Uccle are presented in Fig. 6.9. Missing values within the observed time series are again also set on missing in the ERAINT dataset. The resulting hourly frequencies of Pasquill indices for the winter period 1990/91-1998/99 (Fig. 6.9(a)) consists of $\approx 67\%$ of missing data mainly due to missing observations of cloudiness. Both the observed and ERAINT distribution show the highest frequencies for the neutral D_d and D_n indices, followed by the stable E and F indices. During the winter season these stable E and F classes are mostly relevant, as they reflect stagnant conditions with few dispersion of potentially present air pollutants. The model (ERAINT) slightly under- and overestimates the frequencies in E and F indices. To assess the impact of the great number of missing data on the validity of this result, the frequencies of Pasquill indices are also calculated for the winter period 1981/82-2009/10 (Fig. 6.9(b)). Also for this longer winter period the ERAINT distribution coincides fairly well with the observed distribution, with acceptable relative differences between the frequencies of the relevant and stable Pasquill indices E and F for ERAINT and OBS.
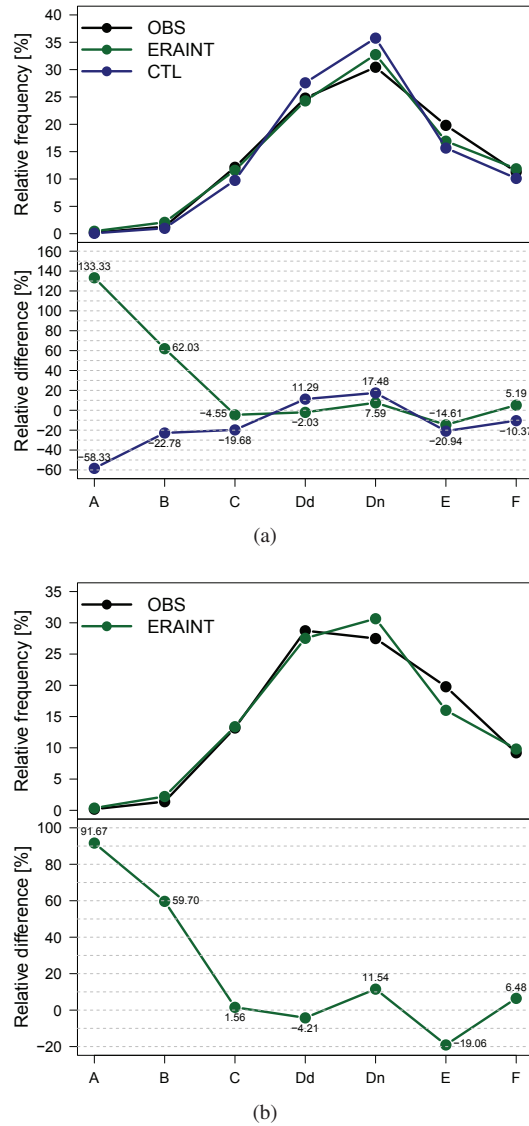
Figure 6.9: *Relative frequencies and differences of observed (OBS, black) and modeled (ERAINT, darkgreen and CTL, darkblue) Pasquill indices in Uccle and its closest model grid point for DJF (a) 1990/91-1998/99 and (b) 1981/82-2009/10. The numbers indicate relative differences between ERAINT and CTL w.r.t. OBS.*

6.4.2 Bias correction

Previous results have demonstrated that stability indices obtained from the downscaled ERA-Interim reanalysis with the ALARO-0 model are a useful tool to infer meteorological conditions that are unfavorable for extreme pollution peaks. Before future changes in oc-currences of the transport- and Pasquill corresponding to such unfavorable meteorological

conditions are assessed, we investigate how well the stability indices derived from present-day downscaled GCM CNRM-CM3 fields (CTL) are reproduced by ALARO-0 w.r.t. to the reference datasets.

The relative frequencies of Pasquill indices obtained from the control simulation are given in Fig. 6.9(a). To ensure a consistent comparison, the frequencies are also derived with the wind speed for the model level corresponding to a height of 21 m. Compared to the relative differences between the frequencies obtained from the ERAINT driven fields and the observed frequencies, the biases between the CTL and OBS frequencies are slightly larger. The introduction of such additional biases when one couples a GCM to a RCM instead of reanalyses is a well known issue in regional climate modeling (see also previous Chapter 5). Menut et al. (2013) evaluated the changes in air quality-related weather variables induced by replacing reanalyses-forced (ERA-Interim) by GCM-forced regional climate simulations. The authors investigated, amongst several meteorological variables that are critical for air quality modeling, two variables which are used here for the Pasquill classification; i.e. the wind speed and short-wave radiation, which are essential for respectively dispersion and photochemistry. Their results demonstrated that in winter, air quality is mostly driven by dispersion, and that no significant differences in wind statistics between GCM-driven and reanalyses-driven regional simulations could be identified (Menut et al., 2013). Although, a different set of models is used here, this finding supports to decide to not apply a bias correction on the Pasquill indices. Furthermore, the overall CTL distribution follows relatively well the observed one, with biases for the relevant E and F classes that vary in an acceptable range.

Figure 6.11(a), Fig. 6.11(c), and Fig. 6.11(e), show respectively the frequencies of low transport length values l calculated from CTL and its absolute and relative differences with ERAINT for the 9-yr DJF period 1990/91-1998/99. It can be seen that the CTL frequencies substantially differ from ERAINT (Fig. 6.6). The CTL ALARO-0 simulation systematically underestimates the number of cases characterized by transport values smaller than 200 m, and one can thus expect to have more transport of pollutants during the winter with CTL. These large differences in the number of $l < 200$ m can be attributed to biases in the driving GCM CNRM-CM3 that are reflected in the downscaled meteorological fields for CTL that are used as input for the transport index calculation. However, as mentioned previously in the introduction (Section 6.1), applying a bias correction on the base meteorological variables (T_a, u, v, q and p_s) would unavoidably alter the physical relationships on which the transport index is based upon. Nevertheless, in order to estimate future frequencies of low transport length values, these differences between the reference ERAINT and CTL transport length values should be addressed. Therefore, we perform a bias correction directly on the transport length values from CTL, which takes into account the differences between the CNRM-CM3 downscaled and ERA-Interim downscaled values. As we are interested here in the tail (i.e. lowest values) of the transport length distributions, Déqué (2007) proposes a variable correction method to be a suitable method. In this variable correction method a particular function f is build with the observation dataset (or in our case the ERAINT dataset) and applied on the CTL dataset as well as on the future climate SCN dataset (Déqué, 2007). By doing so, we assume that the model bias in the future climate stays the same as in the present-day climate.

A linear Quantile-Quantile (Q-Q) function is used as correction function $f(x)$. Figure 6.10 shows the Q-Q plot for downscaled transport length values using ERAINT versus CNRM-CM3 (i.e. CTL) for the DJF period 1990/91-1998/99.
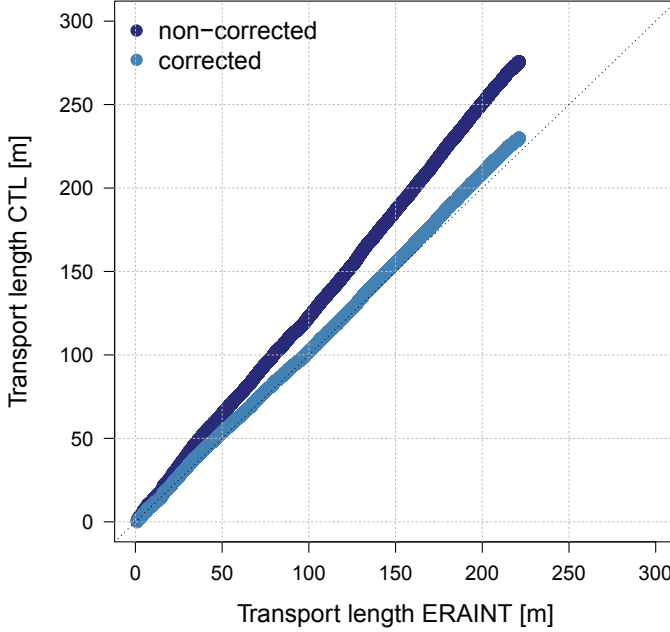
Figure 6.10: *Quantile-Quantile (Q-Q) plot showing downscaled transport length values for the closest model grid point to Uccle using ERAINT versus CNRM-CM3 (i.e. CTL) for the DJF period 1990/91-1998/99. The bias-corrected CTL transport length values are shown in light blue. The dashed line represents the perfect model (1:1 line).*

As we are only interested in the height-duration-frequency of low transport length values smaller than 200 m, quantiles are not plotted for values larger than 300 m. As expected, CTL clearly overestimates the ERAINT quantiles, and this overestimation gets larger as the transport length increases. It can be seen that the Q-Q curve is a straight line, suggesting that a linear Q-Q correction is appropriate here. The following linear correction function is applied on the CTL transport length values:

$$l_{corr} = 0.84 * l_{non-corr} - 0.69 \tag{6.11}$$

The intercept of the linear correction function [Eq. (6.11)] is chosen in such a way that after correction no negative and nonphysically transport length values are present. After the bias correction the results improve significantly. The corrected CTL transport length values are indeed much closer to the perfect model line (1:1 line), and this for all quantiles ranging between 0 and 200 m (light blue quantiles in Fig. 6.10). This result, where we have used the Q-Q plot as a correction function, suggests that the ALARO-0 model driven by CNRM-CM3 global data is now able to predict a ranked category of transport length but not the exact value for this variable (Déqué, 2007).
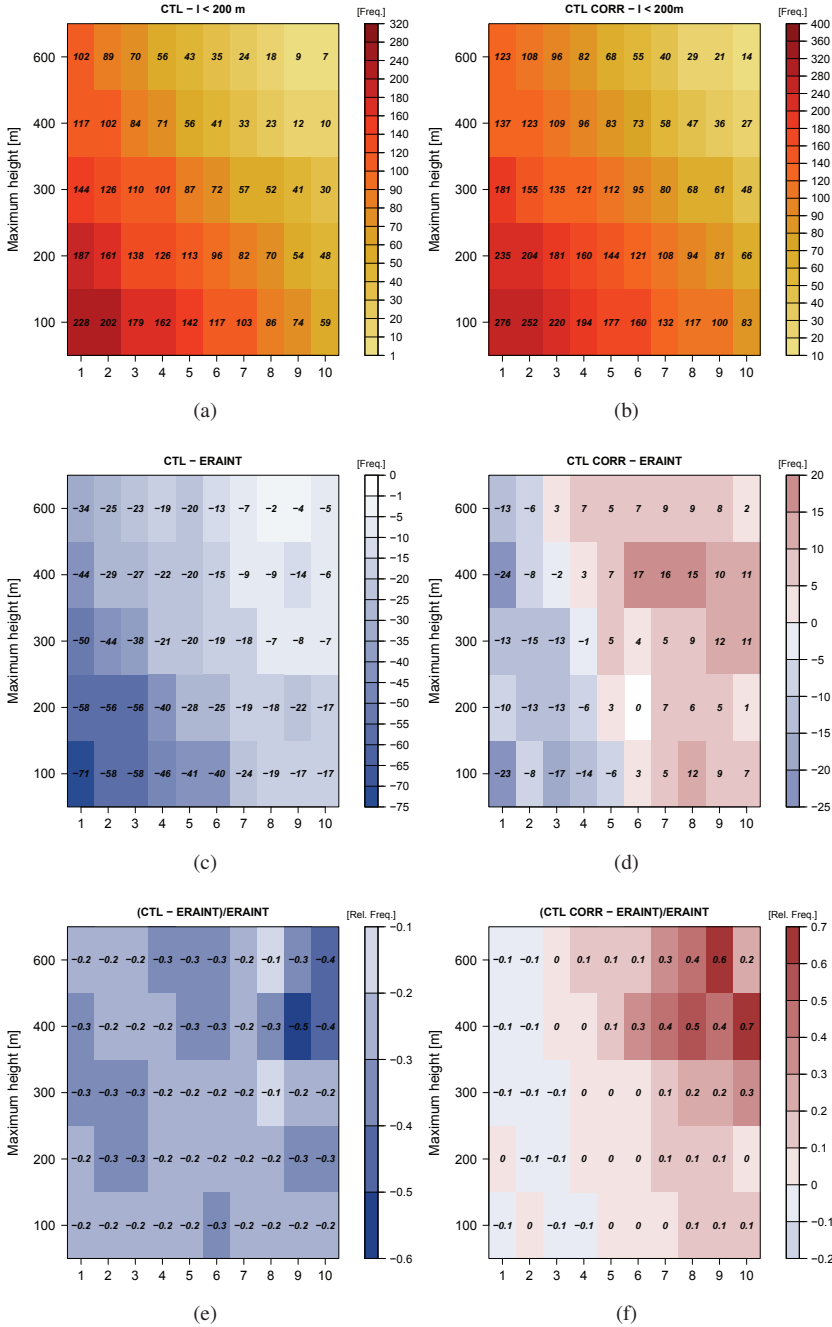
Figure 6.11: *Absolute frequencies of transport length values l < 200 m from (a) the control (CTL) and (b) bias corrected control simulation (CTL CORR), and absolute and relative differences in frequencies between (c, e) CTL and ERAINT and (d, f) CTL CORR and ERAINT. The frequencies and differences are calculated for the 9-yr DJF present climate period 1990/91-1998/99 for the closest model grid point to Uccle, and they are shown for maximum heights ranging between 0 and 600 m, and minimum durations ranging between 1 and 10 h.*

The improvement after correction is also reflected in the height-duration-frequency plots, as displayed in Fig. 6.11. Compared to the non-corrected CTL frequencies, the corrected CTL frequencies of the transport length values < 200 m are now for all durations and heights higher and the absolute and relative differences with the ERAINT frequencies are remarkably smaller. However, it seems that for the longer durations and for the larger heights the Q-Q correction "overcorrected" the CTL transport lengths, while for the shortest durations and for the lowest heights the frequencies are still slightly underestimated by CTL. This may be explained by the fact that the correction method based on Q-Q plot does not correct for the temporal properties of the series (Déqué, 2007), suggesting that this correction is not sufficient for postprocessing frequencies based upon duration. Although a more advanced and non-linear correction might be more appropriate, the bias has been significantly reduced after correction and has an acceptable maximum magnitude of 10% for the most relevant heights of up to 200 m. As there is even after correction, a considerable overestimation in frequencies for the longest durations and largest heights, we do not give much confidence to the results found here and keep in mind that these results are likely to be an upper boundary of (future) frequencies of low transport length values.

6.4.3   Future impact

The possible future changes in the occurrences of the transport- and Pasquill indices, which have been found to provide a confident measure for the dispersion of air pollutants, is explored in the following section. Figure 6.12 presents absolute frequencies of $l < 200$ m derived from the scenario simulation (SCN), together with its absolute and relative differences w.r.t. the control simulation (CTL). The bias corrected SCN frequencies and potential future differences are also shown (Fig. 6.12(b), Fig. 6.12(d), and Fig. 6.12(f)). These corrected SCN frequencies are obtained by applying the same Q-Q based correction as used for the present-day climate to the SCN dataset [Eq. (6.11)]. As mentioned previously, by doing so it is assumed that the model bias in the future will be similar to the bias in the present-day climate. As shown in Fig. 6.12, the future response for both the non-corrected and corrected data shows a systematic increase in frequencies of low transport length values < 200 m.

According to our criterion, our results suggest that an increase of 60 to 70 % in the conditions which are unfavorable for the dispersion of air pollutants can be expected in Brussels by the middle of the 21st century under the A1B SRES scenario. This consistent future climate response as obtained from postprocessed data, in complement with the non-corrected data, narrows the uncertainty range on our potential future impact and thus increases the confidence in the finding of a positive future signal. Furthermore, these results are also supported by the shift towards higher relative frequencies of the stable Pasquill indices E and F in winter for the future scenario period (SCN) (Fig. 6.13). Although, the increase in the relative frequency of E and F Pasquill classes (i.e. $\approx 30\%$) is lower than the 60 - 70 % increase in the frequencies of low transport length values. This can be explained by the fact that the positive response of the E and F Pasquill classes reflects a *general* increase towards more stable cases, while according to our criterion, the changes in the transport length frequencies correspond to an increase of the most *extreme* stable cases.
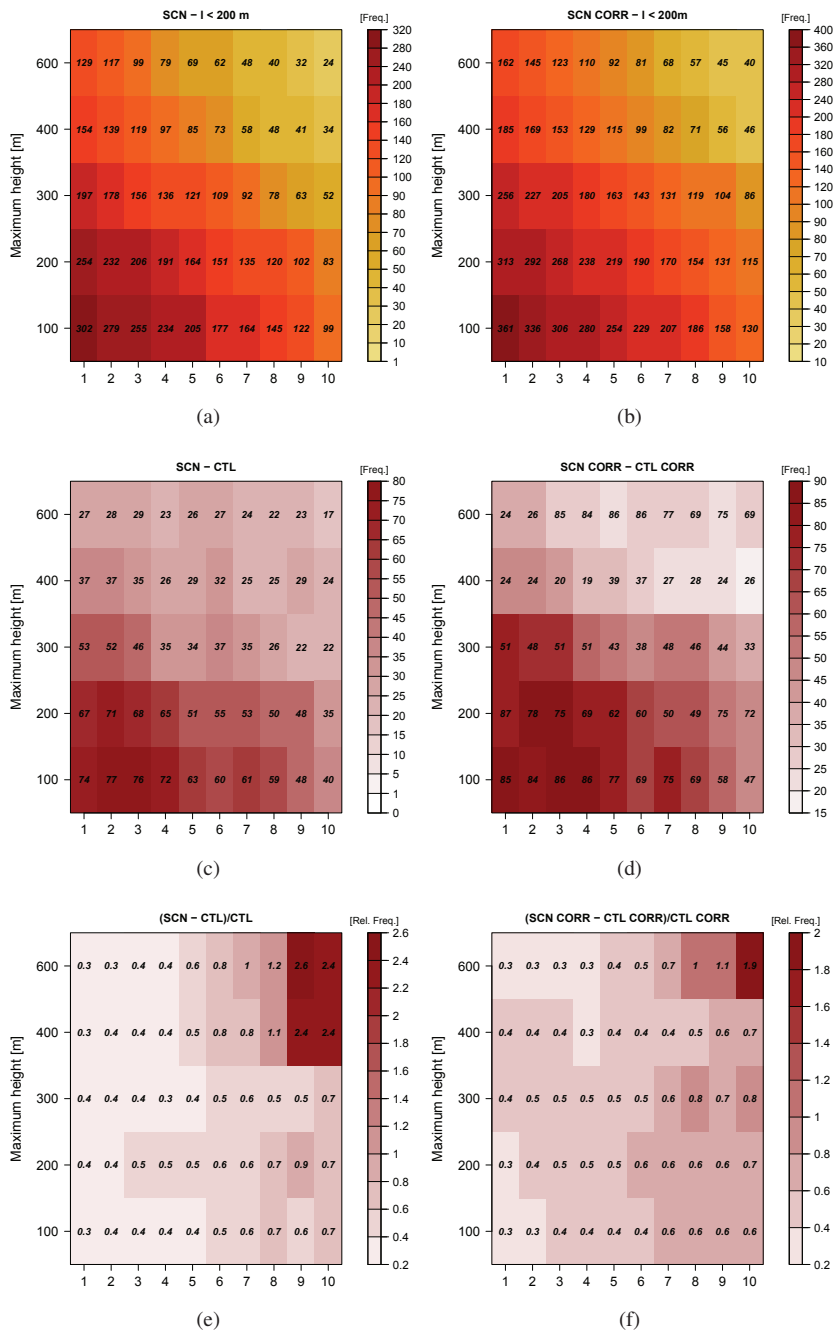
Figure 6.12: *Absolute frequencies of transport length values l < 200 m from (a) the scenario (SCN) and (b) bias corrected scenario simulation (SCN CORR), and absolute and relative differences in frequencies between (c, e) the SCN and the control (CTL) simulation and (d, f) the SCN CORR and the bias corrected control simulation (CTL CORR). The frequencies and differences are calculated for the 9-yr DJF CTL climate and future climate period 1990/91-1998/99 and 2046/47-2054/55 for the closest model grid point to Uccle, respectively, and they are shown for maximum heights ranging between 0 and 600 m, and minimum durations ranging between 1 and 10 h.*

Our results, showing an overall positive tendency towards more stable conditions, are in agreement with the findings of previous studies. As reported by Jacob and Winner (2009), there is a consensus among GCMs that 21st-century climate change will increase the frequency of stagnation episodes over northern mid-latitudes continents. The authors relate this increase in stagnation to the weakening of the general circulation and a northward shift of the mid-latitude cyclone tracks, which decreases the frequency of cold fronts that are the principal ventilation mechanism for eastern North America, Europe, and East Asia (Jacob and Winner, 2009). Furthermore, Giorgi and Meleux (2007) assessed the regional effects of climate change on air quality by analyzing outputs of climatological variables from the Coupled Model Intercomparison Project Phase 3 (CMIP3) dataset for the A1B emissions scenario. The authors analyzed ensemble mean changes for the period 2071-2100 w.r.t. to the reference period 1961-1990 for climatological variables such as temperature, precipitation and sea-level pressure, which are important drivers in the distribution of pollutants. For winter (DJF), the authors also find a pronounced increase in sea-level pressure over the Mediterranean region and central Europe, which are indicative of greater subsidence and stagnant conditions, and which indeed inhibit the dispersion of pollutants (Giorgi and Meleux, 2007).
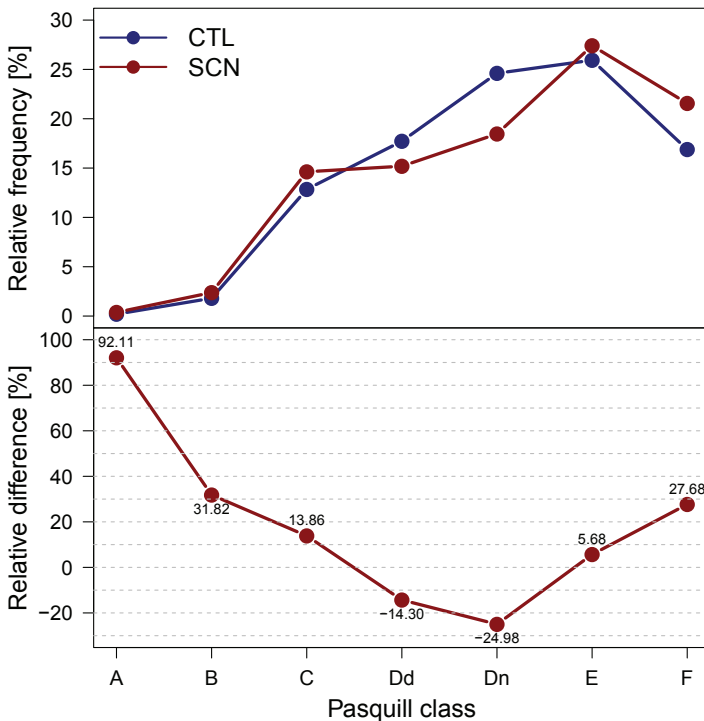
Figure 6.13: *Relative frequencies and differences of control (CTL, darkblue) and scenario (SCN, darkred) Pasquill indices for the closest model grid point to Uccle for respectively DJF 1990/91-1998/99 and 2046/47-2054/55. The numbers indicate relative differences between SCN and CTL.*

Finally, this tendency in more stagnant episodes and consequently increased pollutant

concentrations, can also be associated to changes in the meteorological parameters used to derive the transport length values and Pasquill classes. Frequency distributions of wind speed at 10 m, radiation and cloudiness demonstrate a shift towards lower wind speeds, lower cloudiness and consequently higher solar radiation (see Appendix D, Fig. D.2). This decrease in 10-m wind speed and the increased number of clear-sky days under A1B future climate conditions in Brussels have also been reported in a recent study of Hamdi et al. (2015). Furthermore, similar changes in these parameters are found by Katragkou et al. (2011), who studied future changes in summer surface ozone from regional climate-air quality simulations over Europe for two future decades, 2041-2050 and 2091-2100 under the A1B scenario and the control decade 1991-2000. The authors find that the combination of relatively more stagnant conditions associated with a strong decrease of wind speed, decrease of cloudiness, and increased temperatures and solar radiation within an anticyclonic anomaly, favors an enhanced ozone production in the western part of Europe, where mean surface summer ozone increases mostly (Katragkou et al., 2011). Despite the agreement of the future response in wind speed, radiation and cloudiness as found in our simulations and the other reported studies, it is important to stress that this result can only be interpreted as a sensitivity of the ALARO-0 model to the two different sets of CTL and SCN boundary conditions. For a more certain statement on the future changes, a multi-model approach which allows to quantify the uncertainty around the results is needed.

6.5 Conclusion

In this chapter, we analyzed the impact of climate change on unfavorable meteorological conditions for the dispersion of air pollution associated with winter smog peaks in Brussels (Uccle). For this, we have calculated two stability indices for a 29-yr and 9-yr winter period using present (1981/82-2009/10 and 1990/91-1998/99) and future (2046/47-2054/55) climate data that has been obtained from a dynamically downscaling of GCM data from the ERA-Interim reanalysis as well as from the coupled CNRM-CM3 model using the ALARO-0 model at 4 km spatial resolution.

In a first step, we have assessed whether the ERA-Interim driven ALARO-0 model is able to reproduce observed cases of winter smog alerts by means of the transport length values as well as observed frequency statistics of the Pasquill stability indices. We defined a criterion of unfavorable conditions for the dispersion of nonreactive pollutants, and it was found that the model reproduces almost all the observed smog peaks under consideration. In addition, the simulated frequency distribution of Pasquill indices also showed a close agreement to the observed one.

When one replaces the reanalysis data providing the boundary fields for the RCM with driving fields from a coupled GCM, it is well known that additional model biases coming from the GCM are often introduced. Comparison of frequency statistics of the transport- and Pasquill indices obtained from the control simulation with the observed frequencies indeed revealed significant biases, and in particular for the transport index. To account for these model imperfections, a linear Q-Q bias correction has been applied directly on the transport length values. After correction, the present-day frequencies are significantly improved, with remaining deviations from the observations that fall within an acceptable range of 10%.

In order to quantify the uncertainty on the future changes, the same linear correction method as used to correct present-day transport length values, has been applied to correct the future transport lengths. In line with previous studies, our results suggest a consistent increase in frequencies of low transport length values as well as stable Pasquill classes under future A1B climate conditions, reflecting a tendency towards more stable conditions and a possible degradation of air quality during winter smog episodes. The confidence of these results is supported by the consistent response that is found between corrected and non-corrected scenario results. Nevertheless, it is important to keep in mind that the assumption of constant bias under present and future climate conditions may not be valid, and that a more advanced bias correction method could be more appropriate. Furthermore, it should be taken into account that the future climate change impact on winter smog episodes as found here, is based upon one single emission scenario from the downscaling results of one RCM. Our assessment has been made under the consideration that the actual concentration peaks of the relevant winter smog pollutants are superimposed on their background concentrations, and hence of a much larger order of magnitude than the respectively background concentrations. However, in this consideration, future background concentrations and emissions of GHGs and other pollutants, remain an important source of uncertainty, due to the unknown changes in population vulnerability and human activity patterns (Ebi and McGregor, 2008). In this respect, a multi-model approach is a promising strategy to narrow the uncertainty on the projected model results.

Another important limitation in our assessment of the climate change impact on winter smog episodes, is the relatively short study period of nine winter seasons. Such short periods may indeed bias the results since on the decadal time scales the climate change signal is small compared to natural variability (Maraun et al., 2010). As a future outlook our methodology will be extended for a longer study period of 30 years, and applied to other European cities, using IPCC's most recent RCP emission pathways.

Finally, it is important to highlight the relevance and innovative strength of the methodology of stability indices as proposed in this chapter. In practice, it is generally hard to evaluate the skill of CTM results because of the limited availability of observational data for the evaluation (Menut et al., 2013). To our knowledge, no previous studies have focused on examining the sensitivity of unfavorable conditions for the dispersion of winter smog pollutants to future climate change, by means of frequencies of stability indices, rather than aiming to project actual concentrations of the air pollutants. Therefore, our presented results provide a perfect complement to validate CTM results, and possibly increasing the confidence of the results. Furthermore, the simplicity of our methodology makes it a powerful tool for decision making in the context of air pollution reduction strategies.

7

GENERAL CONCLUSIONS AND PERSPECTIVES

7.1   Main conclusions and limitations

Regional climate modeling based on dynamical downscaling resolves the scale discrepancy between Global Climate Model (GCM) output and the high resolutions required for impact assessment. Compared to lower resolution models, Regional Climate Models (RCMs) provide added value and more appropriate information by simulating the local aspects and underlying subgrid scale climate processes, such as extreme precipitation, more realistically. The ultimate aim of this thesis is to investigate to what extent the Belgian ALARO-0 Numerical Weather Prediction (NWP) model can be applied for regional climate modeling of (i) (extreme) precipitation and of (ii) the unfavorable meteorological conditions for the dispersion of air pollution in Belgium.

The three key research goals of this thesis, which were formulated in the General introduction (Chapter 1), are:

1. *Describe in detail the Belgian operational ALARO-0 NWP model with its revised physical parameterizations in the context of regional climate modeling.* A detailed study and validation of the application of the ALARO-0 NWP model for regional climate purposes in Belgium, has since its use for regional climate simulations in 2010, never been carried out, making it a very important novelty of this thesis.

2. *Investigate at which temporal and spatial scales the downscaling results add value.* As mentioned previously, the downscaling principle aims that RCMs should not alter the simulated climate on scales that can be skillfully reproduced by the resolutions of the global model, but should rather add value on the finer scales, such as for example mesoscale structures and extremes.

3. *Assess qualitatively the uncertainty of the regional downscaling climate change results.* Multi-model ensemble simulations, allowing to quantify the spread of uncertainty due to model formulation and natural variability, is not feasible at a small institute such as the RMI with limited computing resources. Therefore, the uncertainty is explored in a qualitative manner by comparing our downscaled future climate model results in its proper context, i.e. w.r.t. other RCM climate projections, which ideally use the same scenario of natural and anthropogenic forcing.

These goals were addressed by two main research steps, focusing on extreme precipitation and air pollution dispersion during winter smog episodes in Belgium. In a first step, the Belgian operational ALARO-0 NWP model for climatological time scales is validated, by driving the model with "perfect boundary conditions" coming from global reanalyses. In a next step, the ALARO-0 model is applied for a dynamical downscaling of climate change projection, by driving the model with a GCM scenario. The key findings of both research steps are summarized and discussed below.

7.1.1   Validation of the ALARO-0 model for regional climate modeling in Belgium

The first part of this thesis (including Chapters 2 to 4) focused on the validation of the ALARO-0 model in a climate setup, and more specifically on the validation of extreme precipitation. The ALARO-0 model, which is operationally used at the Royal Meteorological

Institute of Belgium (RMI), finds its origins in NWP. In the research department of the RMI the model is since 2010 also used for regional climate simulations. Hence, **Chapter 2** described in detail the Belgian operational ALARO-0 NWP model as used for the regional climate simulations that were performed for this research. The model is a new version of the ALADIN model, centred around an improved physical parameterization package for precipitation and clouds, called Modular Multiscale Microphysics and Transport (3MT). The basic concepts of 3MT which were given in Chapter 2, yet highlighted the importance of the new physics parameterization scheme for the simulation of convective and extreme precipitation.

Furthermore, the different options in the surface and radiation parameterization schemes of the model were discussed in this chapter. Historically, the ALARO-0 model relies for the land-surface parameterization on the ISBA scheme (Interactions between Soil, Biosphere and Atmosphere). However, during the last decade, the more sophisticated land surface scheme called SURFEX (SURFace EXternalisée) has been developed. For radiation there are two different parameterization schemes available within ALARO-0: the Action de Recherche Petite Echelle Grande Echelle (ARPEGE) Calcul Radiatif avec Nebulosité scheme (ACRANEB) scheme, and the European Centre for Medium-Range Weather Forecasts (ECMWF) Fouquart-Morcrette Radiation (FMR) scheme. The greater part of the ALARO-0 simulations carried out in this thesis have used the "default" settings using the ISBA and ACRANEB scheme, which is also the configuration as used for the current operational NWP applications of the model. However, for the analysis of Chapter 6, simulations were performed with the SURFEX and FMR scheme. Therefore, the sensitivity and relative impact to the use of the more sophisticated surface scheme SURFEX and a different radiation scheme was assessed. This was done through comparison of 30-yr daily precipitation and daily 2-meter minimum temperature from ERA-Interim driven ALARO-0 simulations at 4 km resolution, using the different options in the surface and radiation scheme. The results from this assessment demonstrated for precipitation a minor sensitivity to the surface scheme, and smaller differences in precipitation w.r.t. the observations with the ACRANEB scheme, which has been specifically designed for ALARO-0, than with the FMR scheme. From this, we could conclude that w.r.t. the other setups, the "default" settings using the ACRANEB and ISBA scheme, is an acceptable setup for the simulation of precipitation with the ALARO-0 model. In particular, because, except for Chapter 6, the focus in this thesis is mainly limited to precipitation as climatological variable.

The experimental setup used for the regional climate simulations in this thesis are based on the widely used "nesting technique". This technique, which originates from NWP, consists of a pure downscaling with reinitializations based on the concept of one-way nesting. However, it should be stressed that the nesting technique involves several technical issues. One of these issues is related to the Lateral Boundary Condition (LBC) problem, where possible errors in the large-scale circulation produced by the driving model will be transmitted to the nested model. In order to minimize the effects of this problem, it is recommended to first validate the model for the current climate using analyses of observations, i.e. the so-called "perfect boundary conditions". In **Chapter 3** and **Chapter 4**, LBCs from reanalyses of ERA-40 as well as ERA-Interim have indeed been used to drive the ALARO-0 model for the validation of (extreme) precipitation in Belgium at daily and subdaily temporal scales as well as at multiple spatial resolutions. The main feature of the ALARO-0, i.e. the new 3MT physics parameterization package, and its multiscale characteristic, had never

been validated for precipitation in a climate context.

Hence, to study in detail the multiscale characteristic of 3MT, **Chapter 3** elaborated on the relative importance of resolution versus parameterization formulation on the model skill to simulate realistic extreme daily precipitation. This was achieved by comparing at varying horizontal resolutions 30-yr (1961-1990) daily cumulated summer precipitation from the ALADIN model and the ALARO-0 model with respect to observations. We assessed the model performances through standard statistical errors and density, frequency, and quantile distributions as well as Extreme Value Analysis (EVA), using the peak-over-threshold method and Generalized Pareto Distribution. The 40-km simulations of ALADIN and ALARO-0 showed similar results, both reproducing the observations reasonably well. For the high-resolution simulations, we found that ALARO-0 at both 10 and 4 km is in better agreement with the observations than ALADIN. The ALADIN model consistently produces too high precipitation rates. These findings demonstrated that the new 3MT parameterization, and its multiscale characteristic, is responsible for a correct simulation of extreme summer precipitation at various horizontal resolutions.

For impact studies of extreme precipitation events, decision makers often require current and projected future climate information at the local scale and at higher temporal resolutions than the daily scale. To address this issue, **Chapter 4** validated in detail for multiple spatial resolutions different aspects of the characteristics of (extreme) precipitation at the subdaily timescale, ranging between 1-hour aggregation levels up to the daily timescale (24-hour aggregation level). Furthermore, the central question in this assessment was whether the multiscale performance of ALARO-0 in the simulation of daily summer precipitation, as was found in the previous chapter (Chapter 3), is also retained in the simulation of subdaily precipitation. First, our results for the diurnal cycle showed for both high-resolution simulations at 4 and 10 km resolution an improvement in the onset and peak of convective activity w.r.t. the observations. The low resolution 40-km run on the other hand, is not able to reproduce the observed diurnal cycle both in magnitude and phase. Furthermore, compared to the 40-km simulation, we found for the high-resolution ALARO-0 runs a consistent improvement in the simulation of high hourly precipitation amounts. Finally, two scaling properties related to extreme rainfall; i.e. the linear behavior of the Generalized Extreme Value parameters and the Clausius-Clapeyron (CC) relation were examined. Again, the high-resolution simulations showed for all durations a power law dependency that approaches very closely the observed power law. In contrast, the low-resolution 40-km run did not reproduce the observed scaling properties for the lowest durations (1 to 3 hour aggregation times). Only for the 24-h durations both the high-resolution as well as the low, 40-km resolution runs displayed an Intensity-Duration-Frequency relationship based upon this power law which is in agreement with the observations. The CC relation was also relatively well reproduced by the model at a spatial resolution of 4 km. In particular for the most extreme events (i.e. highest percentiles) and for a rather large temperature range, the model is able to reproduce the (super-)CC relation.

These results suggest that the multiscale performance of ALARO-0 in the simulation of daily summer precipitation, as was found in **Chapter 3**, does not hold for the simulation of subdaily precipitation. Furthermore, our results demonstrated that the highest-resolution simulations of ALARO-0 at 4 km benefit from added value in the description of several char-

acteristics of subdaily precipitation, such as the diurnal cycle, heavy precipitation amounts, and important scaling properties. This finding is in line with previous studies that explored w.r.t. observations the added value of low-resolution RCMs where deep convection is parameterized versus high-resolution CPMs where the convection parameterizations are (partly) switched off. The multiscale character of the physics package 3MT for clouds and deep convection allowed us to use consistently the same model physics at a range of different spatial resolutions, which is an important strength in our assessment w.r.t. other studies where one is often obliged to use a different model physics or treatment of deep convection once the horizontal resolution is changed.

Overall the results of the validation of extreme precipitation at the daily and subdaily timescale, as extensively studied in **Chapter 3** and **Chapter 4** demonstrated that the ALARO-0 model is able to consistently capture the relevant precipitation characteristics at a wide range of atmospheric and corresponding temporal scales, varying from the micro- to the mesoscales. It is important to mention that this finding is representative to only one climate variable (i.e. precipitation), one season (i.e. the summer season) and one location (i.e. Belgium or Brussels (Uccle)). It is important to mention that this finding is representative to only one climate variable (i.e. precipitation), one season (i.e. the summer season) and one location (i.e. Belgium or Brussels (Uccle)). It is clear that these limitations in the validation do not assure a similar performance of the model for other variables, seasons or locations. If one desires to assess the performance of the ALARO-0 model for another study area, it is preferable to enlarge or relocate the respective model domain. In a recent validation study of Giot et al. (2016), using the same configuration of the ALARO-0 model but a slightly different setup as used here, the model has indeed been tested for Europe within the framework of the Coordinated Regional Climate Downscaling Experiment (EURO-CORDEX). The performance of the model is quantified through several metrics which are compared to metrics from an ensemble of 17 other EURO-CORDEX experiments. The results demonstrated that the model is capable of correctly representing the precipitation and temperature climate variables for the European region in an acceptable way as most of the ALARO-0 scores lie within the existing ensemble (Giot et al., 2016).

7.1.2   Application of the ALARO-0 model for present and future climate impacts on extreme precipitation and on the unfavorable meteorological conditions for the dispersion of air pollution

The highly promising and encouraging findings of the validation of extreme precipitation increased our confidence to apply the ALARO-0 model to compute Intergovernmental Panel on Climate Change (IPCC) scenarios for climate change studies. In the context of climate change assessment, potential future changes in extreme precipitation are of great importance. In particular because such precipitation extremes are related to e.g. floods and landslides, which have a great impact on many aspects of human society: health, natural and urban environments, buildings and infrastructure, and economy. Hence, **Chapter 5** studied the future response of extreme precipitation over Belgium to increased greenhouse gas (GHG) concentrations under the A1B IPCC scenario. For this, the GCM CNRM-CM3 has been dynamically downscaled for a 30-yr control period (1961-90) and future scenario period (2071-2100), using the ALARO-0 model at 4 km spatial resolution. In a first step, it

was verified whether or not the observed climatology for the control period is correctly represented by the model, which is important if one couples a RCM to a GCM. The results of this evaluation revealed significant biases in the simulation of daily precipitation and temperature, as well as hourly mean precipitation. These biases were found to be mainly related to model errors present in the driving GCM CNRM-CM3. Furthermore, it was found that the downscaling of the GCM data towards higher spatial resolutions of 4 km results in more extreme summer precipitation, and significantly reduces the bias w.r.t. the observations in the control simulation. It is worth mentioning that the formulation of clear reasons for the model biases was out of the scope of this thesis. This can be done by assessing the model sensitivity to changes in the driving GCM model formulation (e.g. physical parameterizations, model resolution, atmosphere-ocean coupling, ...). Nevertheless, in order to address these questions for future research, one has to know the biases.

In a next step, we analyzed the future changes in which we limited ourselves to a sensitivity of the model to the different sets of control and scenario boundary conditions, by quantifying the differences between the future scenario and control simulation without the application of any bias correction. The results from the analysis on the future changes in mean temperature showed a significant warming in 2-meter mean temperature by the end of the 21st century throughout the summer and winter season. For mean precipitation our results demonstrated for summer an overall significant decrease for the whole country, and a significant increase during winter for the coastal and central region of the country. The future changes in extreme precipitation showed for summer little significant but negative changes, which are reflected in a negative response of 5-yr return level values of hourly and daily precipitation. For winter, the 5-yr return level values of daily precipitation showed mostly positive and significant changes between the scenario and control simulation.

However, we did not attempt to *quantify* the uncertainties of these future changes, we compared the changes w.r.t. the model biases as well as w.r.t. other GCM/RCM results. This allowed us to provide a qualitative notion of the uncertainty and confidence of our results. For example, the highest percentiles of hourly winter precipitation in Uccle showed a positive future change which is more than 50% larger than the model bias, which is also reflected in the positive changes in extreme daily winter precipitation as given by the 5-yr return level estimates for Belgium. Furthermore, our results for winter extreme precipitation are in line with other RCM and high-resolution model results for our region, which also show a projected increase in extreme winter precipitation amounts. In this way we could expect, with some level of confidence, an intensification of *extreme* hourly winter precipitation in Belgium. However, for summer the negative changes in extreme and hourly precipitation are much more uncertain, since (i) most of the changes are non-significant and smaller than the respective biases, and (ii) other modeling studies for our regions project an increase in extreme summer precipitation, which is in disagreement with our results. In general, the disagreements and highly varying patterns of projected changes in extreme summer precipitation can be explained by the transition zone in which Belgium is located. Furthermore, the strong dependency of the parameterizations, and in particular the deep convection parameterizations, are also a key source for the uncertainty in future climate projections of extreme summer precipitation.

The last chapter, **Chapter 6**, assessed the impact of climate change on another major

area where climate change may cause adverse effects, i.e. air quality. The effect of climate change on winter smog episodes was studied through the analysis of two different stability indices, i.e. the transport index and the Pasquill classes, which are based on meteorological conditions determining the dispersion of air pollution. This methodology, which is commonly used for alerts of winter smog peaks in a context of operational weather impact, was applied and validated here for the first time with the high-resolution climatological data from the ALARO-0 model. We have calculated both stability indices for a 29-yr and 9-yr winter period using present (1981/82-2009/10 and 1990/91-1998/99) and future (2046/47-2054/55) climate data that has been obtained from a dynamical downscaling of GCM data from the ERA-Interim reanalysis as well as from the coupled CNRM-CM3 model using the ALARO-0 model at 4 km spatial resolution.

In a first step, we assessed whether the ERA-Interim driven ALARO-0 model is able to reproduce observed cases of winter smog alerts by means of the transport length values as well as observed frequency statistics of the Pasquill stability indices. The results showed that the model is able to reproduce almost all the observed smog peaks, and also the simulated frequency distribution of Pasquill indices are in close agreement to the observed one. Comparison of frequency statistics of the transport- and Pasquill indices obtained from the control simulation with the observed frequencies indeed revealed significant biases, and in particular for the transport index. To account for these model imperfections, a linear Q-Q bias correction was applied directly on the transport length values. After correction, the present-day frequencies were significantly improved, with remaining deviations from the observations that fall within an acceptable range of 10%.

In order to quantify the uncertainty on the future changes, the same linear correction method as used to correct present-day transport length values, has been applied to correct the future transport lengths. Our results demonstrated that we can expect an increase of the unfavorable conditions for the dispersion of air pollutants up to 60 - 70% in Brussels by the middle of the 21st century. It is important to keep in mind that the assumption of constant bias under present and future climate conditions may not be valid, and that a more advanced bias correction method could be more appropriate. Furthermore, it should be taken into account that the future climate change impact on winter smog episodes as found here, is based upon one single emission scenario from the downscaling results of one RCM. It is important to highlight the relevance and innovative strength of the methodology of stability indices as proposed in this chapter. In practice, it is generally hard to evaluate the skill of Chemical Transport Model (CTM) results because of the limited availability of observational data for the evaluation. To our knowledge, no previous studies have focused on examining the sensitivity of unfavorable conditions for the dispersion of winter smog pollutants to future climate change, by means of frequencies of stability indices, rather than aiming to project actual concentrations of the air pollutants. Therefore, our presented results provide a perfect complement to validate CTM results, and possibly increasing the confidence of the results. Furthermore, the simplicity of our methodology makes it a powerful tool for decision making in the context of air pollution reduction strategies.

The results from Chapter 6 suggested for the winter a possible change towards more anticyclonic and stable situations, whereas the results from Chapter 5 indicated for the winter season an increase of future precipitation extremes which is often associated with cyclonic conditions. This suggests that, according to our model results, the future climate change

impact for the winter season on the extreme cases should rather be attributed to the fact that the extremes become more extreme, and not so much to a change in the recurrence frequencies of respective anticyclonic or cyclonic weather types.

Finally, it is important to mention, that the results on the future changes in extreme precipitation and the meteorological conditions which are unfavorable for the dispersion of air pollution associated with winter smog peaks in Brussels (Uccle), are *one* possible outcome of future changes at the end or middle of the 21st century. Our assessment on the climate impact of winter smog episodes has been made under the consideration that the actual concentration peaks of the relevant winter smog pollutants are superimposed on their background concentrations, and hence of a much larger order of magnitude than the respectively background concentrations. However, in this consideration, future background concentrations and emissions of GHGs and other pollutants, remain an important source of uncertainty, due to the unknown changes in population vulnerability and human activity patterns. In this respect, a multi-model ensemble approach is a promising strategy to narrow and quantify the uncertainty on the projected model results.

7.2 Perspectives

7.2.1 Future developments in the ALARO model for climate simulations

Extensive developments of the ALARO model have recently led to the release of a new model version, i.e. the ALARO-1 model. The new features in this new model version include amongst others several updates in the different physical parameterization schemes for e.g. radiation, turbulence, and deep convection. As mentioned previously, the results on the validation of the subdaily precipitation of the ALARO-0 model, showed a significant improvement in the representation of the diurnal cycle for the high-resolution runs w.r.t. the low 40-km simulation. However, there is still a time shift towards a too early onset of the convective build-up notable. Recent NWP results demonstrated that changes in the 3MT physics package which are included in ALARO-1, significantly reduced (or even canceled) this time shift. These most recent changes are not implemented in the ALARO-0 version that is used for the simulations in this thesis. Hence, in the perspective of climate applications with the new ALARO-1 model version, the presence of this time shift in the diurnal cycle could be examined.

In NWP there is a general tendency to increase the models' spatial resolution to 1 km or even less. The developments of the ALARO-1 model also attempt to allow NWP applications to such high spatial resolutions. Hence, with the continuing advancements in high-performance computing, we can expect parallel evolutions in the application of the ALARO-1 model for regional climate simulations. It is self-evident that such a refinement of the horizontal grid spacing requires an adaptation and tuning of the physical parameterizations. At the same time these improvements in the physical parameterizations are needed to further investigate the dominant physical processes that are effectively responsible for changes of subdaily extreme rainfall in a future climate.

Another important limitation in the validation and potential future changes of extreme precipitation at the subdaily time scale, is the availability of sufficiently long time series

of observed precipitation at high-temporal resolutions. However, a dense network of such point observations from gauge measurements are generally difficult to find. Gridded hourly radar data could provide a good alternative to gauge measurements, as they have excellent resolutions both in time and space. At the RMI rainfall estimates from a 10-yr (2005-2015) dataset of volumetric weather radar measurements from a radar located in the southeast of Belgium (Wideumont), have recently been processed. For future climate research of the subdaily precipitation characteristics this gridded rainfall product can be used as a reference.

7.2.2   ALARO-0 in the context of the EURO-CORDEX project

As mentioned previously, our experimental setup for the regional climate simulations in this thesis, is based on the nesting technique, which involves several technical issues. One of these issues is related to the problem that the internal climate of the RCM can start to diverge from the large-scale atmospheric circulation given by the driving global model. In our experimental setup we deal with these deviations by means of daily reinitializations, while taking into account a spinup period which allows that the physics can adjust. Although, another and more commonly used approach within regional climate modeling, are continuous and uninterrupted model runs over long periods, so that the RCM finds its own climate equilibrium. Hence, it can be argued that the use of this setup for climate integrations, does not allow for the nested model to find its balance. This may indeed cause some spurious biases. However, this issue has not been explored in the thesis, a recent validation study of Giot et al. (2016) where the ALARO-0 model is validated within the CORDEX framework, used a continuous model integration setup, and revealed similar and positive results in the models ability to reproduce the precipitation climatology over Europe.

Furthermore, for the simulations where SURFEX was employed, the soil variables evolved freely after initialization and were never corrected or nudged in the course of the simulation. This is motivated by the fact that daily restarts would limit the equilibration of the surface physics (soil moisture and temperature), which is particularly desirable in long-term regional climate modeling. In the framework of an ongoing research project at the RMI, the effects of the different nesting approaches (i.e. daily restarts versus continuous integrations) are investigated, and this will be continued in further research.

Finally, the ongoing project CORDEX.be (COmbining the Regional Downscaling EXpertise in BElgium: CORDEX and beyond), aims to produce a set of comparable simulations by the Belgian regional climate modeling groups within the EURO-CORDEX framework. This will allow to extend our research on the future climate impact of extreme precipitation and the meteorological conditions for the dispersion of air pollutants to other European regions, using IPCC's most recent Representative Concentration Pathways (RCP). Furthermore, several statistical downscaling techniques will be applied within the project, to infer the climate uncertainties with the CORDEX.be micro-ensemble, to be properly situated w.r.t. the other runs in the EURO-CORDEX archive.

References

ALADIN international team, 1997: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. *WMO Bull.*, **46**, 317–324.

Allen, M. R. and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232, doi:10.1038/nature01092.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *J. Climate*, **17**, 2493–2525, doi:10.1175/1520-0442(2004)017⟨2493:RATCPP⟩2.0.CO;2.

Arakawa, A. and J.-H. Jung, 2011: Multiscale modeling of the moist-convective atmosphere – A review. *Atmos. Res.*, **102**, 263–285, doi:10.1016/j.atmosres.2011.08.009.

Arakawa, A., J.-H. Jung, and C.-M. Wu, 2011: Toward unification of the multi-scale modeling of the atmosphere. *Atmos. Chem. Phys.*, **11**, 3731–3742, doi:10.5194/acp-11-3731-2011.

Attema, J. J., J. M. Loriaux, and G. Lenderink, 2014: Extreme precipitation response to climate perturbations in an atmospheric mesoscale model. *Environ. Res. Lett.*, **9**, 12, doi:10.1088/1748-9326/9/1/014003.

Baguis, P., E. Roulin, P. Willems, and V. Ntegeka, 2010: Climate change scenarios for precipitation and potential evapotranspiration over central Belgium. *Theor. Appl. Climatol.*, **99**, 273–286, doi:10.1007/s00704-009-0146-5.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res.*, **119**, 7889–7907, doi:10.1002/2014JD021478.

Ban, N., J. Schmidli, and C. Schär, 2015: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophys. Res. Lett.*, **42**, 1165–1172, doi:10.1002/2014GL062588.

Bechtold, P., J. P. Chaboureau, A. Beljaars, A. K. Betts, M. Köhler, M. Miller, and J. L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model. *Quart. J. Roy. Meteor. Soc.*, **130**, 3119–3137, doi:10.1256/qj.03.103.

Belamari, S. and A. Pirani, 2007: Validation of the optimal heat and momentum fluxes using the ORCA2-LIM global ocean-ice model. *Marine environment and security for the European area. Integrated Project (MERSEA IP)*, **Deliverable D4.1.3**, 88.

Beniston, M., D. B. Stephenson, O. B. Christensen, C. A. T. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylhä, B. Koffi, J. Palutikof, R. Schöll, T. Semmler, and K. Woth, 2007: Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change*, **81**, 71–95, doi:10.1007/s10584-006-9226-z.

Berg, P., J. O. Haerter, P. Thejll, C. Piani, S. Hagemann, and J. H. Christensen, 2009: Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature. *J. Geophys. Res.*, **114**, 9, doi:10.1029/2009JD012008.

Berg, P., C. Moseley, and J. O. Haerter, 2013a: Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.*, **6**, 181–185, doi:10.1038/ngeo1731.

Berg, P., S. Wagner, H. Kunstmann, and G. Schädler, 2013b: High resolution regional climate model simulations for Germany: part I - validation. *Climate Dyn.*, **40**, 401–414, doi:10.1007/s00382-012-1508-8.

Best, M., A. Beljaars, J. Polcher, and P. Viterbo, 2004: A proposed structure for coupling tiled surfaces with the planetary boundary layer. *J. Hydrometeor.*, **5**, 1271–1278, doi:10.1175/JHM-382.1.

Betts, A. K., M. Köhler, and Y. Zhang, 2009: Comparison of river basin hydrometeorology in ERA-Interim and ERA-40 reanalyses with observations. *J. Geophys. Res.*, **114**, 12, doi:10.1029/2008JD010761.

Boberg, F., P. Berg, P. Thejll, W. J. Gutowski, and J. H. Christensen, 2010: Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models. *Climate Dyn.*, **35**, 1509–1520, doi:10.1007/s00382-009-0683-8.

Bougadis, J. and K. Adamowski, 2006: Scaling model of a rainfall intensity-duration-frequency relationship. *Hydrol. Process.*, **20**, 3747–3757, doi:10.1002/hyp.6386.

Bougeault, P., 1985: A simple parameterization of the large-scale effects of cumulus convection. *Mon. Wea. Rev.*, **113**, 2108–2121.

Brockhaus, P., D. Lüthi, and C. Schär, 2008: Aspects of the diurnal cycle in a regional climate model. *Meteor. Z.*, **17**, 433–443, doi:10.1127/0941-2948/2008/0316.

Brožková, R., 2015: Deep convection diurnal cycle in ALARO-1. *ALADIN-HIRLAM Newsletter no. 4*, 433–443.

Bubnová, R., G. Hello, P. Bénard, and J.-F. Geleyn, 1995: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Mon. Wea. Rev.*, **123**, 515–535.

Bultynck, H. and L. Malet, 1972: Evaluation of atmospheric dilution factors for effluents diffused from an elevated continuous point source. *Tellus*, **24**, 455–472, doi:10.1111/j.2153-3490.1972.tb01572.x.

Buonomo, E., R. Jones, I. C. Huntingford, and J. Hannaford, 2007: On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quart. J. Roy. Meteor. Soc.*, **133**, 65–81, doi:10.1002/qj.13.

Burlando, P. and R. Rosso, 1996: Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation. *J. Hydrol.*, **187**, 45–64, doi:10.1016/S0022-1694(96)03086-7.

Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala, 2009: Evaluation of a WRF dynamical downscaling simulation over California. *Climatic Change*, **95**, 499–521, doi:10.1007/s10584-009-9583-5.

Castro, C. L., R. A. S. Pielke, and G. Leoncini, 2005: Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *J. Geophys. Res.*, **110**, 21, doi:10.1029/2004JD004721.

Catry, B., J.-F. Geleyn, M. Tudor, P. Bénard, and A. Trojáková, 2007: Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus*, **59A**, 71–79, doi:10.1111/j.1600-0870.2006.00212.x.

Champeaux, J., V. Masson, and R. Chauvin, 2005: Ecoclimap: A global database of land surface parameters at 1 km resolution. *Meteorol. Appl.*, **12**, 29–32, doi:10.1017/S1350482705001519.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, C. A. T. Ferro, and D. B. Stephenson, 2013: Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Clim. Dynam.*, **41**, 1475–1495, doi:10.1007/s00382-012-1568-9.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, N. M. Roberts, and C. A. T. Ferro, 2014: The value of high-resolution Met Office regional climate models in the simulation of multihourly precipitation extremes. *J. Climate*, **27**, 6155–6174, doi:10.1175/JCLI-D-13-00723.1.

Charnock, H., 1955: Wind stress over a water surface. *Quart. J. Roy. Meteor. Soc.*, **81**, 639–640, doi:10.1002/qj.49708135027.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer, Berlin Heidelberg, 210 pp.

Csima, G. and A. Horányi, 2008: Validation of ALADIN-Climate regional climate model at the Hungarian Meteorological Service. *Quarterly Journal of the Hungarian Meteorological Service*, **112**, 155–177.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate*, **19**, 4605–4630, doi:10.1175/JCLI3884.1.

Dai, A., F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.: Atmospheres*, **104**, 6377–6402, doi:10.1029/98JD02720.

Davies, H. C., 1976: A lateral boundary formulation for multilevel prediction models. *Quart. J. Roy. Meteor. Soc.*, **102**, 405–418.

de Elía, R., R. Laprise, and D. Bertrand, 2002: Forecasting skill limits of nested, limited-area models: A perfect-model approach. *Mon. Wea. Rev.*, **130**, 2006–2023, doi:10.1175/1520-0493(2002)130⟨2006:FSLONL⟩2.0.CO;2.

De Troch, R., R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia, 2013: Multi-scale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. *J. Climate*, **26**, 8895–8915, doi:10.1175/JCLI-D-12-00844.1.

De Troch, R., H. Van de Vyver, O. Giot, R. Hamdi, and P. Termonia, 2014a: Assessment of subdaily precipitation from the ALARO-0 model at different spatial resolutions. Presented at the European Geosciences Union General Assembly 2014, Vienna, Austria.

De Troch, R., O. Giot, R. Hamdi, S. Saeed, H. Tabari, M. T. Taye, P. Termonia, N. van Lipzig, and P. Willems, 2014b: Overview of a few regional climate models and climate scenarios for Belgium. *Scientific and technical publication, nr. 65*, D. Gellens, Ed., Royal Meteorological Institute of Belgium.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kllberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Delcloo, A., C. Andersson, B. Forsberg, T. Nawrot, and M. Valari, 2014: ACCEPTED: An Assessment of Changing Conditions, Environmental Policies, Time - Activities, Exposure and Disease. *Air Pollution Modeling and its Application XXIII*, D. Steyn and R. Mathur, Eds., Springer International Publishing, 55–59, doi:10.1007/978-3-319-04379-1_9.

Denis, B., R. Laprise, and D. Caya, 2003: Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. *Clim. Dynam.*, **20**, 107–126, doi:10.1007/s00382-002-0264-6.

Denis, B., R. Laprise, D. Caya, and J. Côté, 2002: Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Clim. Dynam.*, **18**, 627–646, doi:10.1007/s00382-001-0201-0.

Déqué, M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global Planet. Change*, **57**, 16–26, doi:10.1016/j.gloplacha.2006.11.030.

Déqué, M. and S. Somot, 2008: Analysis of heavy precipitation for France using high resolution ALADIN RCM simulations. *Quarterly Journal of the Hungarian Meteorological Service*, **112**, 179–190.

Déqué, M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. de Castro, and B. van den Hurk, 2007: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, **81**, 53–70, doi:10.1007/s10584-006-9228-x.

Domínguez, M., R. Romera, E. Sánchez, L. Fita, J. Fernández, P. Jiménez-Guerrero, J. P. Montávez, W. D. Cabos, G. Liguori, and M. A. Gaertner, 2011: Present climate precipitation and temperature extremes over Spain from a set of high resolution RCMs. *Climate Research*, submitted, available online at http://www.meteo.unican.es/en/node/73026.

Duerinckx, A., R. Hamdi, J.-F. Mahfouf, and P. Termonia, 2015: Study of the Jacobian of an extended Kalman filter for soil analysis in SURFEXv5. *Geosci. Model Dev.*, **8**, 845–863, doi:10.5194/gmd-8-845-2015.

Dulière, V., Y. Zhang, and E. P. J. Salathé, 2011: Extreme precipitation and temperature over the U.S. Pacific Northwest: A comparison between observations, reanalysis data, and regional models. *J. Climate*, **24**, 1950–1964, doi:10.1175/2010JCLI3224.1.

Durman, C. F., J. M. Gregory, D. C. Hassell, R. G. Jones, and J. M. Murphy, 2001: A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates. *Quart. J. Roy. Meteor. Soc.*, **127**, 1005–1015, doi:10.1002/qj.49712757316.

Ebi, K. L. and G. McGregor, 2008: Climate change, tropospheric ozone and particulate matter, and health impacts. *Environ. Health Perspect.*, **116**, 1449–1455, doi:10.1289/ehp. 11463.

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert, 2012: HESS Opinions "Should we apply bias correction to global and regional climate model data?". *Hydrol. Earth Syst. Sci.*, **16**, 3391–3404, doi:10.5194/hess-16-3391-2012.

European Centre for Medium-Range Weather Forecasts, 2010: IFS Documentation - Cy36r1. Operational implementation 26 january 2010. Part III: Dynamics and numerical procedures. IFS Documentation - Cy36r1, 29 pp.

European Community, 1985: Council Directive 85/580/eec of 20 december 1985, amendment of Council Directive 85/203/eec on air quality standards for nitrogen dioxide. Official Journal of the European Communities L372, 36 pp. [Available online at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31985L0580&from=EN].

European Community, 1999: Council Directive 1999/30/ec of 22 april 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air. Official Journal of the European Communities L163, 41-60 pp. [Available online at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31999L0030&from=EN].

European Community, 2008: Directive 2008/50/ec of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Communities L152, 1-44 pp. [Available online at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0050&from=EN].

European Environment Agency, 1998: Assessment and management of urban air quality in europe. EEA Monograph 05/1998, 150 pp. ISBN: 92-9167-103-7.

European Environment Agency, 2014: European environment agency: Air quality in Europe - 2014 report. EEA report 05/2014, 80 pp. Copenhagen (Denmark), doi:10.2800/22847.

Feldmann, H., G. Schädler, H.-J. Panitz, and C. Kottmeier, 2012: Near future changes of extreme precipitation over complex terrain in Central Europe derived from high resolution RCM ensemble simulations. *Int. J. Climatol.*, **134**, 2180–2190, doi:10.1175/MWR3183.1.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. (eds.), Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 741–866.

Fosser, G., S. Khodayar, and P. Berg, 2015: Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Clim. Dynam.*, **44**, 45–60, doi:10.1007/s00382-014-2242-1.

Fouquart, Y. and B. Bonnel, 1980: Computations of solar heating of the earths atmosphere: A new parameterization. *Quart. J. Roy. Meteor. Soc.*, **53**, 35–62.

Fowler, H. J. and M. Ekström, 2009: Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *Int. J. Climatol.*, **29**, 385–416, doi:10. 1002/joc.1827.

Fowler, H. J., M. Ekström, S. Blenkinsop, and A. P. Smith, 2007: Estimating change in extreme European precipitation using a multimodel ensemble. *J. Geophys. Res.*, **112**, 20, doi:10.1029/2007JD008619.

Fowler, H. J., M. Ekström, C. Kilsby, and P. Jones, 2005: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *J. Hydrol.*, **300**, 212–233, doi:10.1016/j.jhydrol.2004.06.017.

Frei, C., R. Schöll, S. Fukutome, J. Schmidli, and P. L. Vidale, 2006: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *J. Geophys. Res.*, **111**, 22, doi:10.1029/2005JD005965.

Geleyn, J.-F., P. Bénard, and R. Fournier, 2005: A general-purpose extension of the Malk-mus band-model average equivalent width to the case of the Voigt line profile. *Quart. J. Roy. Meteor. Soc.*, **131**, 2757–2768, doi:10.1256/qj.04.107.

Geleyn, J.-F., B. Catry, Y. Bouteloup, and R. Brožková, 2008: A statistical approach for sedimentation inside a microphysical precipitation scheme. *Tellus*, **60A**, 649–662, doi: 10.1111/j.1600-0870.2008.00323.x.

Gerard, L., 2007: An integrated package for subgrid convection, clouds and precipitation compatible with the meso-gamma scales. *Quart. J. Roy. Meteor. Soc.*, **133**, 711–730, doi: 10.1002/qj.58.

Gerard, L. and J.-F. Geleyn, 2005: Evolution of a subgrid deep convection parameterization in a limited area model with increasing resolution. *Quart. J. Roy. Meteor. Soc.*, **131**, 2293–2312, doi:10.1256/qj.04.72.

Gerard, L., J.-M. Piriou, R. Brožková, J.-F. Geleyn, and D. Banciu, 2009: Cloud and precipitation parameterization in a meso-gamma-scale operational weather prediction model. *Mon. Wea. Rev.*, **137**, 3960–3977, doi:10.1175/2009MWR2750.1.

Giorgi, F., X. Bi, and J. S. Pal, 2004: Mean, interannual variability and trends in a regional climate change experiment over Europe. I. Present-day climate (1961-1990). *Clim. Dynam.*, **22**, 733–756, doi:10.1007/s00382-004-0409-x.

Giorgi, F. and M. R. Marinucci, 1996: An investigation of the sensitivity of simulated precipitation to model resolution and its implications for climate studies. *Mon. Wea. Rev.*, **124**, 148–166, doi:10.1175/1520-0493(1996)124⟨0148:AIOTSO⟩2.0.CO;2.

Giorgi, F. and L. O. Mearns, 1999: Introduction to special section: Regional Climate Modeling Revisited. *J. Geophys. Res.*, **104**, 6335–6352.

Giorgi, F. and F. Meleux, 2007: Modelling the regional effects of climate change on air quality. *Compt. Rendus Geosci.*, **339**, 721733, doi:10.1016/j.crte.2007.08.006.

Giot, O., R. De Troch, R. Hamdi, A. Deckmyn, and P. Termonia, 2014: Impact of using the surface scheme SURFEX within ALARO-0 for the ERA-INTERIM high-resolution dynamical downscaling over belgium. *3rd Lund Regional-scale Climate Modelling Workshop, 21st Century Challenges in Regional Climate Modelling, Lund, Sweden*, International Baltic Earth Secretariat Publications, 101–102.

Giot, O., P. Termonia, D. Degrauwe, R. De Troch, S. Caluwaerts, G. Smet, J. Berckmans, A. Deckmyn, L. De Cruz, P. De Meutter, A. Duerinkcx, L. Gerard, R. Hamdi, J. Van den Bergh, M. Van Ginderachter, and B. Van Schaeybroeck, 2016: Validation of the ALARO-0 model within the EURO-CORDEX framework. *Geosci. Model Dev.*, **9**, 1143–1152, doi:10.5194/gmd-9-1143-2016.

Goudenhoofdt, E. and L. Delobbe, 2015: Verification of rainfall estimates from 10-year volumetric weather radar measurements. *J. Hydrometeor.*, in preparation.

Guichard, F., J. C. Petch, J.-L. Redelsperger, P. Bechtold, J.-P. Chaboureau, S. Cheinet, W. Grabowski, H. Grenier, C. G. Jones, M. Köhler, J.-M. Piriou, R. Tailleux, and M. Tomasini, 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quart. J. Roy. Meteor. Soc.*, **130**, 3139–3172, doi:10.1256/qj.03.145.

Haerter, J. O. and P. Berg, 2009: Unexpected rise in extreme precipitation caused by a shift in rain type? *Nat. Geosci.*, **2**, 372–373, doi:10.1038/ngeo523.

Hamdi, R., O. Giot, R. De Troch, A. Deckmyn, and P. Termonia, 2015: Future climate of Brussels and Paris for the 2050s under the A1B scenario. *Urban Climate*, **12**, 160–182, doi:10.1016/j.uclim.2015.03.003.

Hamdi, R., H. Van de Vyver, R. De Troch, and P. Termonia, 2014: Assessment of three dynamical urban climate downscaling methods: Brussels's future urban heat island under an A1B emission scenario. *Int. J. Climatol.*, **34**, 978–999, doi:10.1002/joc.3734.

Hamdi, R., H. Van de Vyver, and P. Termonia, 2012: New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: application for summer extreme temperature over Belgium. *Int. J. Climatol.*, **32**, 2051–2065, doi:10.1002/joc.2409.

Hanel, M. and T. A. Buishand, 2010: On the value of hourly precipitation extremes in regional climate model simulations. *J. Hydrol.*, **393**, 265–273, doi:10.1016/j.jhydrol.2010.08.024.

Hardwick Jones, R., S. Westra, and A. Sharma, 2010: Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geophys. Res. Lett.*, **37**, 5, doi:10.1029/2010GL045081.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New, 2008: An European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *J. Geophys. Res.*, **113**, 12, doi:10.1029/2008JD010201.

Hedegaard, G. B., J. H. Christensen, and J. Brandt, 2013: The relative importance of impacts from climate change vs. emissions change on air pollution levels in the 21st century. *Atmos. Chem. Phys.*, **13**, 3569–3585, doi:10.5194/acp-13-3569-2013.

Heikkilä, U., A. Sandvik, and A. Sorterberg, 2011: Dynamical downscaling of ERA-40 in complex terrain using WRF regional climate model. *Clim. Dynam.*, **37**, 1551–1564, doi:10.1007/s00382-010-0928-6.

Herrera, S., L. Fita, J. Fernández, and J. M. Gutiérrez, 2010: Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. *J. Geophys. Res.*, **115**, 13, doi:10.1029/2010JD013936.

Hofstra, N., M. New, and C. McSweeney, 2010: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Clim. Dynam.*, **35**, 841–858, doi:10.1007/s00382-009-0698-1.

Hohenegger, C., P. Brockhaus, and C. Schär, 2008: Towards climate simulations at cloud-resolving scales. *Meteor. Z.*, **17**, 383–394, doi:10.1127/0941-2948/2008/0303.

Jacob, D., L. Bärring, O. B. Christensen, J. H. Christensen, M. de Castro, M. Déqué, F. Giorgi, S. Hagemann, M. Hirschi, R. Jones, E. Kjellström, G. Lenderink, B. Rockel, E. Sánchez, C. Schär, S. I. Seneviratne, S. Somot, A. van Ulden, and B. van den Hurk, 2007: An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change*, **81**, 31–52, doi:10.1007/s10584-006-9213-4.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, E. Georgopoulou, A. Gobiet, L. Menut, G. Nikulin, A. Haensler, N. Hempelmann, C. Jones, K. Keuler, S. Kovats, N. Kröner, S. Kotlarski, A. Kriegsmann, E. Martin, E. van Meijgaard, C. Moseley, S. Pfeifer, S. Preuschmann, C. Radermacher, K. Radtke, D. Rechid, M. Rounsevell, P. Samuelsson, S. Somot, J.-F. Soussana, C. Teichmann, R. Valentini, R. Vautard, B. Weber, and P. Yiou, 2014: EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg. Environ. Change*, **14**, 563–578, doi:10.1007/s10113-013-0499-2.

Jacob, D. J. and D. A. Winner, 2009: Effect of climate change on air quality. *Atmos. Env.*, **43**, 51–63, doi:10.1016/j.atmosenv.2008.09.051.

Jones, P. D. and P. A. Reid, 2001: Assessing future changes in extreme precipitation over Britain using regional climate model integrations. *Int. J. Climatol.*, **21**, 1337–1356, doi:10.1002/joc.677.

Jones, R. G., J. M. Murphy, and M. Noguer, 1995: Simulation of climate change over europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. *Quart. J. Roy. Meteor. Soc.*, **121**, 1413–1449, doi:10.1002/qj.49712152610.

Jones, R. G., M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins, and J. F. B. Mitchell, 2004: *Generating high resolution climate change scenarios using PRECIS*. Exeter, UK, Met Office Hadley Centre, ISBN 0 86180 372 8.

Juda-Rezler, K., M. Reizer, P. Huszar, B. C. Krüger, P. Zanis, D. Syrakov, E. Katragkou, W. Trapp, D. Melas, H. Chervenkov, I. Tegoulias, and T. Halenka, 2012: Modelling the effects of climate change on air quality over Central and Eastern Europe: concept, evaluation and projections. *Climate Res.*, **53**, 179–203, doi:10.3354/cr01072.

Katragkou, E., P. Zanis, I. Kioutsioukis, I. Tegoulias, D. Melas, B. C. Krüger, and E. Coppola, 2011: Future climate change impacts on summer surface ozone from regional climate-air quality simulations over Europe. *J. Geophys. Res.*, **116**, 14, doi:10.1029/2011JD015899.

Kelly, J., P. A. Makar, and D. A. Plummer, 2012: Projections of mid-century summer air-quality for north america: effects of changes in climate and precursor emissions. *Atmos. Chem. Phys.*, **12**, 5367–5390, doi:10.5194/acp-12-5367-2012.

Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior, 2014: Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Clim. Change*, **4**, 570–576, doi:10.1038/nclimate2258.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of rainfall in a very high-resolution regional climate model. *J. Climate*, **25**, 5791–5806, doi:10.1175/JCLI-D-11-00562.1.

Kendon, E. J., D. P. Rowell, and R. G. Jones, 2010: Mechanisms and reliability of future projected changes in daily precipitation. *Climate Dyn.*, **35**, 489–509, doi:10.1007/s00382-009-0639-z.

Kharin, V. V. and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphereocean GCM. *J. Climate*, **13**, 3760–3788, doi:10.1175/1520-0442(2000)013⟨3760:CITEIA⟩2.0.CO;2.

Knote, C., G. Heinemann, and B. Rockel, 2010: Changes in weather extremes: Assessment of return values using high resolution climate simulations at convection-resolving scale. *Meteor. Z.*, **19**, 11–23, doi:10.1127/0941-2948/2010/0424.

Kyselý, J. and R. Beranová, 2009: Climate-change effects on extreme precipitation in central Europe: uncertainties of scenarios based on regional climate models. *Theor. Appl. Climatol.*, **95**, 361–374, doi:10.1007/s00704-008-0014-8.

Lamarque, J.-F., G. P. Kyle, M. Meinshausen, K. Riahi, S. J. Smith, D. P. van Vuuren, A. J. Conley, and F. Vitt, 2011: Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways. *Climatic Change*, **109**, 191–212, doi:10.1007/s10584-011-0155-0.

Langhans, W., J. Schmidli, O. Fuhrer, S. Bieri, and C. Schär, 2013: Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. *J. Appl. Meteor. Climatol.*, **52**, 1490–1510, doi:10.1175/JAMC-D-12-0167.1.

Lauwaet, D., P. Viaene, E. Brisson, N. van Lipzig, T. van Noije, A. Strunk, S. V. Looy, N. Veldeman, L. Blyth, K. D. Ridder, and S. Janssen, 2014: The effect of climate change and emission scenarios on ozone concentrations over Belgium: a high-resolution model study for policy support. *Atmos. Chem. Phys.*, **14**, 5893–5904, doi:10.5194/acp-14-5893-2014.

Leduc, M. and R. Laprise, 2009: Regional climate model sensitivity to domain size. *Clim. Dynam.*, **32**, 833–854, doi:10.1007/s00382-008-0400-z.

Lenderink, G. and E. van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nat. Geosci.*, **1**, 511–514, doi:10.1038/ngeo262.

Lenderink, G. and E. van Meijgaard, 2010: Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environ. Res. Lett.*, **5**, 9, doi:10.1088/1748-9326/5/2/025208.

Li, F., D. Rosa, W. D. Collins, and M. F. Wehner, 2012: "Super-parameterization": A better way to simulate regional extreme precipitation? *Journal of Advances in Modeling Earth Systems*, **4**, 10, doi:10.1029/2011MS000106.

Loriaux, J. M., G. Lenderink, S. R. De Roode, and A. P. Siebesma, 2013: Understanding convective extreme precipitation scaling using observations and an entraining plume model. *J. Atmos. Sci.*, **70**, 3641–3655, doi:10.1175/JAS-D-12-0317.1.

Luca, A. D., R. de Ela, and R. Laprise, 2012: Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations. *Clim. Dynam.*, **38**, 1229–1247, doi:10.1007/s00382-011-1068-3.

Lucas-Picher, P., F. Boberg, J. Christensen, and P. Berg, 2013: Dynamical downscaling with reinitializations: a method to generate fine-scale climate data sets suitable for impact studies. *J. Hydrometeor.*, **14**, 1159–1174, doi:10.1175/JHM-D-12-063.1.

Lucas-Picher, P., J. Cattiaux, A. Bougie, and R. Laprise, 2015: How does large-scale nudging in a regional climate model contribute to improving the simulation of weather regimes and seasonal extremes over North America? *Clim. Dynam.*, 1–20, doi:10.1007/s00382-015-2623-0.

Lynch, P., 2008: The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, 3431–3444, doi:10.1016/j.jcp.2007.02.034.

Lynn, B. H., C. Rosenzweig, R. Goldberg, D. Rind, C. Hogrefe, L. Druyan, R. Healy, J. Dudhia, J. Rosenthal, and P. Kinney, 2010: Testing GISS-MM5 physics configurations for use in regional impacts studies. *Climatic Change*, **99**, 567–587, doi:10.1007/s10584-009-9729-5.

Ma, L., T. Zhang, O. W. Frauenfeld, B. Ye, D. Yang, D. Yang, and D. Qin, 2009: Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *J. Geophys. Res.*, **114**, 20, doi:10.1029/2008JD011178.

Mahoney, K., M. Alexander, J. D. Scott, and J. Barsugli, 2013: High-resolution downscaled simulations of warm-season extreme precipitation events in the Colorado front range under past and future climate. *J. Climate*, **26**, 8671–8689, doi:10.1175/JCLI-D-12-00744.1.

Mailhot, A., S. Duchesne, D. Caya, and G. Talbot, 2007: Assessment of future change in intensity-duration-frequency (IDF) curves for Southern Quebec using the Canadian Regional Climate Model (CRCM). *J. Hydrol.*, **347**, 197–210, doi:10.1016/j.jhydrol.2007.09.019.

Mailhot, A., S. Lachance-Cloutier, G. Talbot, and A.-C. Favre, 2013: Regional estimates of intense rainfall based on the Peak-Over-Threshold (POT) approach. *J. Hydrol.*, **476**, 188–199, doi:10.1016/j.jhydrol.2012.10.036.

Malkmus, W., 1967: Random Lorentz band model with exponential-tailed S-1 line intensity distribution function. *Journal of the Optical Society of America*, **57**, 323–329, doi:10.1364/JOSA.57.000323.

Maraun, D., F. Wetterhall, A. M. Ireson, R. E. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. W. Rust, T. Sauter, M. Themeßl, V. K. C. Venema, K. P. Chun, C. M. Goodess, R. G. Jones, C. Onof, M. Vrac, and I. ThieleEich, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, 34, doi:10.1029/2009RG000314.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430, doi:10.1175/1520-0477(2002)083⟨0407:DIHRPM⟩2.3.CO;2.

Masson, V., 2000: A physically-based scheme for the urban energy budget in atmospheric models. *Bound.-Lay. Meteorol.*, **94**, 357–397, doi:10.1023/A:1002463829265.

Masson, V., J. L. Champeaux, F. Chauvin, C. Meriguet, and R. Lacaze, 2003: A global database of land surface parameters at 1 km resolution in meteorological and climate models. *J. Climate*, **16**, 1261–1282, doi:10.1175/1520-0442-16.9.1261.

Masson, V., P. L. Moigne, E. Martin, S. Faroux, A. Alias, R. Alkama, S. Belamari, A. Barbu, A. Boone, F. Bouyssel, P. Brousseau, E. Brun, J.-C. Calvet, D. Carrer, B. Decharme, C. Delire, S. Donier, K. Essaouini, A.-L. Gibelin, H. Giordani, F. Habets, M. Jidane, G. Kerdraon, E. Kourzeneva, M. Lafaysse, S. Lafont, C. Lebeaupin Brossier, A. Lemonsu, J.-F. Mahfouf, P. Marguinaud, M. Mokhtari, S. Morin, G. Pigeon, R. Salgado, Y. Seity, F. Taillefer, G. Tanguy, P. Tulet, B. Vincendon, V. Vionnet, and A. Voldoire, 2013: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of Earth surface variables and fluxes. *Geosci. Model Dev.*, **6**, 929–960, doi:10.5194/gmd-6-929-2013.

Mašek, J., J.-F. Geleyn, R. Brožková, O. Giot, H. O. Achom, and P. Kuma, 2015: Single interval shortwave radiation scheme with parameterized optical saturation and spectral overlaps. *Quart. J. Roy. Meteor. Soc.*, Online Version of Record published before inclusion in an issue, doi:10.1002/qj.2653.

May, W., 2008: Potential future changes in the characteristics of daily precipitation in Europe simulated by the HIRHAM regional climate model. *Climate Dyn.*, **30**, 581–603, doi:10.1007/s00382-007-0309-y.

Mayer, S., C. Maule, S. Sobolowski, O. Christensen, H. Sørup, M. Sunyer, K. Arnbjerg-Nielsen, and I. Barstad, 2015: Identifying added value in high-resolution climate simulations over Scandinavia. *Tellus A*, **67**, 18, doi:10.3402/tellusa.v67.24941.

Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell, 2007: The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394, doi:10.1175/BAMS-88-9-1383.

Menut, L., O. P. Tripathi, A. Colette, R. Vautard, E. Flaounas, and B. Bessagnet, 2013: Evaluation of regional climate simulations for air quality modelling purposes. *Climate Dyn.*, **40**, 2515–2533, doi:10.1007/s00382-012-1345-9.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock, 2004: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. *J. Geophys. Res.*, **109**, 14, doi:10.1029/2003JD004495.

Mishra, V., J. M. Wallace, and D. P. Lettenmaier, 2012: Relationship between hourly extreme precipitation and local air temperature in the United States. *Geophys. Res. Lett.*, **39**, 7, doi:10.1029/2012GL052790.

Mitchell, T. D. and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.*, **25**, 693–712, doi:10.1002/joc.1181.

Mlawer, E. J., S. J. Taubman, P. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682, doi:10.1029/97JD00237.

Morcrette, J., 1990: Impact of changes to the radiation transfer parameterizations plus cloud optical properties in the ECMWF model. *Mon. Wea. Rev.*, **118**, 847–873, doi:10.1175/1520-0493(1990)118⟨0847:IOCTTR⟩2.0.CO;2.

Muller, C. J., P. A. O'Gorman, and L. E. Back, 2011: Intensification of precipitation extremes with warming in a cloud-resolving model. *J. Climate*, **24**, 2784–2800, doi:10.1175/2011JCLI3876.1.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, 659–740.

Nakićenović, N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grübler, T. Y. Jung, T. Kram, E. L. L. Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Riahi, A. Roehrl, H.-H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, and Z. Dadi, 2000: Emissions scenarios. *IPCC 2000: Special Report on Emissions Scenarios (SRES). Working Group III, Intergovernmental Panel on Climate Change (IPCC)*, N. Nakićenović and R. Swart, Eds., Cambridge University Press, Vol. 1, 600.

Nicolis, C., 2003: Dynamics of model error: Some generic features. *J. Atmos. Sci.*, **60**, 2208–2218, doi:10.1175/1520-0469(2003)060⟨2208:DOMESG⟩2.0.CO;2.

Nicolis, C., 2004: Dynamics of model error: The role of unresolved scales revisited. *J. Atmos. Sci.*, **61**, 1740–1753, doi:10.1175/1520-0469(2004)061⟨1740:DOMETR⟩2.0.CO;2.

Niehörster, F., I. Fast, H. Huebener, and U. Cubasch, 2008: The stream one ENSEMBLES projections of future climate change. ENSEMBLES Technical Report no. 3, 30 pp.

Nikulin, G., E. Kjellström, U. Hansson, G. Strandberg, and A. Ullerstig, 2011: Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations. *Tellus A*, **63**, 41–55, doi:10.1111/j.1600-0870.2010.00466.x.

Noilhan, J. and J.-F. Mahfouf, 1996: The ISBA land parameterization scheme. *Global Planet. Change*, **13**, 145–159, doi:10.1016/0921-8181(95)00043-7.

Noilhan, J. and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.*, **117**, 536–549, doi:10.1175/1520-0493(1989)117⟨0536:ASPOLS⟩2.0.CO;2.

Ntegeka, V. and P. Willems, 2008: Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resour. Res.*, **44**, 15, doi:10.1029/2007WR006471.

O'Gorman, P. A. and T. Schneider, 2009: The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proc. Natl. Acad. Sci. (USA)*, **106**, 14 773–14 777, doi:10.1073/pnas.0907610106.

Overeem, A., A. Buishand, and I. Holleman, 2008: Rainfall depth-duration-frequency curves and their uncertainties. *J. Hydrol.*, **348**, 124–134, doi:10.1016/j.jhydrol.2007.09.044.

Pan, L.-L., S.-H. Chen, D. Cayan, M.-Y. Lin, Q. Hart, M.-H. Zhang, Y. Liu, and J. Wang, 2011: Influences of climate change on California and Nevada regions revealed by a high-resolution dynamical downscaling study. *Climate Dyn.*, **37**, 2005–2020, doi:10.1007/s00382-010-0961-5.

Pasquill, F., 1961: The estimation of the dispersion of windborne material. *Meteor. Mag.*, **90**, 33–49.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate*, **20**, 4356–4376, doi:10.1175/JCLI4253.1.

Persson, A., 2005: Early operational numerical weather prediction outside the USA: an historical introduction: Part II: Twenty countries around the world. *Meteor. Appl.*, **12**, 269–289, doi:10.1017/S1350482705001751.

Prein, A. F., A. Gobiet, M. Suklitsch, H. Truhetz, N. K. Awan, K. Keuler, and G. Georgievski, 2013a: Added value of convection permitting seasonal simulations. *Clim. Dynam.*, **41**, 2655–2677, doi:10.1007/s00382-013-1744-6.

Prein, A. F., G. J. Holland, R. M. Rasmussen, J. Done, K. Ikeda, M. P. Clark, and C. H. Liu, 2013b: Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters. *J. Climate*, **26**, 4848–4857, doi:10.1175/JCLI-D-12-00727.1.

Prein, A. F., W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli, N. P. M. van Lipzig, and R. Leung, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, doi:10.1002/2014RG000475.

Qian, J.-H., A. Seth, and S. Zebiak, 2003: Reinitialized versus continuous simulations for regional climate downscaling. *Mon. Wea. Rev.*, **131**, 2857–2874, doi:10.1175/1520-0493(2003)131⟨2857:RVCSFR⟩2.0.CO;2.

Radnóti, G., 1995: Comments on "a spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations". *Mon. Wea. Rev.*, **123**, 3122–3123.

Radu, R., M. Déqué, and S. Somot, 2008: Spectral nudging in a spectral regional climate model. *Tellus A*, **60**, 898–910, doi:10.1111/j.1600-0870.2008.00341.x.

Randall, D. A., R. A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R. J. Stouffer, A. Sumi, and K. E. Taylor, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, 589–662.

Ribatet, M., T. B. M. J. Ouarda, E. Sauquet, and J.-M. Gresillon, 2009: Modeling all exceedances above a threshold using an extremal dependence structure: Inferences on several flood characteristics. *Water Resour. Res.*, **45**, 15, doi:10.1029/2007WR006322.

Richard, J., B. Kirtman, R. Laprise, H. von Storch, and W. Wergen, 2002: Atmospheric regional climate models (RCMs): A multiple purpose tool? Report of the "Joint WGNE/WGCM ad hoc Panel on Regional Climate Modelling", 19 pp.

Ritter, B. and J.-F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.*, **120**, 303–325.

Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, doi:10.1175/2007MWR2123.1.

Roth, M., T. A. Buishand, G. Jongbloed, A. M. G. K. Tank, and J. H. van Zanten, 2012: A regional peaks-over-threshold model in a nonstationary climate. *Water Resour. Res.*, **48**, 12, doi:10.1029/2012WR012214.

Rowell, D. P. and R. G. Jones, 2006: Causes and uncertainty of future summer drying over Europe. *Climate Dyn.*, **27**, 281–299, doi:10.1007/s00382-006-0125-9.

Rummukainen, M., 2010: State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 82–96, doi:10.1002/wcc.8.

Russo, S. and A. Sterl, 2012: Global changes in seasonal means and extremes of precipitation from daily climate model data. *J. Geophys. Res.*, **117**, 11, doi:10.1029/2011JD016260.

Salas-Mélia, D., F. Chauvin, M. Déqué, H. Douville, J. F. Guérémy, P. Marquet, S. Planton, J. F. Royer, and S. Tyteca, 2005: Description and validation of the CNRM-CM3 global coupled model. *Climate Dyn.*, submitted manuscript draft.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio, 2009: Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372, doi:10.1175/2009MWR2924.1.

Seth, A. and F. Giorgi, 1998: The effects of domain choice on summer precipitation simulation and sensitivity in a regional climate model. *J. Climate*, **11**, 2698–2712, doi: 10.1175/1520-0442(1998)011⟨2698:TEODCO⟩2.0.CO;2.

Shepherd, T. G., 2014: Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.*, **7**, 703–708, doi:10.1038/ngeo2253.

Sillman, S., 1999: The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmos. Environ.*, **33**, 1821–1845, doi:10.1016/S1352-2310(98) 00345-8.

Skalák, P., P. Stepánek, and A. Farda, 2008: Validation of ALADIN-Climate/CZ for present climate (1961-1990) over the Czech Republic. *Quarterly Journal of the Hungarian Meteorological Service*, **112**, 191–201.

Soares, P. M. M., R. M. Cardoso, P. M. A. Miranda, J. Medeiros, M. Belo-Pereira, and F. Espirito-Santo, 2012a: WRF high resolution dynamical downscaling of ERA-Interim for Portugal. *Clim. Dynam.*, **39**, 2497–2522, doi:10.1007/s00382-012-1315-2.

Soares, P. M. M., R. M. Cardoso, P. M. A. Miranda, P. Viterbo, and M. Belo-Pereira, 2012b: Assessment of the ENSEMBLES regional climate models in the representation of precipitation variability and extremes over Portugal. *J. Geophys. Res.*, **117**, 18, doi: 10.1029/2011JD016768.

Stephenson, A., 2002: evd: Extreme value distributions. *R News*, **2**, 31–32.

Sylla, M. B., E. Coppola, L. Mariotti, F. Giorgi, P. M. Ruti, A. DellAquila, and X. Bi, 2010: Multiyear simulation of the African climate using a regional climate model (RegCM3) with the high resolution ERA-interim reanalysis. *Clim. Dynam.*, **35**, 231–247, doi:10. 1007/s00382-009-0613-9.

Tabari, H., R. De Troch, O. Giot, R. Hamdi, P. Termonia, S. Saeed, E. Brisson, N. van Lipzig, and P. Willems, 2016: Local impact analysis of climate change on precipitation extremes: Are high-resolution climate models needed for realistic simulations? *Hydrol. Earth Syst. Sci. Discuss.*, **In review**, doi:10.5194/hess-2016-106.

Termonia, P., A. Deckmyn, and R. Hamdi, 2009: Study of the lateral boundary condition temporal resolution problem and a proposed solution by means of boundary error restarts. *Mon. Wea. Rev.*, **137**, 3551–3566, doi:10.1175/2009MWR2964.1.

Termonia, P. and A. Quinet, 2004: A new transport index for predicting episodes of extreme air pollution. *J. Appl. Meteor.*, **43**, 631640, doi:10.1175/1520-0450(2004)043⟨0631: ANTIFP⟩2.0.CO;2.

Termonia, P., F. Voitus, D. Degrauwe, S. Caluwaerts, and R. Hamdi, 2012: Application of boyd's periodization and relaxation method in a spectral atmospheric limited-area model – Part I: implementation and reproducibility tests. *Mon. Wea. Rev.*, **140**, 3137–3148, doi: 10.1175/MWR-D-12-00033.1.

Thunis, P. and R. Bornstein, 1996: Hierarchy of mesoscale flow assumptions and equations. *J. Atmos. Sci.*, **53**, 380–397, doi:10.1175/1520-0469(1996)053⟨0380:HOMFAA⟩2.0.CO;2.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1217, doi:10.1175/BAMS-84-9-1205.

Tustison, B., D. Harris, and E. Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *J. Geophys. Res.*, **106**, 11,775–11,784, doi:10.1029/2001JD900066.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. V. D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.

Van Der Auwera, L., 1991a: *Histograms of wind speed (part A) and statistics of Pasquill stability classes (part B)*. Miscellanea SERIE B-65, Koninklijk Meteorologisch Instituut van België, 113 pp.

Van Der Auwera, L., 1991b: *Statistics of Pasquill stability classes (part B, recalculated)*. Miscellanea SERIE B-65 bis, Koninklijk Meteorologisch Instituut van België.

van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakićenović, S. J. Smith, and S. K. Rose, 2011: The representative concentration pathways: an overview. *Climatic Change*, **109**, 5–31, doi:10.1007/s10584-011-0148-z.

Váňa, F., P. Bénard, J.-F. Geleyn, A. Simon, and Y. Seity, 2008: Semi-lagrangian advection scheme with controlled damping: An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Quart. J. Roy. Meteor. Soc.*, **134**, 523–537, doi:10.1002/qj.220.

Warner, T., R. Peterson, and R. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **78**, 2599–2617.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548, doi:10.1175/1520-0493(1997)125⟨0527:TRDOEM⟩2.0.CO;2.

Westra, S., H. J. Fowler, J. P. Evans, L. V. Alexander, P. Berg, F. Johnson, E. J. Kendon, G. Lenderink, and N. M. Roberts, 2014: Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.*, **52**, 522–555, doi:10.1002/2014RG000464.

Willems, P., 2000: Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. *J. Hydrol.*, **233**, 189–205, doi:10.1016/S0022-1694(00)00233-X.

Willems, P., 2011: Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *J. Hydrol.*, **402**, 193–205, doi:10.1016/j.jhydrol.2011.02.030.

Willems, P., 2013: Revision of urban drainage design rules after assessment of climate change impacts on precipitation extremes at Uccle, Belgium. *J. Hydrol.*, **496**, 166–177, doi:10.1016/j.jhydrol.2013.05.037.

Willems, P., K. Arnbjerg-Nielsen, J. Olsson, and V. T. V. Nguyen, 2012: Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. *Atmos. Res.*, **103**, 106–118, doi:10.1016/j.atmosres.2011.04.003.

Zolina, O., A. Kapala, C. Simmer, and S. K. Gulev, 2004: Analysis of extreme precipitation over Europe from different reanalyses: a comparative assessment. *Global Planet. Change*, **44**, 129–161, doi:10.1016/j.gloplacha.2004.06.009.

A
MULTISCALE PERFORMANCE OF THE ALARO-0 MODEL FOR SIMULATING EXTREME DAILY SUMMER PRECIPITATION CLIMATOLOGY IN BELGIUM

De Troch, R., R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia

Multiscale Performance of the ALARO-0 Model for Simulating Extreme Summer Precipitation Climatology in Belgium

ROZEMIEN DE TROCH

Royal Meteorological Institute, Brussels, and Department of Physics and Astronomy, Ghent University, Ghent, Belgium

RAFIQ HAMDI AND HANS VAN DE VYVER

Royal Meteorological Institute, Brussels, Belgium

JEAN-FRANÇOIS GELEYN

Czech Hydrometeorological Institute, Prague, Czech Republic, and Department of Physics and Astronomy, Ghent University, Ghent, Belgium

PIET TERMONIA

Royal Meteorological Institute, Brussels, and Department of Physics and Astronomy, Ghent University, Ghent, Belgium

ABSTRACT

Daily summer precipitation over Belgium from the Aire Limitée Adaptation Dynamique Développement International (ALADIN) model and a version of the model that has been updated with physical parameterizations, the so-called ALARO-0 model [ALADIN and AROME (Application de la Recherche à l'Opérationnel à Meso-Echelle) combined model, first baseline version released in 1998], are compared with respect to station observations for the period 1961–90. The 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) is dynamically downscaled using both models on a horizontal resolution of 40 km, followed by a one-way nesting on high spatial resolutions of 10 and 4 km. This setup allows us to explore the relative importance of spatial resolution versus parameterization formulation on the model skill to correctly simulate extreme daily precipitation. Model performances are assessed through standard statistical errors and density, frequency, and quantile distributions as well as extreme value analysis, using the peak-over-threshold method and generalized Pareto distribution. The 40-km simulations of ALADIN and ALARO-0 show similar results, both reproducing the observations reasonably well. For the high-resolution simulations, ALARO-0 at both 10 and 4 km is in better agreement with the observations than ALADIN. The ALADIN model consistently produces too high precipitation rates. The findings demonstrate that the new parameterizations within the ALARO-0 model are responsible for a correct simulation of extreme summer precipitation at various horizontal resolutions. Moreover, this study shows that ALARO-0 is a good candidate model for regional climate modeling.

1. Introduction

Extreme precipitation events have a large impact on societies through damage caused by floods, landslides, and snow events. Precipitation is thus an important meteorological variable in weather prediction and climate studies. Herrera et al. (2010) studied the ability of regional climate models (RCMs) to reproduce the mean and extreme precipitation regimes over Spain using a state-of-the-art ensemble of RCM simulations. The RCMs show good agreement with the observed mean precipitation regime, but for the extreme regimes the models reveal important limitations.

As described in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), the model skill to simulate realistic extreme daily precipitation strongly depends on the spatial resolution and convective parameterization of the model (Randall et al. 2007). However, it is not straightforward to quantify the relative contribution of an increase in spatial resolution

Corresponding author address: Rozemien De Troch, Royal Meteorological Institute, Ringlaan 3, B-1180 Brussels, Belgium.
E-mail: rozemien.detroch@meteo.be

versus an improvement in physical parameterization of deep convection on the overall performance of the model.

On the other hand, precipitation is one of the most sensitive quantities in the different parameterization schemes of the climate models and to their interplay with the dynamics of the atmosphere represented in the models. For this variable it has been shown that RCMs are able to add significant information to the driving global simulations, both in space and time (e.g., Jones et al. 1995; Durman et al. 2001; Jones et al. 2004). In general terms, the RCMs produce an intensification of precipitation with respect to the driving global climate model (GCM), related to the intensification of the hydrological cycle (Jones et al. 1995; Durman et al. 2001; Buonomo et al. 2007). Lynn et al. (2010) tested a regional climate model with different physics components at two different spatial resolutions. Their results demonstrated a sensitivity of the RCM to the choice of the convective parameterization, leading to significantly different summer precipitation outcomes. The authors conclude that these differences are due to differences in the convective parameterizations and not because of the change in spatial resolution of the model.

The aim of the present paper is to elaborate on the relative importance of resolution versus parameterization formulation on the model skill to simulate realistic extreme daily precipitation. This is achieved by comparing at varying horizontal resolutions the Aire Limitée Adaptation Dynamique Développement International (ALADIN) model with a version of the model that has been updated with physical parameterizations, the so-called ALARO-0 model [ALADIN and AROME (Application de la Recherche à l'Opérationnel à Meso-Echelle) combined model, first baseline version released in 1998]. The ALADIN model is the limited area model (LAM) version of the Action de Recherche Petite Echelle Grande Echelle Integrated Forecast System (ARPEGE-IFS) (Bubnová et al. 1995; ALADIN International Team 1997). Since the 1990s the model has been widely used by the numerical weather prediction (NWP) community and, more recently, in regional climate modeling (e.g., Radu et al. 2008; Skalák et al. 2008). Furthermore, the model uses a diagnostic-type deep convection and microphysics parameterization based on Bougeault (1985) with upgrades from Gerard and Geleyn (2005). The new physical parameterizations within the ALARO-0 model, as proposed by Gerard et al. (2009), were specifically designed to be used from mesoscale to the convection-permitting scales (so-called gray-zone scales) and are centered around an improved convection and cloud scheme. For this study we use the version of the ALARO-0 model that was adopted for the operational applications in the Royal Meteorological Institute (RMI)

of Belgium in 2010. Since then this model has undergone systematic verification with respect to observations at 7-km resolution. Gerard et al. (2009) tested the new parameterizations within the ALARO-0 model in a 1-day case study over Belgium, which was characterized by heavy convective precipitation. From this study an improvement of ALARO-0 at varying horizontal scales has been demonstrated.

Basically, the "nesting" strategy, or climate downscaling technique, in which a LAM or RCM is driven by either a GCM or by analyses of observations, is the most widely used strategy to produce high resolution over a region of interest (Denis et al. 2002). Hence, limiting the geographical domain of these atmospheric models reduces the total number of grid points and allows one to perform simulations at high resolutions with an affordable computational cost. Because of the ability of these high-resolution LAMs or RCMs to reproduce meaningful small-scale features over a limited region (Denis et al. 2002; Giorgi et al. 2004), they have become a popular tool in both the NWP and the climate community for studying extreme events at regional and local scales (e.g., Jones and Reid 2001; Buonomo et al. 2007; Déqué and Somot 2008; Dulière et al. 2011).

However, studies show that RCMs do not necessarily improve their driving GCM simulations or global reanalyses (e.g., Castro et al. 2005; Jacob et al. 2007; Sylla et al. 2010). The use of nested LAMs or RCMs as a climate downscaling technique, indeed, involves a number of issues, one of which is related to the lateral boundary conditions (LBCs) (Giorgi and Mearns 1999; Denis et al. 2002). This drawback of RCMs is related to the fact that one is obliged to impose imperfect LBCs, inducing various errors at the boundaries (e.g., Warner et al. 1997; Termonia et al. 2009). Despite this, past and current applications with RCMs have shown that the one-way nesting strategy is a workable solution (Giorgi and Mearns 1999). To minimize the effects of the LBC problem, Giorgi and Mearns (1999) recommend to first validate the model for the current climate using analyses of observations, that is, the so-called perfect boundary conditions.

Interesting work has been carried out by de Elía et al. (2002) and Denis et al. (2002) with a perfect-model approach, showing that, in a downscaling with a one-way nesting, a LAM or RCM is able to regenerate the correct amount of variability at the scales smaller than the ones of the driving model in which the high-resolution variability had been removed by filtering. However, de Elía et al. (2002) found that the LAM is not capable of reproducing the correct details with sufficient precision required by the rms errors (RMSEs), that is, that the variables locally in space and time do not fully reproduce the ones of the perfect model run. Whereas de Elía et al. concentrated on the short-term evolution of weather

systems and quantified the models' ability to simulate the data in a deterministic day-by-day basis by means of RMSEs, Denis et al. focused on climate time scales and demonstrated the ability of high-resolution RCMs to gain accuracy in a climatic–statistical sense.

Therefore, for studying the climate of weather extremes it is rather the statistics of the extremes that are important, provided the large-scale evolution is consistent with the large-scale flow of the driving model. This is an important additional criterion in deciding to use RCMs with respect to global ones.

For long-range runs at temporal scales of multiple decades, there is also the problem that the internal climate can start to diverge from the climate of the global model (Nicolis 2003; Qian et al. 2003; Nicolis 2004). One can deal with this by either (i) interrupting the model runs of the LAM after a few days and restarting them, while allowing a spinup period so that the physics can adjust, or (ii) carrying out uninterrupted model runs over long periods, allowing the LAM to find its own climate equilibrium (Qian et al. 2003). In the second case, one can for instance apply a spectral nudging of the large scales to the large scale of the driving global model. In the present paper, we will also study whether the internal climate variability generated by the higher resolution of the RCM and its model physics, as identified by Denis et al. (2002) and de Elía et al. (2002), reproduces the correct statistics. For this we want to avoid imposing an upper-air spectral nudging; hence, we will merely carry out a pure downscaling with reinitializations using a one-way nesting approach. Lucas-Picher et al. (2013) demonstrated that dynamical downscaling with reinitializations has lower systematic errors than with a standard continuous model configuration.

The 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) (Uppala et al. 2005) is used as large-scale coupling data to drive the coupled models, ALARO-0 and ALADIN. As suggested by Giorgi and Mearns (1999), atmospheric reanalyses, such as ERA-40, can be used in climate studies to provide the "perfect boundary conditions" for RCMs (e.g., Csima and Horányi 2008; Déqué and Somot 2008; Skalák et al. 2008; Heikkilä et al. 2011; Hamdi et al. 2012). These reanalyses are produced by means of data assimilation methods in order to find optimal estimates for past atmospheric states that are consistent with meteorological observations and the model dynamics.

In a recent study of Hamdi et al. (2012) the use of high-resolution dynamical downscaling of ALARO-0 at 4-km horizontal resolution is explored by means of the summer maximum surface air temperature over Belgium. Our study extends the work of Hamdi et al. in the sense that, instead of temperature, precipitation is now

analyzed. Daily summer precipitation from different model runs is compared with respect to station observations, with an emphasis on extreme precipitation. This approach by which model output is directly compared against station observations can be motivated by the fact that the station-level observations provide the closest representation of extreme events (Dulière et al. 2011). Furthermore, the motivation for only considering summer precipitation is threefold: (i) other regional climate studies (e.g., Caldwell et al. 2009; Soares et al. 2012a,b) show difficulties of RCMs to simulate summer precipitation; (ii) the new parameterization scheme within ALARO-0 mostly modifies convection, which is the process most relevant for (extreme) precipitation events in summer (Kyselý and Beranová 2009; Soares et al. 2012a); and (iii) the relatively small scale on which these convective processes often occur better corresponds to the high-resolution ALARO-0 simulation (Kyselý and Beranová 2009).

We add to our evaluation the ALADIN-Climate model developed by the Centre National de Recherches Météorologiques (CNRM), which took part in the European ENSEMBLES project (www.ensembles-eu.org). The ALADIN-Climate model is an ALADIN model version that is specifically used for regional climate modeling. The Ensemble-Based Predictions of Climate Changes and their Impacts (ENSEMBLES) project was finished near the end of 2009 and is aimed to develop an ensemble climate forecast system to produce probabilistic scenarios of future climate so as to provide detailed, quantitative, and policy-relevant information to the European society and economy. Several experiments were performed with some 10 state-of-the-art European and Canadian high-resolution, global, and regional climate models. The ENSEMBLES ALADIN-Climate/CNRM simulations use a long uninterrupted model run, which is a different setup than our ALADIN and ALARO-0 simulations. Hence, a direct comparison with the ALADIN-Climate/CNRM simulation is not possible, and these uninterrupted climate runs are merely added as a reference for regional climate modeling in order to make the present paper complete.

The model simulations, experimental design, and observational data used in this study are described in the next section. Section 3 gives a description of the applied methods, and the results are discussed in section 4. The results are summarized in the conclusions in section 5.

2. Model description and data

a. Experimental design

The experimental design is summarized in Table 1. The ERA-40 reanalysis data (Uppala et al. 2005) are

TABLE 1. Overview of the experimental design.

	Reference	Daily cumulated precipitation	Model	Daily cumulated precipitation
1) Effect of downscaling	Station observations	0800 LT (day) → 0800 LT (day + 1)	ERA-40 ALR40 ALD40	0600 UTC (day) → 0600 UTC (day + 1)
2) Multiscale performance of ALARO-0	Station observations	0800 LT (day) → 0800 LT (day + 1)	ALR40 ALD40 ALR10 ALD10 ALR04	0600 UTC (day) → 0600 UTC (day + 1)
3) Reference for regional climate modeling	Station observations	0800 LT (day) → 0800 LT (day + 1)	CNRM	mean (0000–2400 UTC)

dynamically downscaled using the limited-area models ALADIN and ALARO-0.

The physics parameterization package of the ALARO-0 model has been specifically designed to be run at convection-permitting resolutions. The key concept behind the package lies in the precipitation and cloud scheme called Modular Multiscale Microphysics and Transport (3MT) developed by Gerard and Geleyn (2005), Gerard (2007), and Gerard et al. (2009). With mesh sizes mostly below the Rossby radius of deformation for convective phenomena, the parameterization schemes must take into consideration that the return current from updrafts is happening in a multitude of grid boxes. Therefore, each individual grid-box realization of the parameterization has a statistical view of the ''compensating subsidence'' happening inside its area. As long as the updraft computation can also be considered as statistical with respect to its population of updrafts of various depths and sizes, it seems not to matter much that the compensating subsidence is computed on the basis of a purely local closure. But when mesh sizes become so small that only a few updraft realizations happen inside each grid box, and with area fractions that cease to be negligible with respect to ''one,'' the whole concept of ''classical'' convective parameterization schemes collapses. In the 3MT scheme this problem is addressed by combining three key features of the scheme: (i) the separately computed deep convective condensation and large-scale condensation are merged as single input for a ''prognostic–geometric'' set of microphysical computations (sedimentation, auto-conversion, collection and melting–evaporation during fall); (ii) the convective detrainment is not diagnosed independently but becomes the result of the combined computations of closure, entrainment, and condensation; and (iii) the closure assumption (core of the physics–dynamics coupling) is a prognostic-type one with memory of the updraft area fraction and of the updraft vertical velocity of previous time steps. These three interrelated characteristics of 3MT induce a good multiscale performance of 3MT, in particular in the gray zone. The latter can be defined as the range of horizontal mesh sizes for which the precipitating convection is partly parameterized and partly simulated by the resolved motions of the model. If nothing specific is done (i.e., using the classical diagnostic-type schemes of, e.g., ALADIN at gray-zone scales), this ambivalence results in double-counting or double-void situations, leading to several negative ''gray-zone syndromes.'' In convective situations drizzle appears nearly everywhere, and the precipitation maxima are too intense and too scattered. This happens especially over mountainous areas.

The multiscale performance of 3MT has been validated in a numerical weather prediction context up to a spatial resolution of 4 km (see Gerard et al. 2009). The ALARO-0 model utilizes 1) the Action de Recherche Petite Echelle Grande Echell (ARPEGE) Calcul Radiatif avec Nebulosité (ACRANEB) scheme for radiation (Ritter and Geleyn 1992, recast in a Net Exchanged Rate framework), 2) a semi-Lagrangian horizontal diffusion scheme (SLHD) (Váňa et al. 2008), 3) some pseudo-prognostic turbulent kinetic energy (pTKE) scheme (i.e., a Louis-type scheme for stability dependencies, but with memory, advection, and autodiffusion of the overall intensity of turbulence), and 4) a statistical sedimentation scheme for precipitation within a prognostic-type scheme for microphysics (Geleyn et al. 2008). The physics package of the ALARO-0 model is coupled to the dynamics of the ALADIN model (Bubnová et al. 1995) via a physics–dynamics interface based on a flux-conservative formulation of the equations proposed by Catry et al. (2007).

For the present study, the same land surface model—Interactions between Soil, Biosphere, and Atmosphere (ISBA) (Noilhan and Planton 1989)—is used in both the ALARO-0 and ALADIN models. Furthermore, both models can be run with different schemes to impose the lateral-boundary conditions (Davies 1976; Radnóti 1995;

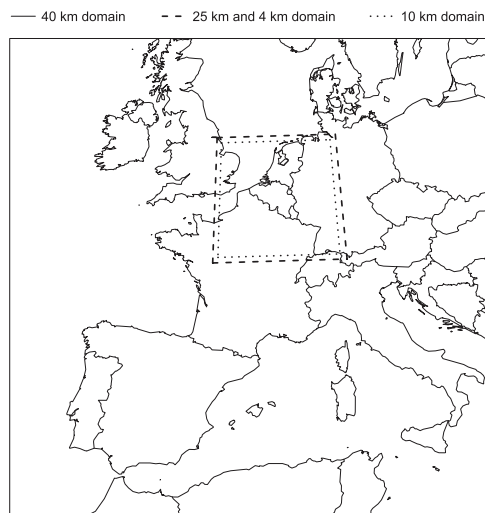—— 40 km domain   – – 25 km and 4 km domain   ···· 10 km domain

FIG. 1. Domains corresponding to the different simulations at 40-, 25-, 10-, and 4-km horizontal resolution.

Termonia et al. 2012). For this study, the version of Radnóti (1995) is used in both models.

The ALARO-0 model runs operationally in a number of countries of the ALADIN and High-Resolution Limited-Area Model (HIRLAM) consortia (Austria, Belgium, Czech Republic, Croatia, Hungary, Norway, Portugal, Romania, Sweden, Slovenia, Slovakia, and Turkey) for the national NWP applications, the first of them already since 2008. More recently, the model is also used for climate runs. The ALARO-0 model is developed and maintained mainly through a collaboration between the RMI of Belgium and the Regional Co-operation for Limited Area Modelling for Central Europe (RC LACE). The developments of the ALARO-0 model (intentionally targeted at the gray-zone scales) are centered around the 3MT basic concept, which means that many other parameterization schemes must be adapted to the use of 3MT, but also sometimes the reverse. Thus, a rather wide international effort is needed.

As the first step of this study, the improvement of the downscaling by means of the ALADIN and ALARO-0 models is examined. This is done by comparing recent past (1961–90) summer precipitation data from an ALARO-0 and ALADIN simulation performed at 40-km spatial resolution (ALR40 and ALD40) (Fig. 1) with summer precipitation from the driving ERA-40 reanalysis data (Uppala et al. 2005).

Despite the fact that reanalysis data products are more continuous in space and time than station data, they inevitably contain biases. A number of evaluations for ERA-40 reanalysis precipitation have been performed (e.g., Zolina et al. 2004; Ma et al. 2009). The ERA-40 precipitation has distinct regional limitations: most of them are generally related to the coarse horizontal resolution of the ERA-40 model, on one hand, and to its strong model dependency, on the other (Ma et al. 2009). All physical parameterizations within ERA-40, including those of precipitation, were run on a spatial resolution of about 125 km (Zolina et al. 2004; Ma et al. 2009). The model diagnostics precipitation in ERA-40 is produced by parameterized microphysical processes in clouds, which are formed at supersaturation by convective or large-scale processes (Ma et al. 2009). Total precipitation is then simply the sum of the convective precipitation generated by convective clouds and large-scale stratiform precipitation, associated with frontal or dynamical systems (Zolina et al. 2004). Hence, ERA-40 precipitation is a pure model product. Due to the poor skill of operational NWP models to account for all important physical mechanisms that affect the atmospheric water cycle, it appears to be one of the most uncertain forecasted parameters in the reanalysis (Zolina et al. 2004; Ma et al. 2009; Heikkilä et al. 2011). The 6-hourly forecasts from the ERA-40 reanalysis are used to calculate daily cumulated summer precipitation between 0600 and 0600 UTC of the next day. For coupling to the regional model we use a linear interpolation in time. This may produce errors at the lateral boundaries on our small domains (Fig. 1) but, as shown by Termonia et al. (2009), such errors only occur very rarely, and the impact on the statistics of extreme precipitation should be very minor.

To explore further the multiscale performance of ALARO-0, as found by Gerard et al. (2009) but now for climate time scales, we evaluate in a second step recent past simulations (1961–90) of the ALADIN and ALARO-0 models at varying horizontal resolutions against different station datasets.

(i) and (ii) The ALADIN and ALARO-0 models are driven by ERA-40 and run at a horizontal resolution of 40-km spatial resolution with 69×69 grid points on a domain that encompasses most of western Europe (ALD40 and ALR40, respectively; Fig. 1). These 40-km outputs are then used to perform a one-way nesting on a domain centered on Belgium (Fig. 1) using the following spatial resolutions:

(iii) and (iv) 10-km spatial resolution on a 67×67 grid (ALD10 and ALR10) and

(v) 4-km spatial resolution on a 181×181 grid (ALR04). That we did not run any ALD04 configuration is obviously linked to the corresponding gray-zone-type resolution, where the diagnostic

parameterization of convection would have become completely irrelevant (see section 4 for the first syndromes already noticeable in ALD10).

Finally, we also include ALADIN-Climate/CNRM simulations within our analysis so as to provide a reference for regional climate modeling. One part of the performed experiments within the ENSEMBLES project aimed to validate the models for the recent past climate. The results from this experiment, including 40 years of 25-km resolution ALADIN-Climate/CNRM simulations driven by the ERA-40 reanalysis (hereafter denoted as CNRM), are used in our analysis for the period 1961–90. From the ENSEMBLES data archive we have only selected the CNRM precipitation data for the grid points that coincide with the ALR04 domain (Fig. 1). The precipitation data correspond to daily means calculated for the interval 0000–2400 UTC. As mentioned in section 1, the model setup of CNRM and our simulations are different. The number of vertical levels that is used in our runs with the ALADIN and ALARO-0 models is 46 with a model top that extends up to 72 km. The CNRM simulations from ENSEMBLES have used 31 vertical levels. Furthermore, the CNRM simulations use a long-term and free run setup. Our procedure is to interpolate the original ERA-40 files to 40-km resolution. These 6-h files serve as initial and boundary conditions for 48-h ALD40 and ALR40 runs. These are started at 0000 UTC every day. The (3 h) output from these first runs serves as input for the high-resolution 10- and 4-km runs (ALD10, ALR10, and ALR04). However, to exclude spinup problems, the first 12 h are not taken into account. So we have 36 h of data left for the 4- and 10-km runs (which thus start at 1200 UTC). Finally, we again dismiss the first 12 h of the runs, to arrive at 24 h of output at 4- and 10-km resolution, and then integrate/reinitialize over each subsequent 24-h period during the summer period of June–August, 1961–90.

b. Observations

The observation dataset comprises 93 climatological stations with daily accumulated precipitation, selected from the climatological network of the RMI of Belgium. The data have undergone a manual quality control by operators, and the stations were chosen so that continuous data for the 30-yr study period (1961–90) are available. The stations cover all of Belgium, thus representing conditions of coastal, inland, and higher orographic locations (Fig. 2).

3. Methods

a. Data processing and analysis

Model validation against observations can either be done with station data or gridded station data. Both
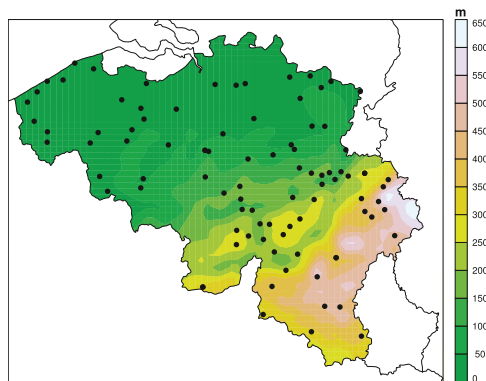

FIG. 2. Topography (m) of Belgium showing the location of the 93 selected climatological stations (black dots).

validation methods have their disadvantages (Hofstra et al. 2010). Model evaluation against observations at station level often raise issues related to the scale difference between the model and observation field (Tustison et al. 2001; Dulière et al. 2011). The model grid cell values correspond to spatially averaged values representing the area of the whole grid cell. Furthermore, the spatial variability of these averaged model fields will always be lower than the one of the observation field. These differences in spatial variability depend on the area of the grid cell as well as on the inherent variability of the field variable. Precipitation, for example, is known to have a relatively high spatial variability. To illustrate the differences in spatial variability in this study, Fig. 3 shows the different grid cell areas of the models together with the 93 climatological stations (i.e., observation points). The grid cell areas in this study range from $1600 \, km^2$ for the 40-km horizontal resolution to $16 \, km^2$ for the 4-km horizontal resolution (Fig. 3). Hence, reducing those spatially averaged model values with an originally greater heterogeneity to a single station point value leads to an inconsistent comparison. However, for long time periods, such as 30 years, we can assume that the spatial variability within a grid cell would be reduced in such a way that the spatial variability of both model and observation fields tends to converge (Dulière et al. 2011).

Another common way to overcome this scale inconsistency is the use of gridded data. The Climate Research Unit (CRU) and the European ENSEMBLES project provide daily gridded observation datasets (Mitchell and Jones 2005; Haylock et al. 2008). However, these gridded datasets are in some regions constructed by interpolation or area-averaging of station observations from a small number of stations, which

40 km model points

25 km model points (CNRM)
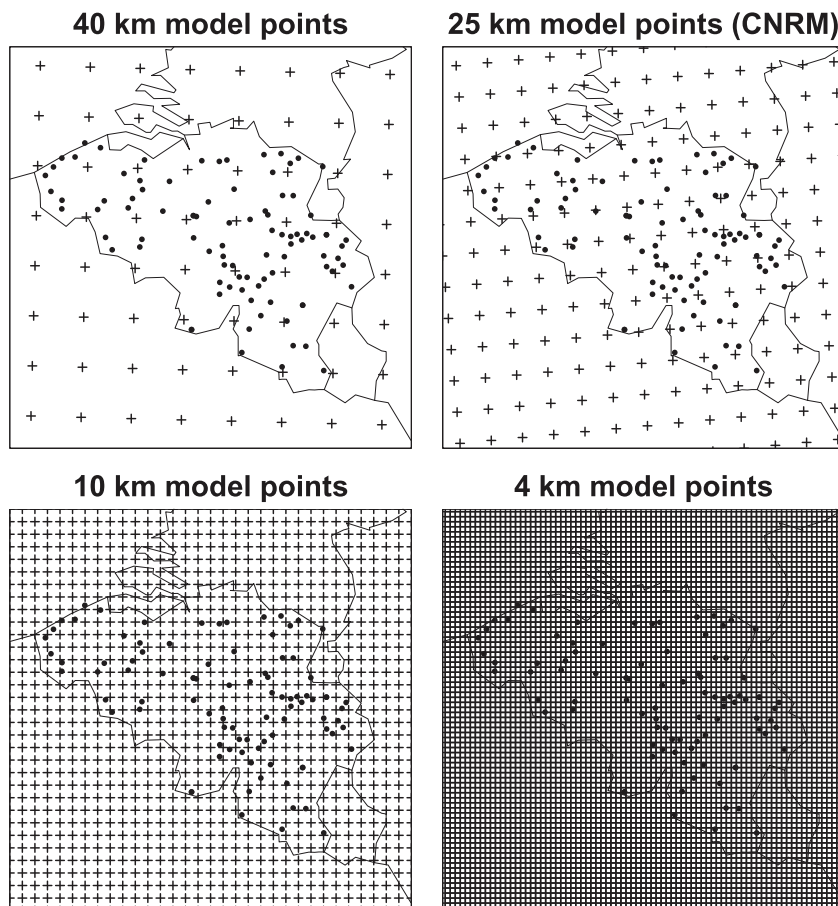
10 km model points

4 km model points

FIG. 3. Model grid points over Belgium for each of the horizontal resolutions for which the simulations are performed. The black dots represent the 93 climatological stations.

smooths and possibly affects the extreme values within the dataset (Hofstra et al. 2010). Since this study aims to examine extreme precipitation events, the models are evaluated against station observations. This is done by comparison of daily observed station-level precipitation with modeled daily precipitation of the nearest grid box over land. The 93 resulting precipitation time series selected from the model simulations are not corrected for topography with respect to altitude of the nearest station. It is difficult to apply such correction for precipitation because of its dependency on topography, humidity, buoyancy, and other local variables (Soares et al. 2012a).

Time discrepancy between computations of daily cumulated precipitation from station observations and model output is an important, but rarely highlighted, problem within precipitation evaluation studies. To deal with this problem, the error analysis can be performed on longer than daily time scales, such as monthly, seasonal, or annual time scales (Ma et al. 2009; Soares et al. 2012b). However, in this study the model evaluation is done on a daily basis, requiring a consistent calculation of the daily precipitation values. Daily observed precipitation corresponds to the total accumulated precipitation between 0800 and 0800 local time (LT) of the following day. Hence, the daily model values for all

simulations (ALR40, ALD40, ALR10, ALD10, and ALR04) have been calculated based on the definition of observed daily accumulation, which corresponds to 0600 and 0600 UTC of the following day (Table 1).

b. *Extreme value analysis and peak-over-threshold methods*

The methods used for the modeling of extreme events are similar to those used in Hamdi et al. (2012). Threshold models and peak-over-threshold (POT) methods are useful tools for the modeling of extreme events. A well-known distribution that may describe the behavior of the excesses or POT events is the generalized Pareto distribution (GPD) (Coles 2001). Recently, several authors have modeled extreme precipitation with the GPD (e.g., Ribatet et al. 2009; Roth et al. 2012; Mailhot et al. 2013).

Consider a sequence of independent and identically distributed random variables $X_1, X_2, \ldots, X_i$ from an unknown distribution F. We are interested in the extreme events that exceed a certain high threshold u. The distribution function of such an extreme event X from the X_i sequence can then be defined as

$$F_u(y) = P\{X > u + y \mid X > u\} = \frac{1 - F(u + y)}{1 - F(u)}, \quad (1)$$

with $y > 0$. Equation (1) is the conditional probability that the threshold u is exceeded by no more than an amount y, given that the threshold u is exceeded. Given that $X > u$, the GPD of the excesses $(X - u)$ is then given by

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi}, \quad (2)$$

where ξ is the shape parameter and σ is the scale parameter. The GPD with parameters ξ and σ describes the limiting distribution for the distribution of excesses [Eq. (1)] and can be used to model the exceedances of a threshold u by a variable X. Thus, for $x > u$,

$$P\{X > x \mid X > u\} = \left[1 + \xi\left(\frac{x - u}{\sigma}\right)\right]^{-1/\xi}. \quad (3)$$

It follows that

$$F(x) = P\{X > x\} = \zeta_u\left[1 + \xi\left(\frac{x - u}{\sigma}\right)\right]^{-1/\xi}, \quad (4)$$

where $\zeta_u = P\{X > u\}$. In this study the parameters of the GPD are estimated by the maximum-likelihood method, following the definitions of Stephenson (2002). The level
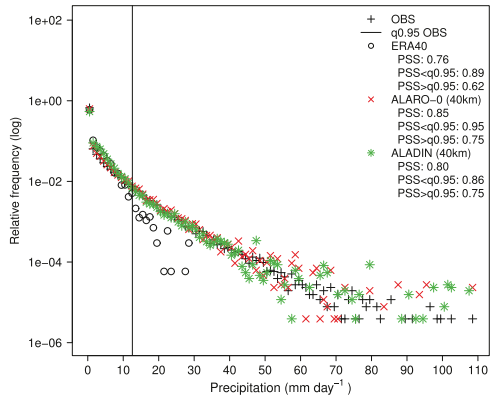

FIG. 4. Relative frequencies of observations, ERA-40, ALR40, and ALD40. Frequencies are computed with the 30-yr (1961–90) daily cumulated summer precipitation given for each station separately and are displayed on a logarithmic scale. Numbers for PSS correspond to the average of the Perkins skill score [Eq. (7)] calculated for precipitation amounts below and above the 0.95th quantile of the observations (PSS < q0.95 and PSS > q0.95). The black line indicates the 0.95th quantile of the observations.

x_m that is on average exceeded once every m observations is the solution of

$$\zeta_u\left[1 + \xi\left(\frac{x_m - u}{\sigma}\right)\right]^{-1/\xi} = \frac{1}{m}. \quad (5)$$

The x_m return level, which gives the amount of extreme precipitation corresponding to a given number of observations m, is then given by

$$x_m = u + \frac{\sigma}{\xi}[(m\zeta_u)^\xi - 1]. \quad (6)$$

4. Results and discussion

a. *Effect of downscaling*

As a first step we validate the effect of the downscaling of the ERA-40 with the ALADIN and ALARO-0 models. Figure 4 shows the relative frequencies calculated for daily precipitation amounts of ERA-40, ALR40, and ALD40, which are binned into bins of 1 mm day^{-1}. As a reference the relative frequencies of the observations are also shown. A logarithmic scale has been used for better representation of the extreme values. From both ERA-40 data and the ALR40 and ALD40 data 93 grid points, corresponding to the closest grid points to the observation stations, have been selected. It should be noted that the ERA-40 only has two grid points over

Belgium. For low precipitation amounts (i.e., <0.95th quantile of the observations) the ERA-40 as well as ALR40 and ALD40 coincide well with the observations. However, for the higher rainfall rates ERA-40 starts to diverge from the observations, while ALR40 and ALD40 still approach the observations. Both 40-km models are able to reproduce rainfall rates up to 108 mm day^{-1}, while the reanalysis does not capture the higher precipitation amounts due to the low spatial resolution of the ERA-40 data. To provide a measure of similarity between observed and modeled frequencies, the Perkins skill score (PSS) has been calculated (Perkins et al. 2007):

$$PSS = \sum_1^n \min(Z_1, Z_2), \qquad (7)$$

where n is the number of bins and $Z_{1,2}$ is the frequency of values in a given bin from the observation and model data, respectively. This metric measures how well the observations and modeled frequencies coincide, with a PSS ranging from zero for no overlap to a skill score of one for a perfect overlap. Similar to Boberg et al. (2010) and Domínguez et al. (2013), the PSS has been calculated for daily precipitation amounts going from 0 mm day^{-1} up to the 0.95th quantile of the observations (PSS < q0.95) and for precipitation amounts above the 0.95th quantile of the observations (PSS > q0.95). In this way, the skill score is to a larger extent influenced by the more extreme precipitation values (Boberg et al. 2010). The skill scores are calculated for each station separately. The final PSS is then simply the mean value of the average of PSS < q0.95 and PSS > q0.95 over the 93 stations. The 0.95th quantile of the observations, which is used as a threshold for the calculation of the modified PSS, is also shown in Fig. 4. The Perkins skill scores for ERA-40 are relatively low, and for the higher precipitation amounts ERA-40 has a much lower PSS (PSS > q0.95: 0.62) than ALR40 and ALD40 (PSS > q0.95: 0.75). ALR40 and ALD40 perform very similar with respect to the observations and have relatively high PSS, which are close to one. To summarize, the downscaling with the ALARO-0 and ALADIN models is significantly different from the driving ERA-40 and is closer to the observations. In particular, ALR40 and ALD40 produce more extreme precipitation than their driving ERA-40.

b. Multiscale performance of ALARO-0

To investigate the multiscale performance of ALARO-0, 40-, 10-, and 4-km horizontal resolution simulations of ALARO-0 together with 40- and 10-km horizontal

resolution simulations of ALADIN are compared with respect to station observations.

1) SPATIAL AND TEMPORAL DISTRIBUTION

Figure 5 shows the observed and simulated spatial distribution of the 30-yr-averaged summer precipitation. On top of each subfigure average values over the 93 stations for the cumulated summer precipitation are given. On average all models except for CNRM overpredict the observed cumulated summer precipitation. Both observation and simulation fields show a clear topographical dependency, with a gradual increase in precipitation going from the northwest (low altitudes) to the southeast (high altitudes) of the country. The ALARO-0 and ALADIN simulation at 40 km show a very similar distribution. Obviously, the precipitation fields for the simulations with low spatial resolution are less heterogeneous than the ones with high spatial resolution. However, the 25-km spatial resolution CNRM plot illustrates less variability than the 40-km simulations: also, the local maximum in the southeast cannot been seen on the CNRM plot. For the higher-resolution simulations ALARO-0 approaches much better the observations than ALADIN. For instance, ALD10 overpredicts cumulated summer precipitation with values that are, on average, over all stations almost 100 mm higher than observed. On the contrary, the average values for ALR10 and ALR04 differ only slightly from the observations, and the observed local maximum at the higher altitudes is very well simulated by both models.

The scatterplots presented in Fig. 6 are consistent with the spatial distributions shown in Fig. 5. Each point in the scatterplots represents the summer cumulated precipitation for each year in the 30-yr period averaged for the 93 stations. The linear regression line (solid line) and its determination coefficient (R^2) is also presented for each of the five models. Except for ALD10, summer precipitation is relatively well simulated by all models. The ALD10 model shows again a clear overestimation of observed summer precipitation. This is an indirect confirmation that, with 10-km mesh sizes, the syndromes linked to the gray-zone performance are already present (see section 2a).

2) ERROR STATISTICS

The previous analysis showed the ability of the models to represent the spatial and temporal pattern of mean annual summer precipitation. To quantify this ability we have computed some important error statistics. Figure 7 shows the spatial distribution of the 30-yr average summer biases of the daily cumulated precipitation, as well as the mean bias over the 93 climatological stations.
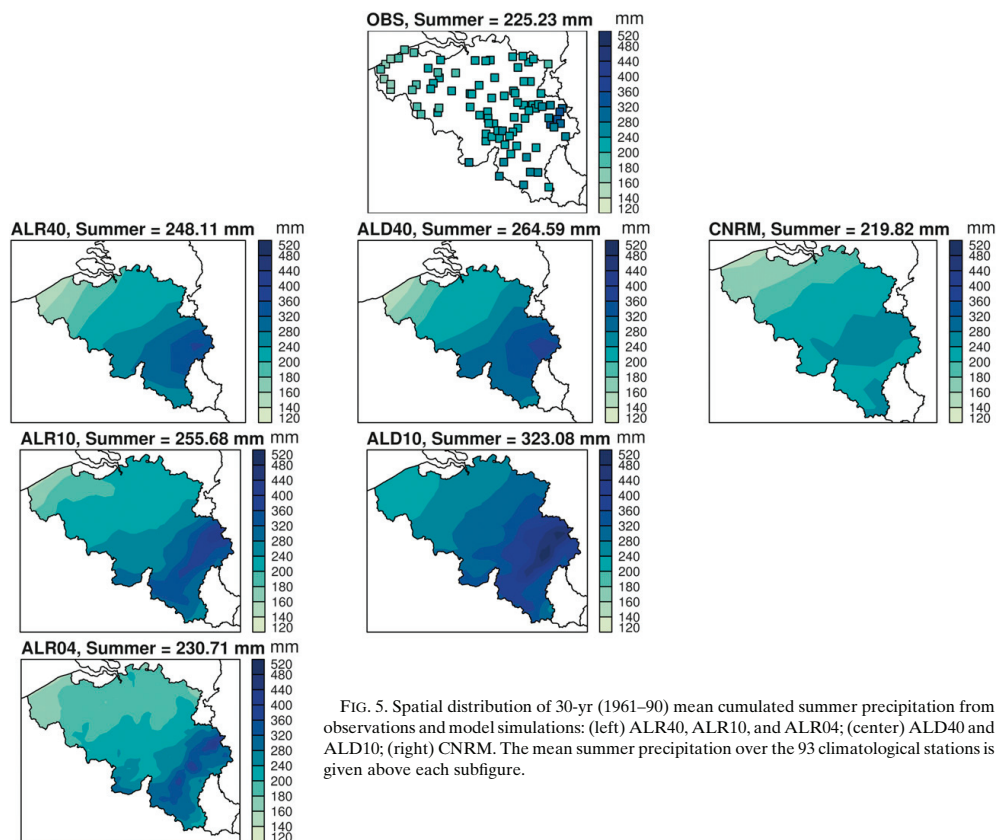
FIG. 5. Spatial distribution of 30-yr (1961–90) mean cumulated summer precipitation from observations and model simulations: (left) ALR40, ALR10, and ALR04; (center) ALD40 and ALD10; (right) CNRM. The mean summer precipitation over the 93 climatological stations is given above each subfigure.

Average values over the 93 stations of other 30-yr mean summer statistics are also given: the RMSE and the mean absolute error (MAE). The statistics are calculated with daily values for each station separately. Both 40-km simulations ALR40 and ALD40 again perform similar. Overall, the biases are remarkably lower for ALARO-0 than for ALADIN. The bias over the 93 climatological stations between model simulations and observations is 0.25 mm day^{-1} for ALR40, 0.43 mm day^{-1} for ALD40, -0.06 mm day^{-1} for CNRM, 0.33 mm day^{-1} for ALR10, 1.06 mm day^{-1} for ALD10, and 0.06 mm day^{-1} for ALR04. The error statistics for all three ALARO-0 simulations show a similar improvement, suggesting a multiscale performance of ALARO-0. However, one should also keep in mind that error statistics are not entirely fair when validating models with different spatial resolution. Small displacements of precipitation maxima and minima in higher-resolution models

are highly penalized by error statistics because of the so-called double penalty effect (Soares et al. 2012a).

The aforementioned underestimation by CNRM is confirmed by the spatial distribution of its bias. Furthermore, the coastal precipitation is by all other models generally better simulated than the inland precipitation (Fig. 7). The larger and positive differences at the higher elevations can partly be assigned to higher uncertainties in the measurements of the observations due to rain gauge undercatchment (Buonomo et al. 2007). However, this overestimation, which is pronounced more strongly for ALD10 (Fig. 7), can also be attributed to the model or the driving ERA-40 data. All three ALARO-0 simulations (40-, 10-, and 4-km horizontal resolution) produce the lowest deviations from the observations, with a tendency to slightly overestimate (underestimate) in the southern (northern) part of the country. ALARO-0 values for RMSE and MAE lie in the same range as those for
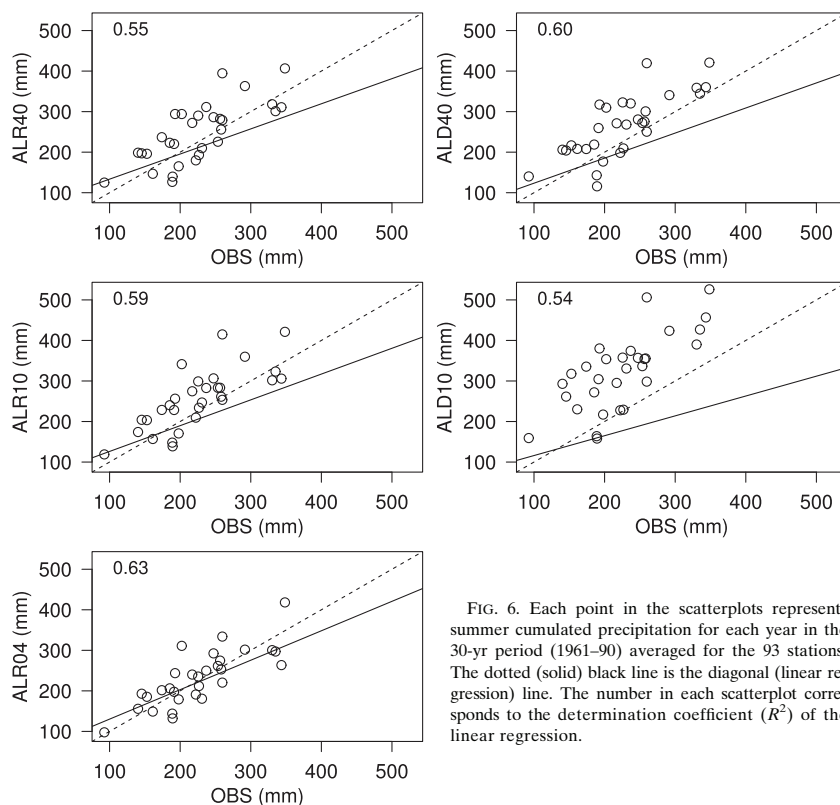
FIG. 6. Each point in the scatterplots represents summer cumulated precipitation for each year in the 30-yr period (1961–90) averaged for the 93 stations. The dotted (solid) black line is the diagonal (linear regression) line. The number in each scatterplot corresponds to the determination coefficient (R^2) of the linear regression.

ALADIN, indicating that the low mean biases of ALARO-0 are possible owing to cancellation effects arising from the bias computation. Nevertheless, the overall errors of the ALARO-0 simulations are still smaller than those of ALD10.

To get an understanding of the trend of frequency and intensity of extreme precipitation, density curves and frequency and quantile distributions of all six simulations have been created (Figs. 8–10). The densities in Fig. 8 have been calculated with the square root of the daily precipitation since the majority of the precipitation rates are less than $10 \, mm \, day^{-1}$. All models tend to overestimate the amount of "drizzle" and low precipitation (i.e., $<1 \, mm \, day^{-1}$). In the $1–2 \, mm \, day^{-1}$ range, both ALADIN simulations as well as CNRM overestimate the observed density almost by 2 times, while ALARO-0 starts to approach closely the observed density (Fig. 8, center). The latter continues to do this up to the right-end tail of the observed density curve (Fig. 8,

right). Perkins et al. (2007) use probability density functions (PDFs) for the evaluation of simulated daily precipitation over Australia from 14 different climate models. Similarly to the density curves of ALADIN and CNRM, the PDFs in Perkins et al. show for all models an overestimation of "drizzle," with most models over-estimating the observed density of rainfall in the $1–2 \, mm \, day^{-1}$ range by 2–3 times.

The relative frequencies, shown in Fig. 9, are again calculated for daily precipitation amounts of the observations and model data, which are binned into bins of $1 \, mm \, day^{-1}$. For the low precipitation rates all models manage to reproduce the observed frequencies relatively well. Once the 0.95th quantile of the observations (indicated by the vertical black line) is exceeded, CNRM shows an increasing departure from the observations with frequencies left shifted from the observations. ALARO-0 and ALADIN at 40-km horizontal resolution reveal again a similar result, while for the higher
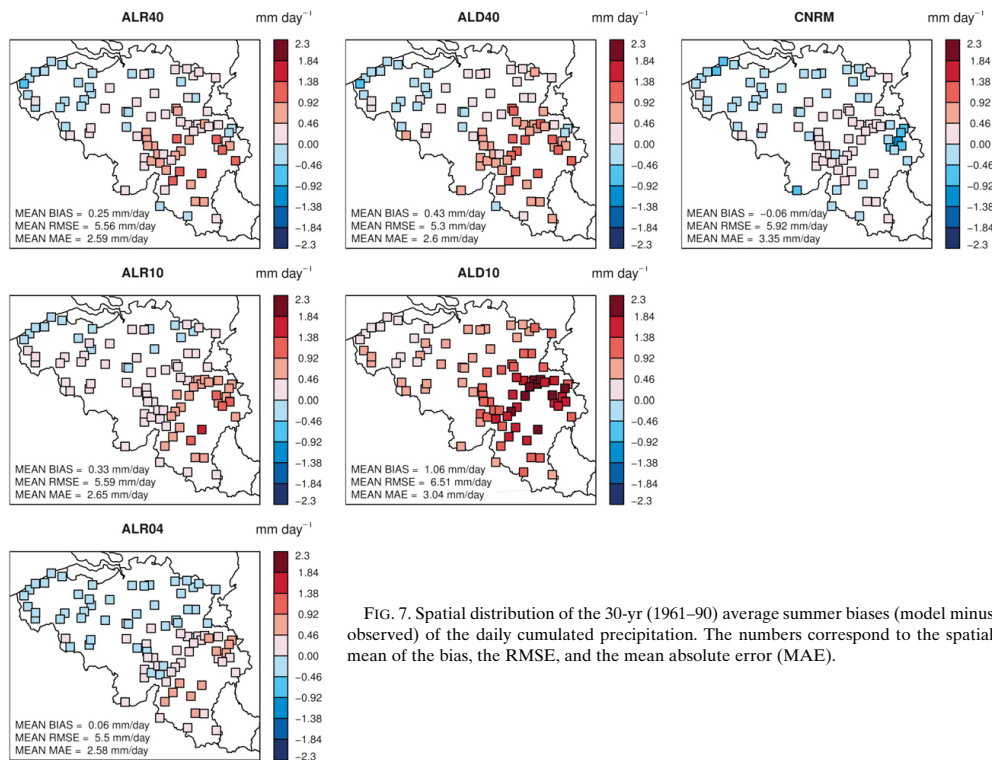
FIG. 7. Spatial distribution of the 30-yr (1961–90) average summer biases (model minus observed) of the daily cumulated precipitation. The numbers correspond to the spatial mean of the bias, the RMSE, and the mean absolute error (MAE).

10-km resolution a clear difference between both models is apparent. The small overestimation of ALD10 for the low precipitation rates persists and becomes larger for the higher rates. The model clearly rains too often, both with very small and very high quantities of rainfall. On the other hand, the frequencies of ALR04 and ALR10 nicely follow the observations, showing their ability to capture the occurrence of extreme and rare precipitation events, with values around 100 mm, quite well. As a measure for similarity between the observed and modeled frequencies, the PSS [Eq. (7)] are also given in Fig. 9. The overall PSS, as well as PSS for precipitation amounts below and above the 0.95th quantile of the observations, is higher for ALARO-0 than for ALADIN and CNRM.

The quantile distributions confirm the ability of ALR04, ALR10, and even ALR40 to reproduce extreme rainfall rates (Fig. 10). Only the highest 99.9 quantile (i.e., strongest events) is slightly overestimated by ALARO-0. It is evident that such events, which are situated in the very end of the distribution, might correspond to

outliers. Consistently with the frequency plots, the higher quantiles are over- and underestimated by ALD10 and CNRM, respectively.

Previous results can be qualified in the context of other regional downscaling studies; however, a direct comparison is difficult because of differences in study area and model design. Soares et al. (2012a) performed a dynamical downscaling of 20 years of the ECMWF Interim Re-Analysis (ERA-Interim) (1989–2008) for Portugal using the Weather Research and Forecasting (WRF) model. Two WRF high-resolution simulations (9 and 27 km) and ERA-Interim are compared with station observations. For summer precipitation, their results show a different frequency distribution for the 9- and 27-km simulation. The 9-km frequencies of summer precipitation follow well the observed frequencies and show a clear improvement compared to the driving reanalysis. Our results show a coherent performance of the ALARO-0 model across all resolutions and the good model performances as displayed in Figs. 8–10 can be practically attributed to the quality of the physics
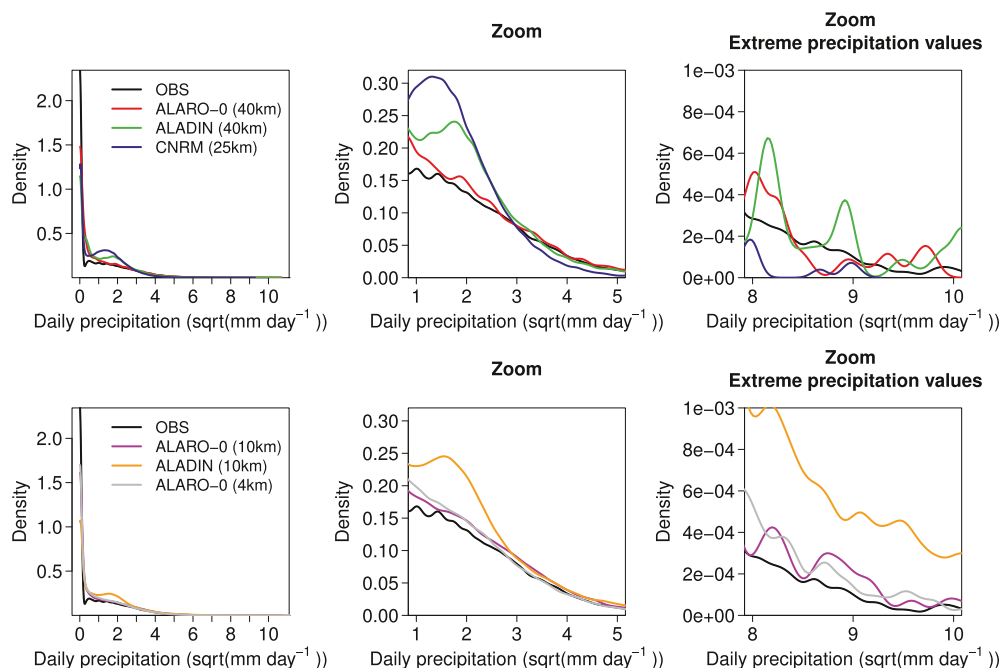
FIG. 8. Density curves of (top) observations, ALR40, ALD40, and CNRM and (bottom) observations, ALR10, ALD10, and ALR04. Densities are computed with the 30-yr (1961–90) daily cumulated summer precipitation given for each station separately. The x axes represent the square root of the daily precipitation since the majority of the precipitation rates are less than 10 mm day^{-1}.

parameterizations unrelated to the increase of the resolution. Finally, the persistent positive biases of the ALADIN model ALD10 are in accordance with other studies where recent past (1961–90) ALADIN simulations at 10-km horizontal resolution, driven by ERA-40 data, are validated against gridded observations (see Csíma and Horányi 2008; Skalák et al. 2008). According to Skalák et al. (2008), these positive (summer) precipitation biases can be linked with the tendency of the model "to precipitate more often than in the station observations."

3) EXTREME VALUE ANALYSIS

The extreme value analysis has been performed for each station separately, using the 30-yr daily summer data. The use of a generalized Pareto distribution as a model for threshold excesses assumes independent excesses (Coles 2001). In practice this is rarely the case. Exceedances over a certain threshold often occur in clusters. To account for these clusters of POT events, the data have been declustered by selecting the maximum value within each cluster. The independence of two clusters of POT events is determined by a combination

of the threshold and the separation time between both clusters. However, the choice of a suitable threshold and separation time is relatively arbitrary. The threshold has to be high enough in order to ensure *extreme* events and to avoid dependency between the events, but a threshold that is too high prevents statistical significance owing to a loss of information (Kyselý and Beranová 2009; Heikkilä et al. 2011). Similar to the study of Heikkilä et al., the threshold has been defined for each station separately as the 0.95th quantile of daily summer precipitation so that spatial differences in the precipitation amount (see Fig. 5) are taken into account.

The results obtained by using cluster maxima defined with different separation times (e.g., 1, 2, or 4 days) do not differ much from the results when the original non-declustered data have been used (not shown). Hence, in accordance with another study on extreme precipitation of Kyselý and Beranová (2009), two POT events are considered to be independent when the minimum separation time between both events is one day.

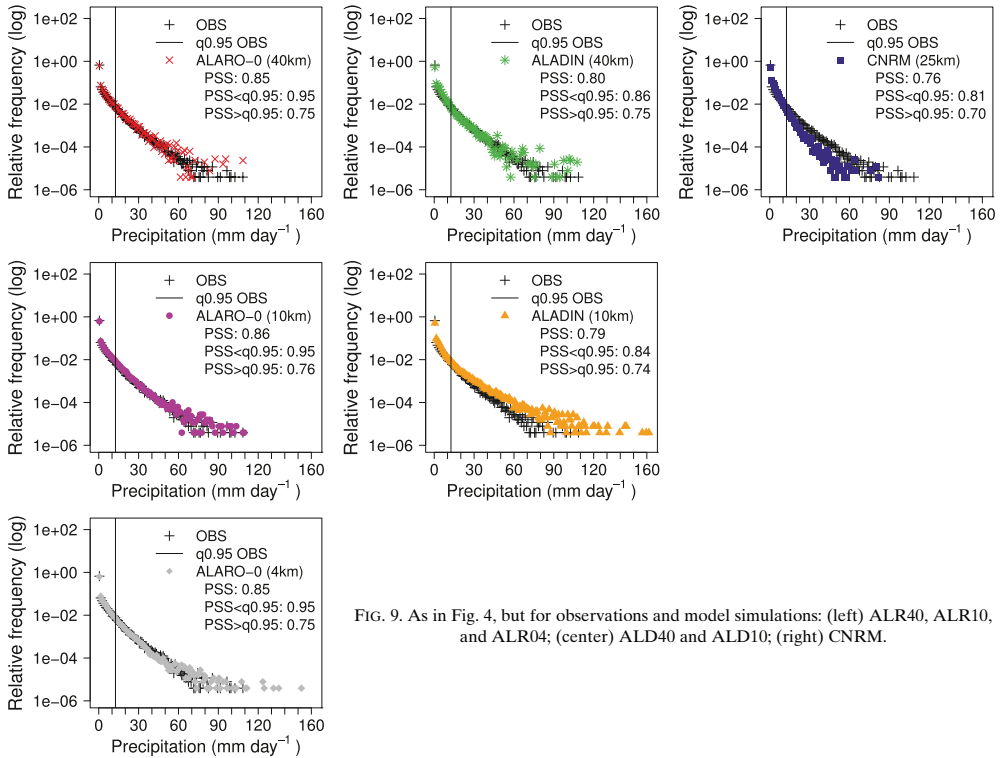To investigate if the underlying probability distribution of the (declustered) peak-over-threshold

FIG. 9. As in Fig. 4, but for observations and model simulations: (left) ALR40, ALR10, and ALR04; (center) ALD40 and ALD10; (right) CNRM.

events of the observations and models significantly differs, a Kolmogorov–Smirnov test has been applied. The Kolmogorov–Smirnov test statistic is defined as the maximum absolute difference between two distribution functions:

$$D_{n1,n2} = \max|F_{n1}(x) - F_{n2}(x)|, \qquad (8)$$

where $F_{n1}(x)$ and $F_{n2}(x)$ are the empirical distribution functions of the observations and the model, respectively, and n_i refers to the number of samples. The null hypothesis (H_0) that the distribution of the observed POT events equals the distribution of the modeled POT events is rejected at significance level $\alpha = 0.05$ if

$$K = \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} D_{n1,n2} > K_\alpha, \qquad (9)$$

where K_α is the critical α level of the Kolmogorov distribution:

$$Pr(K \leq K_\alpha) = 1 - \alpha. \qquad (10)$$

Figure 11 shows for each station the K statistic of the observations and models. In general, the K values for the ALARO-0 model at all three spatial resolutions are much smaller than ALADIN and CNRM. H_0 is accepted at the 95% level at 35 and 16 stations for ALD40 and ALD10, respectively. For ALARO-0 at 40, 10, and 4 km, H_0 is accepted at 46, 47, and 38 locations, respectively. Compared to ALD10, there are for ALARO-0 more stations at the high altitudes for which the distribution of the POT events equals the observed distribution of the POT events. This indicates that an increase in resolution does not necessarily contribute to a better representation of orographic precipitations. In the case of CNRM, H_0 is rejected for all stations. Thus, consistent with the results from the frequency and quantile distributions, the Kolmogorov–Smirnov test confirms that the ALARO-0 simulations yield more reliable statistics of the extreme events.

The GPD equation [Eq. (2)] is then fitted through the selected cluster maxima of the observations and the six model simulations ALR40, ALD40, CNRM, ALR10, ALD10, and ALR04. The 5- and 20-yr return levels of
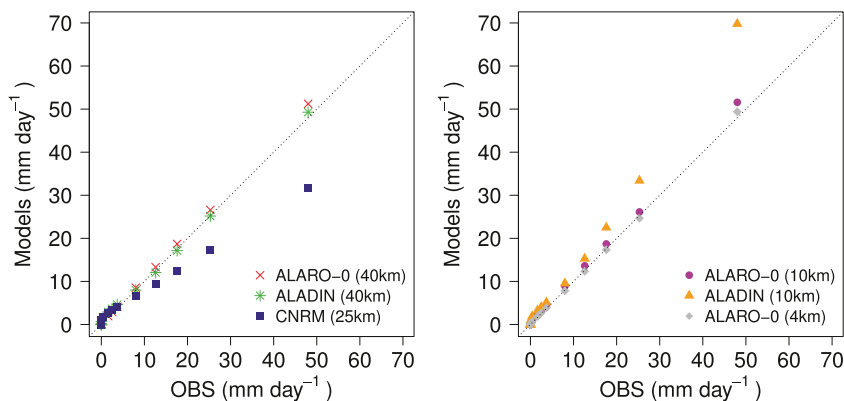
FIG. 10. Quantiles (2.5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95, 97.5, 99, and 99.9) of observations vs (left) the ALR40, ALD40, and CNRM models and (right) ALR10, ALD10 and ALR04 models. Quantiles are computed with the 30-yr (1961–90) daily cumulated summer precipitation given for each station separately.

the POT models for the observations and six simulations are shown in Figs. 12 and 13. The return levels x_m are calculated by Eq. (6) using the declustered data with 1-day separation time and a threshold u, defined as the 0.95th quantile. Since the return levels x_m are calculated on an annual basis, the value for m equals 92

observations, corresponding to the number of summer days within one year of the study period. The return levels for both return periods are generally larger at the higher elevations. The 95% confidence levels of the observed return levels are also indicated. It appears that for most stations the return levels of ALARO-0 lie
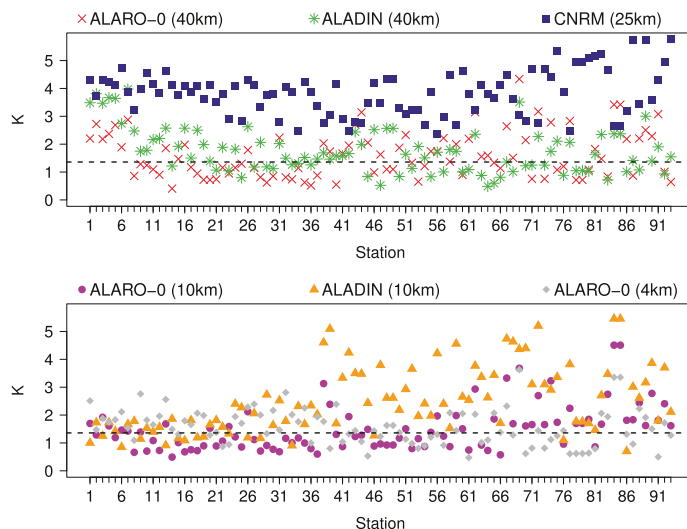

FIG. 11. The K statistic from a Kolmogorov–Smirnov test [Eq. (9)]. The 93 stations (abscissa) are shown by ascending altitude (from left to right). The test is performed on the POT events of the observations vs (top) ALR40, ALD40, and CNRM and (bottom) ALR10, ALD10, and ALR04 model simulations. The horizontal dotted line represents the critical K level with significance $\alpha = 0.05$.
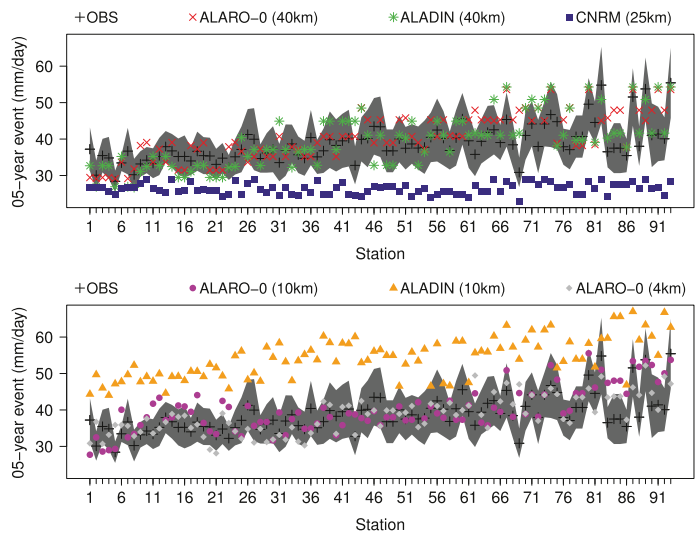
FIG. 12. The 5-yr return levels of the POT models for the observations and model simulations: (top) ALR40, ALD40, and CNRM and (bottom) ALR10, ALD10, and ALR04. The 93 stations (abscissa) are shown by ascending altitude (from left to right), and the shaded area represents the 95% confidence interval of observed return levels.

within the 95% confidence range of the observed return levels. In contrast to ALARO-0, ALD10 and CNRM are not able to produce the observed 5- and 20-yr return events. Their estimated return levels lie for a great

number of stations outside the observed confidence interval.

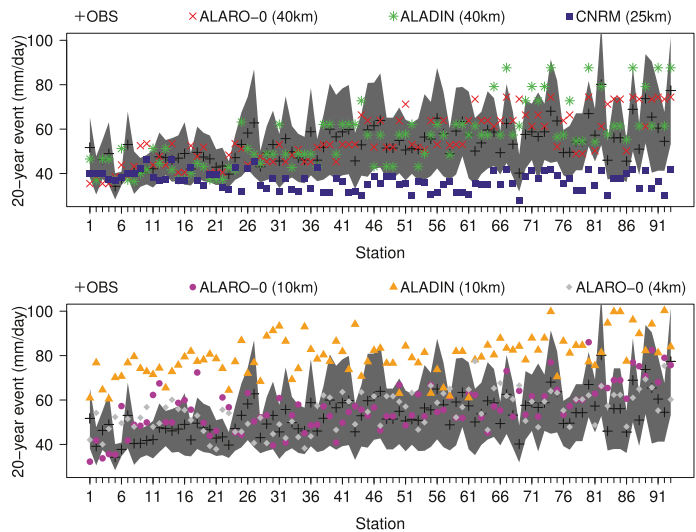In line with what Hamdi et al. (2012) found for summer maximum temperature, previous results from the

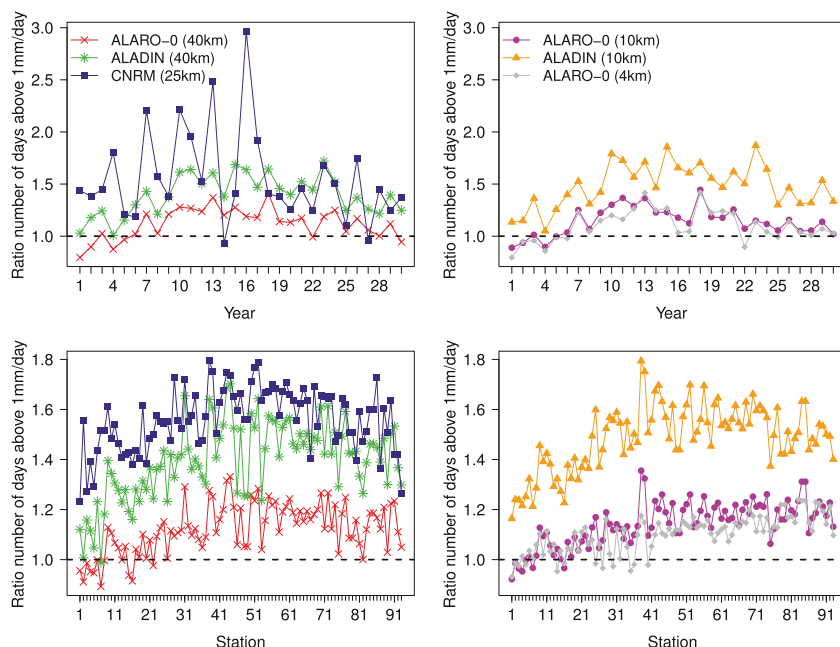FIG. 13. As in Fig. 12, but for 20-yr return levels.

FIG. 14. (top) Spatial mean of ratio of number of days above 1 mm day^{-1} (i.e., WD) in models to observations. (bottom) Temporal mean of ratio of number of days above 1 mm day^{-1} (i.e., WD) in models to observations. Station sequence as in Fig. 12.

extreme value analysis show for ALARO-0 at the high resolutions of 4 and 10 km, as well as at 40-km horizontal resolution, a clear improvement in simulating extreme summer precipitation. Extreme events are also often investigated by means of climate indices (e.g., Herrera et al. 2010; Domínguez et al. 2013; Dulière et al. 2011; Soares et al. 2012b). To complete the extreme value analysis, two main precipitation indices have been calculated: the number of wet days and the number of very heavy precipitation days. Both indices are explained below and are calculated for each year (i.e., summer season) and each climatological station.

4) NUMBER OF WET DAYS

The number of wet days (WD) for the observations and models are defined as the annual count of days when precipitation is >1 mm. Figure 14 shows the ratio of WD in models to observations. As the model values represent a whole grid box, we could assume that the models, and especially the lower resolution models, will poorly reproduce the indices at the station points. However, the low-resolution ALR40 model (left) reproduces

relatively well the observed WD. On the other hand, ALADIN and CNRM show an overestimation for WD. This can be explained by the fact that precipitation may occur more systematically at the model grid box level, which gives rise to a WD even when no precipitation has been observed at the station location. Compared to ALADIN and CNRM, the ALARO-0 model (at 4-, 10-, and 40-km horizontal resolution) is able to better reproduce the number of wet days.

5) NUMBER OF VERY HEAVY PRECIPITATION DAYS

The number of very heavy precipitation days is derived by an annual counting of days with precipitation rates >20 mm. The temporal and spatial means of the number of very heavy precipitation days are consistent with the results from foregoing extreme value analysis. Overall, ALR04, ALR10, and ALR40 can reproduce the number of days with precipitation >20 mm day^{-1} very well (Fig. 15). ALR04 and ALR10 have the highest correlations, and for three out of the 93 stations ALR10 predicts exactly the same number of days with heavy precipitation rates as have been observed.
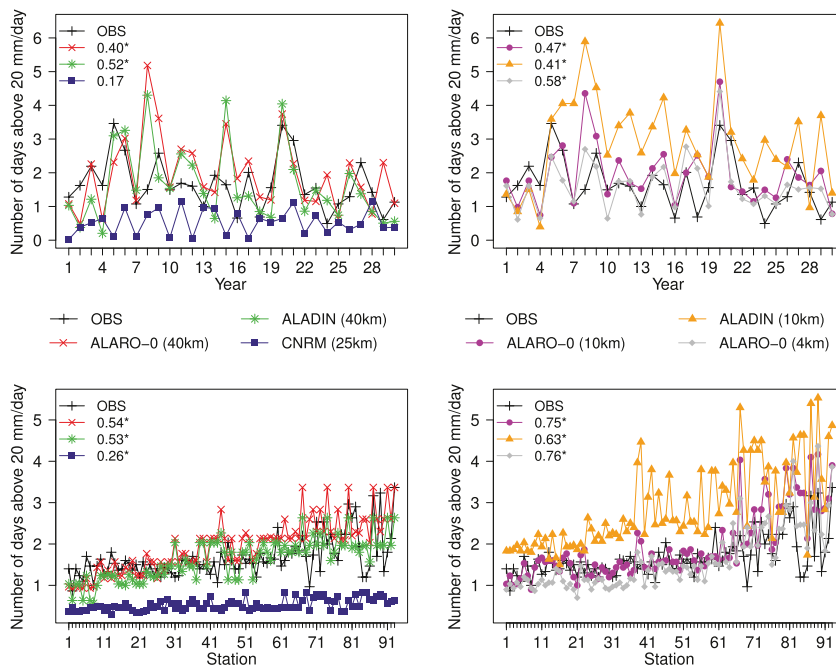
FIG. 15. (top) Spatial mean of number of days above $20\,\mathrm{mm\,day}^{-1}$. (bottom) Temporal mean of number of days above $20\,\mathrm{mm\,day}^{-1}$. The numbers correspond to Pearson correlation coefficients. Significant correlation coefficients at significance level 0.05 based on t statistics are indicated with an asterisk. Station sequence as in Fig. 12.

5. Conclusions

Extreme value analysis, using the peak-over-threshold method and generalized Pareto distribution, was performed in order to explore the relative importance of resolution versus parameterization formulation on the simulation of extreme daily summer precipitation. The results show that dynamical downscaling of the ERA-40 reanalysis using the ALARO-0 model adds value to the prediction of extreme daily summer precipitation when compared to the ERA-40 results. Hence, running a limited area model with the adapted parameterization, which was originally motivated to perform in the convection-permitting resolutions, statistically outperforms the global data in the output of extreme precipitation events of the ERA-40 reanalysis. The main strength of these tests is that, by the choice of the setup, we are considering the pure effect of the downscaling, without being obliterated by issues such as spectral nudging. Moreover, the model regenerates the precipitation instead of letting it evolve from its initial state. The regional nature keeps the computing cost within reach of

a typical small center, like the RMI, while reproducing the correct statistics of the extreme precipitation events consistently with the large-scale forcing imposed by the initial conditions and lateral boundaries. Furthermore, it should be stressed that the present model version has been developed and tuned in a context of NWP, is used as a 12-member component of the Grand Limited Area Model Ensemble Prediction System (GLAMEPS), and has been taken as such to downscale ERA-40 data. This can be seen as an extra indirect validation of the NWP applications running ALARO-0, in the sense that the model has a more correct climatology of convective rain. It is clear that there are several components, such as the physics–dynamics interaction, the interaction between model physics, and the numerics, that may influence the climatology of the precipitation. However, it is difficult to isolate the importance of these components, and it is beyond the scope of this study to address the relative impact of the different parameterization updates within ALARO-0. It should be kept in mind, though, that all of these factors play a crucial role in the model performance at gray-zone resolutions.

ALARO-0 simulations at 40-, 10-, and 4-km horizontal resolution with a new parameterization scheme of deep convection and microphysics and 40- and 10-km horizontal resolution output from the ALADIN model with an old parameterization scheme were compared with respect to station observation data. We find for ALARO-0 at high spatial resolutions of 10 and 4 km an improvement in the spatial distribution of summer precipitation, such that the distinct local maximum at the highest elevations is well resolved by the model, a feature strongly overestimated by the ALADIN model at 10-km resolution. Furthermore, the results from the extreme value analysis suggest that the new parameterization scheme of ALARO-0 contributes to the improvement in the modeling of extreme precipitation events at varying horizontal resolutions, rather than the increase in spatial resolution. Thus, the nature of the parameterization is more important than the resolution, which confirms the previous findings of Lynn et al. (2010) and Hamdi et al. (2012). As an outlook, the ALARO-0 model will be used to compute IPCC scenarios.

REFERENCES

ALADIN International Team, 1997: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. *WMO Bull.,* **46,** 317–324.

Boberg, F., P. Berg, P. Thejll, W. J. Gutowski, and J. H. Christensen, 2010: Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models. *Climate Dyn.,* **35,** 1509–1520, doi:10.1007/s00382-009-0683-8.

Bougeault, P., 1985: A simple parameterization of the large-scale effects of cumulus convection. *Mon. Wea. Rev.,* **113,** 2108–2121.

Bubnová, R., G. Hello, P. Bénard, and J.-F. Geleyn, 1995: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/ALADIN NWP system. *Mon. Wea. Rev.,* **123,** 515–535.

Buonomo, E., R. Jones, I. C. Huntingford, and J. Hannaford, 2007: On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quart. J. Roy. Meteor. Soc.,* **133,** 65–81, doi:10.1002/qj.13.

Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala, 2009: Evaluation of a WRF dynamical downscaling simulation over California. *Climatic Change,* **95,** 499–521, doi:10.1007/s10584-009-9583-5.

Castro, C. L., R. A. S. Pielke, and G. Leoncini, 2005: Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *J. Geophys. Res.,* **110,** D05108, doi:10.1029/2004JD004721.

Catry, B., J.-F. Geleyn, M. Tudor, P. Bénard, and A. Trojáková, 2007: Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus,* **59A,** 71–79, doi:10.1111/j.1600-0870.2006.00212.x.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values.* Springer, 210 pp.

Csima, G., and A. Horányi, 2008: Validation of the ALADIN-Climate regional climate model at the Hungarian Meteorological Service. *Idojaras,* **112,** 155–177.

Davies, H. C., 1976: A lateral boundary formulation for multilevel prediction models. *Quart. J. Roy. Meteor. Soc.,* **102,** 405–418.

de Elía, R., R. Laprise, and D. Bertrand, 2002: Forecasting skill limits of nested, limited-area models: A perfect-model approach. *Mon. Wea. Rev.,* **130,** 2006–2023.

Denis, B., R. Laprise, D. Caya, and J. Côté, 2002: Downscaling ability of one-way nested regional climate models: The Big-Brother Experiment. *Climate Dyn.,* **18,** 627–646, doi:10.1007/s00382-001-0201-0.

Déqué, M., and S. Somot, 2008: Analysis of heavy precipitation for France using high resolution ALADIN RCM simulations. *Idojaras,* **112,** 179–190.

Domínguez, M., and Coauthors, 2013: Present climate precipitation and temperature extremes over Spain from a set of high resolution RCMs. *Climate Res.,* in press.

Dulière, V., Y. Zhang, and E. P. J. Salathé, 2011: Extreme precipitation and temperature over the U.S. Pacific Northwest: A comparison between observations, reanalysis data, and regional models. *J. Climate,* **24,** 1950–1964.

Durman, C. F., J. M. Gregory, D. C. Hassell, R. G. Jones, and J. M. Murphy, 2001: A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates. *Quart. J. Roy. Meteor. Soc.,* **127,** 1005–1015, doi:10.1002/qj.49712757316.

Geleyn, J.-F., B. Catry, Y. Bouteloup, and R. Brožková, 2008: A statistical approach for sedimentation inside a microphysical precipitation scheme. *Tellus,* **60A,** 649–662, doi:10.1111/j.1600-0870.2008.00323.x.

Gerard, L., 2007: An integrated package for subgrid convection, clouds and precipitation compatible with the meso-gamma scales. *Quart. J. Roy. Meteor. Soc.,* **133,** 711–730, doi:10.1002/qj.58.

——, and J.-F. Geleyn, 2005: Evolution of a subgrid deep convection parameterization in a limited area model with increasing resolution. *Quart. J. Roy. Meteor. Soc.,* **131,** 2293–2312, doi:10.1256/qj.04.72.

——, J.-M. Piriou, R. Brožková, J.-F. Geleyn, and D. Banciu, 2009: Cloud and precipitation parameterization in a meso-gamma-scale operational weather prediction model. *Mon. Wea. Rev.,* **137,** 3960–3977.

Giorgi, F., and L. O. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *J. Geophys. Res.,* **104,** 6335–6352.

——, X. Bi, and J. S. Pal, 2004: Mean, interannual variability and trends in a regional climate change experiment over Europe. I. Present-day climate (1961–1990). *Climate Dyn.,* **22,** 733–756, doi:10.1007/s00382-004-0409-x.

Hamdi, R., H. Van de Vyver, and P. Termonia, 2012: New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: Application for summer extreme temperature over Belgium. *Int. J. Climatol.,* **32,** 2051–2065, doi:10.1002/joc.2409.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.,* **113,** D20119, doi:10.1029/2008JD010201.

Heikkilä, U., A. Sandvik, and A. Sorteberg, 2011: Dynamical downscaling of ERA-40 in complex terrain using WRF regional climate model. *Climate Dyn.,* **37,** 1551–1564, doi:10.1007/s00382-010-0928-6.

Herrera, S., L. Fita, J. Fernández, and J. M. Gutiérrez, 2010: Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. *J. Geophys. Res.,* **115,** D21117, doi:10.1029/2010JD013936.

Hofstra, N., M. New, and C. McSweeney, 2010: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Climate Dyn.,* **35,** 841–858, doi:10.1007/s00382-009-0698-1.

Jacob, D., and Coauthors, 2007: An inter-comparison of regional climate models for Europe: Model performance in present-day climate. *Climatic Change,* **81,** 31–52, doi:10.1007/s10584-006-9213-4.

Jones, P. D., and P. A. Reid, 2001: Assessing future changes in extreme precipitation over Britain using regional climate model integrations. *Int. J. Climatol.,* **21,** 1337–1356, doi:10.1002/joc.677.

Jones, R. G., J. M. Murphy, and M. Noguer, 1995: Simulation of climate change over europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. *Quart. J. Roy. Meteor. Soc.,* **121,** 1413–1449, doi:10.1002/qj.49712152610.

——, M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins, and J. F. B. Mitchell, 2004: Generating high resolution climate change scenarios using PRECIS. National Communications Support Unit Workbook, Met Office Hadley Centre, Exeter, United Kingdom, 32 pp. [Available online at http://www.uncclearn.org/sites/www.uncclearn.org/files/inventory/UNDP17.pdf.]

Kyselý, J., and R. Beranová, 2009: Climate-change effects on extreme precipitation in central Europe: Uncertainties of scenarios based on regional climate models. *Theor. Appl. Climatol.,* **95,** 361–374, doi:10.1007/s00704-008-0014-8.

Lucas-Picher, P., F. Boberg, J. Christensen, and P. Berg, 2013: Dynamical downscaling with reinitializations: A method to generate fine-scale climate datasets suitable for impact studies. *J. Hydrometeor.,* **14,** 1159–1174.

Lynn, B. H., and Coauthors, 2010: Testing GISS-MM5 physics configurations for use in regional impacts studies. *Climatic Change,* **99,** 567–587, doi:10.1007/s10584-009-9729-5.

Ma, L., T. Zhang, O. W. Frauenfeld, B. Ye, D. Yang, D. Yang, and D. Qin, 2009: Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *J. Geophys. Res.,* **114,** D09105, doi:10.1029/2008JD011178.

Mailhot, A., S. Lachance-Cloutier, G. Talbot, and A.-C. Favre, 2013: Regional estimates of intense rainfall based on the peak-over-threshold (POT) approach. *J. Hydrol.,* **476,** 188–199, doi:10.1016/j.jhydrol.2012.10.036.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.,* **25,** 693–712, doi:10.1002/joc.1181.

Nicolis, C., 2003: Dynamics of model error: Some generic features. *J. Atmos. Sci.,* **60,** 2208–2218.

——, 2004: Dynamics of model error: The role of unresolved scales revisited. *J. Atmos. Sci.,* **61,** 1740–1753.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.,* **117,** 536–549.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate,* **20,** 4356–4376.

Qian, J.-H., A. Seth, and S. Zebiak, 2003: Reinitialized versus continuous simulations for regional climate downscaling. *Mon. Wea. Rev.,* **131,** 2857–2874.

R Development Core Team, cited 2011: *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. [Available online at http://www.R-project.org/.]

Radnóti, G., 1995: Comments on "A spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations." *Mon. Wea. Rev.,* **123,** 3122–3123.

Radu, R., M. Déqué, and S. Somot, 2008: Spectral nudging in a spectral regional climate model. *Tellus,* **60A,** 898–910, doi:10.1111/j.1600-0870.2008.00341.x.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 589–662.

Ribatet, M., T. B. M. J. Ouarda, E. Sauquet, and J.-M. Gresillon, 2009: Modeling all exceedances above a threshold using an external dependence structure: Inferences on several flood characteristics. *Water Resour. Res.,* **45,** W03407, doi:10.1029/2007WR006322.

Ritter, B., and J.-F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.,* **120,** 303–325.

Roth, M., T. A. Buishand, G. Jongbloed, A. M. G. K. Tank, and J. H. van Zanten, 2012: A regional peaks-over-threshold model in a nonstationary climate. *Water Resour. Res.,* **48,** W11533, doi:10.1029/2012WR012214.

Skalák, P., P. Stepánek, and A. Farda, 2008: Validation of ALADIN-Climate/CZ for present climate (1961–1990) over the Czech Republic. *Idojaras,* **112,** 191–201.

Soares, P. M. M., R. M. Cardoso, P. M. A. Miranda, J. Medeiros, M. Belo-Pereira, and F. Espirito-Santo, 2012a: WRF high resolution dynamical downscaling of ERA-Interim for Portugal. *Climate Dyn.,* **39,** 2497–2522, doi:10.1007/s00382-012-1315-2.

——, ——, ——, P. Viterbo, and M. Belo-Pereira, 2012b: Assessment of the ENSEMBLES regional climate models in the representation of precipitation variability and extremes over Portugal. *J. Geophys. Res.,* **117,** D07114, doi:10.1029/2011JD016768.

Stephenson, A., 2002: evd: Extreme value distributions. *R News,* Vol. 2 (2), R Project, 31–32. [Available online at http://cran.r-project.org/doc/Rnews/Rnews_2002-2.pdf.]

Sylla, M. B., E. Coppola, L. Mariotti, F. Giorgi, P. M. Ruti, A. Dell'Aquila, and X. Bi, 2010: Multiyear simulation of the African climate using a regional climate model (RegCM3) with the high resolution ERA-interim reanalysis. *Climate Dyn.,* **35,** 231–247, doi:10.1007/s00382-009-0613-9.

Termonia, P., A. Deckmyn, and R. Hamdi, 2009: Study of the lateral boundary condition temporal resolution problem and a proposed solution by means of boundary error restarts. *Mon. Wea. Rev.,* **137,** 3551–3566.

——, F. Voitus, D. Degrauwe, S. Caluwaerts, and R. Hamdi, 2012: Application of Boyd's periodization and relaxation method in a spectral atmospheric limited-area model. Part I: Implementation and reproducibility tests. *Mon. Wea. Rev.,* **140,** 3137–3148.

Tustison, B., D. Harris, and E. Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *J. Geophys. Res.,* **106,** 11 775–11 784.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012, doi:10.1256/qj.04.176.

Váňa, F., P. Bénard, J.-F. Geleyn, A. Simon, and Y. Seity, 2008: Semi-Lagrangian advection scheme with controlled damping: An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Quart. J. Roy. Meteor. Soc.,* **134,** 523–537, doi:10.1002/qj.220.

Warner, T., R. Peterson, and R. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc.,* **78,** 2599–2617.

Zolina, O., A. Kapala, C. Simmer, and S. K. Gulev, 2004: Analysis of extreme precipitation over Europe from different reanalyses: A comparative assessment. *Global Planet. Change,* **44,** 129–161, doi:10.1016/j.gloplacha.2004.06.009.

B

SUPPLEMENTARY FIGURES TO CHAPTER 4

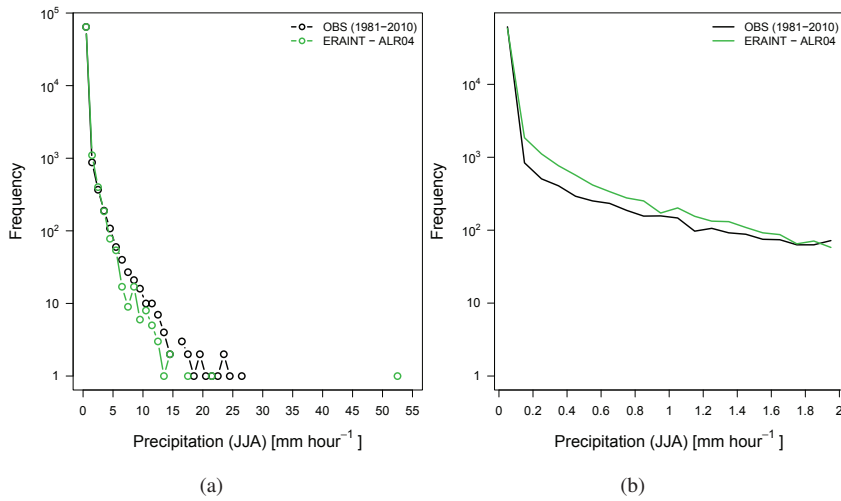**Figure B.1:** *Frequencies of (a) hourly summer precipitation, and (b) low values of hourly summer precipitation (i.e. [0-2] mm hour^{-1}) of observations and ERAINT-ALR04. Frequencies are computed with the 30-yr (1961-90) hourly summer precipitation given for the station of Uccle and its nearest model grid box. Frequencies are displayed on a logarithmic scale.*
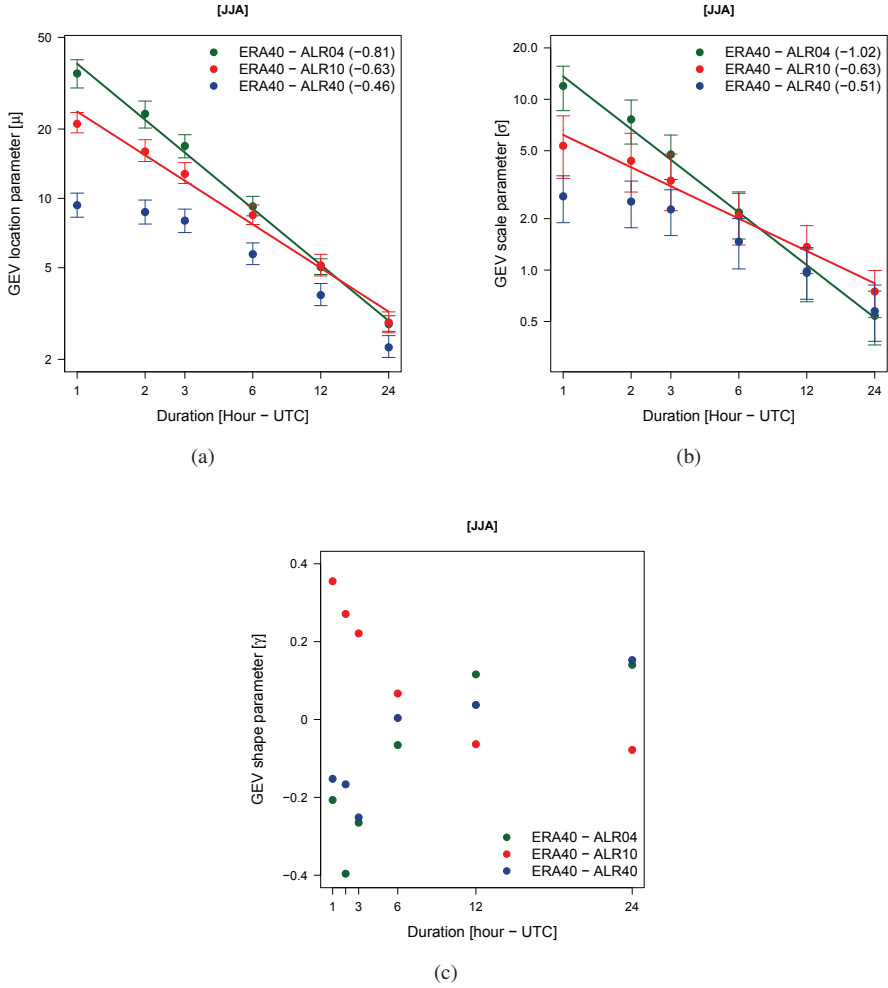
Figure B.2: *Estimates of (a) GEV location parameter (μ), (b) GEV scale parameter (σ), and (c) GEV shape parameter (γ). GEV estimates (colored dots) are obtained by the L-moments method for different durations d of observed (black) and modeled (green for ERA40-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40) summer annual maxima precipitation intensities for the upscaled 4-km and 10-km simulations, and 40-km simulation, across all 40-km grid boxes of the common 8 × 7 subdomain. The error bars show the 95% confidence interval around the location and scale parameter estimates based on 1000 parametric bootstrap iterations and solid lines correspond to the linear regression lines between the logarithm of GEV parameter estimates and the logarithm of the different durations d, with the values of the slope given between parentheses in the legend. GEV location and scale parameter estimates and durations are displayed on a logarithmic scale.*
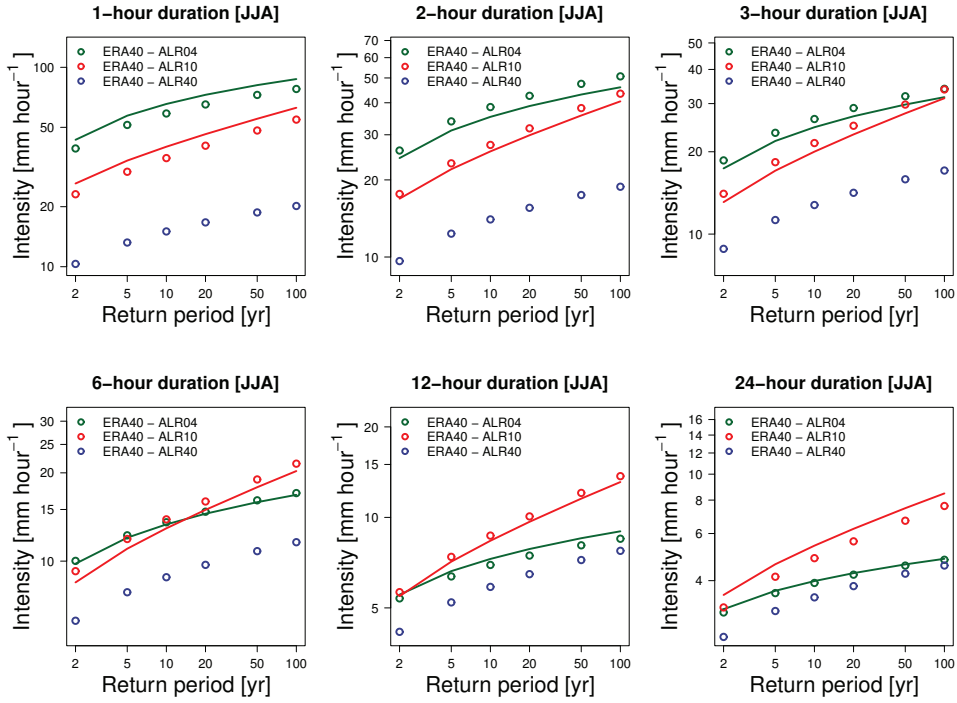
Figure B.3: *JJA IDF-relationships with the intensities (return levels) calculated with (i) [Eq. (4.8)] using the GEV-estimated parameters (circles), and with (ii) [Eq. (4.23)] using the location- and scale parameter derived from the power law given by [Eqs. (4.21) to (4.22)] (solid lines). For both intensity calculations (i) and (ii) one and the same mean value over all durations for the shape parameter (γ) has been used. Intensities are given as a function of return period (T) for observations (black) and models (green for ERA4O-ALR04, red for ERA40-ALR10, and blue for ERA40-ALR40), and plotted on log-log graphs. Intensities are computed for the upscaled 4-km and 10-km simulations, and 40-km simulation, across all 40-km grid boxes of the common 8 × 7 subdomain. The different panels represent different durations (1, 2, 3, 6, 12, and 24 h).*
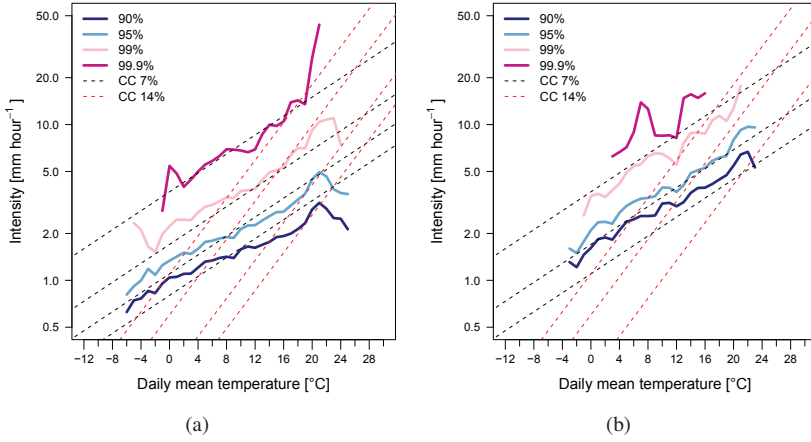
(a) (b)

Figure B.4: *Dependencies of different extreme percentiles (90th-99.9^{th}) of the distribution of modeled (ERAINT-ALR04) (a) hourly and (b) daily maximum of hourly precipitation on temperature, for a region of 5×5 (25) grid points surrounding the closest model grid point to Uccle. Exponential relations given by a 7% and a 14% increase per degree are given by the black and red dotted lines, respectively. Percentiles are displayed on a logarithmic y axis.*
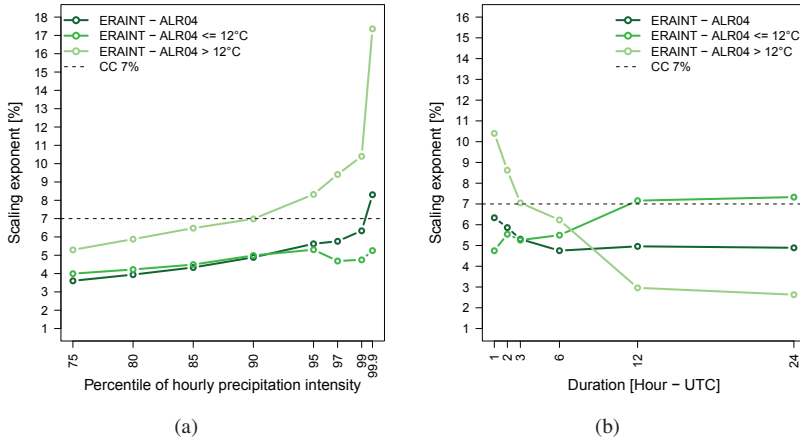

(a) (b)

Figure B.5: *(a) Variation of the modeled (ERAINT-ALR04, green) scaling exponent with (a) percentile of hourly precipitation, and (b) 99th percentile scaling exponent with precipitation duration. The scaling exponents are calculated for the whole temperature range up to $22°C$, as well as for temperatures below and equal to $12°C$ and for temperatures above $12°C$. The horizontal black dotted line corresponds to the theoretical 7% Clausius-Clapeyron (CC) increase per degree.*
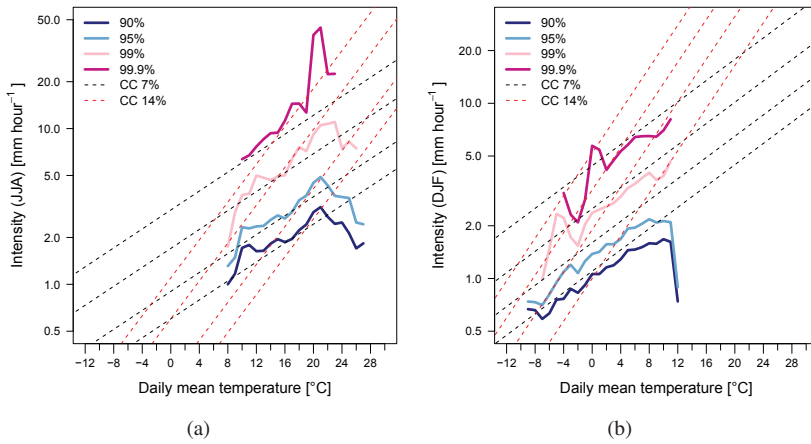
Figure B.6: *Dependency of hourly precipitation extremes on temperature computed from model data (ERAINT-ALR04) of a region of 5 × 5 (25) grid points surrounding the closest model grid point to Uccle of (a) the summer season JJA, and (b) the winter season DJF. Lines and axes are similar to Fig. 4.14.*

C
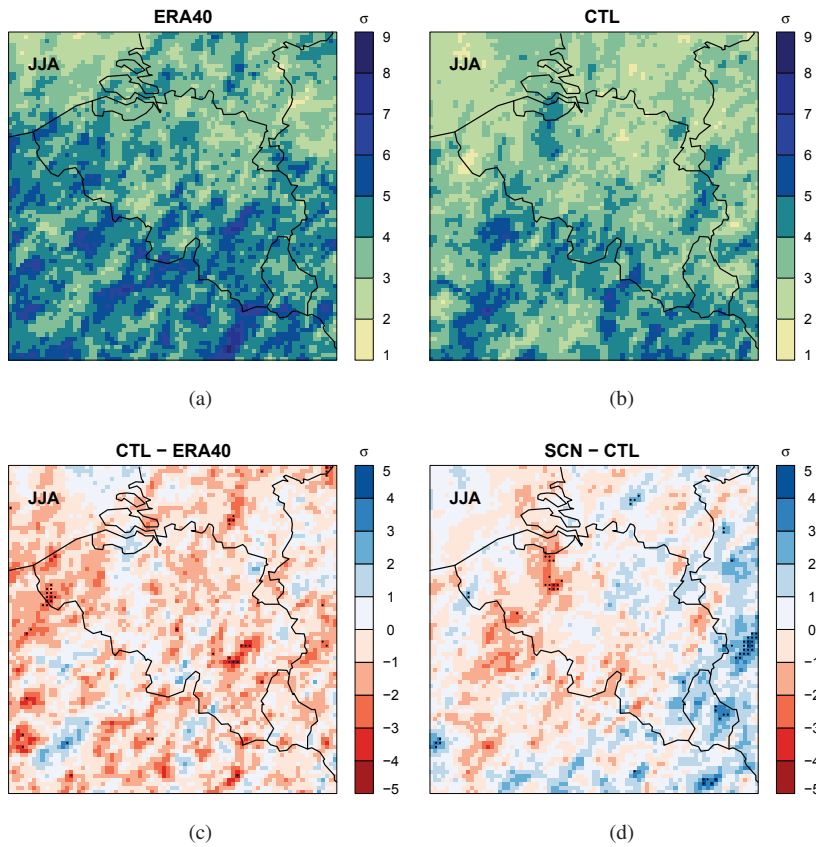
SUPPLEMENTARY FIGURES TO CHAPTER 5

Figure C.1: *Spatial distributions of the estimated scale parameter σ from the GEV fit of 1-hour precipitation in summer (JJA, 1961-1990 and 2071-2100) for (a) the ERA-40 driven simulation (ERA40), and (b) the control simulation (CTL), and (c-d) absolute differences between CTL and ERA40, and SCN and CTL, giving the bias and projected change, respectively. Dotted areas indicate regions where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and ERA40, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*

Figure C.2: *Spatial distributions of the estimated shape parameter γ from the GEV fit of 1-hour precipitation in summer (JJA, 1961-1990 and 2071-2100) for (a) the ERA-40 driven simulation (ERA40), and (b) the control simulation (CTL), and (c-d) absolute differences between CTL and ERA40, and SCN and CTL, giving the bias and projected change, respectively. Dotted areas indicate regions where the bias or change is statistically significant at the 5% significance level for which the 95% confidence intervals of CTL and ERA40, and SCN and CTL, as inferred from 1000 bootstrap samples, do not overlap.*
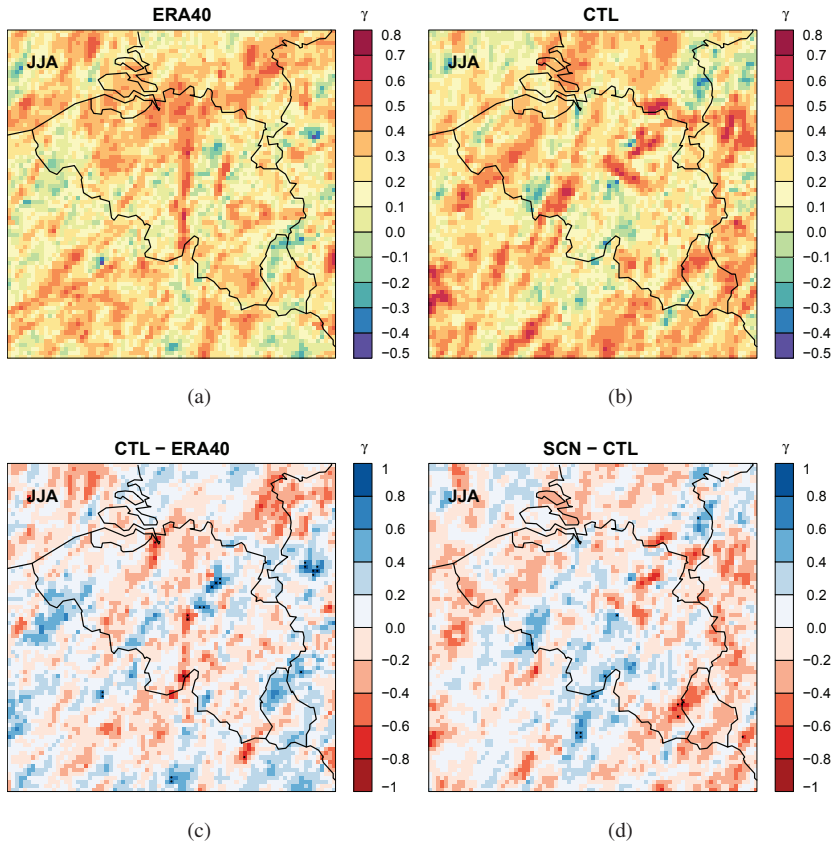
D
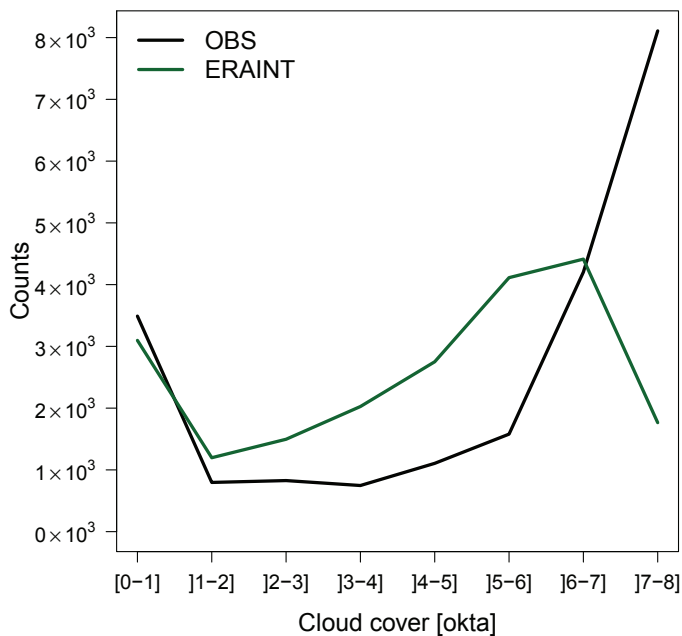
SUPPLEMENTARY FIGURES TO CHAPTER 6

Figure D.1: *Frequency distribution of cloudiness (okta) for the 29-yr DJF period 1981/82-2009/10. Frequencies are calculated for observed (OBS) values in Uccle and modeled (ERAINT) values for the closest model grid point to Uccle, by binning the hourly cloudiness values into bins of 1 okta.*

(a)

(b)

(c)
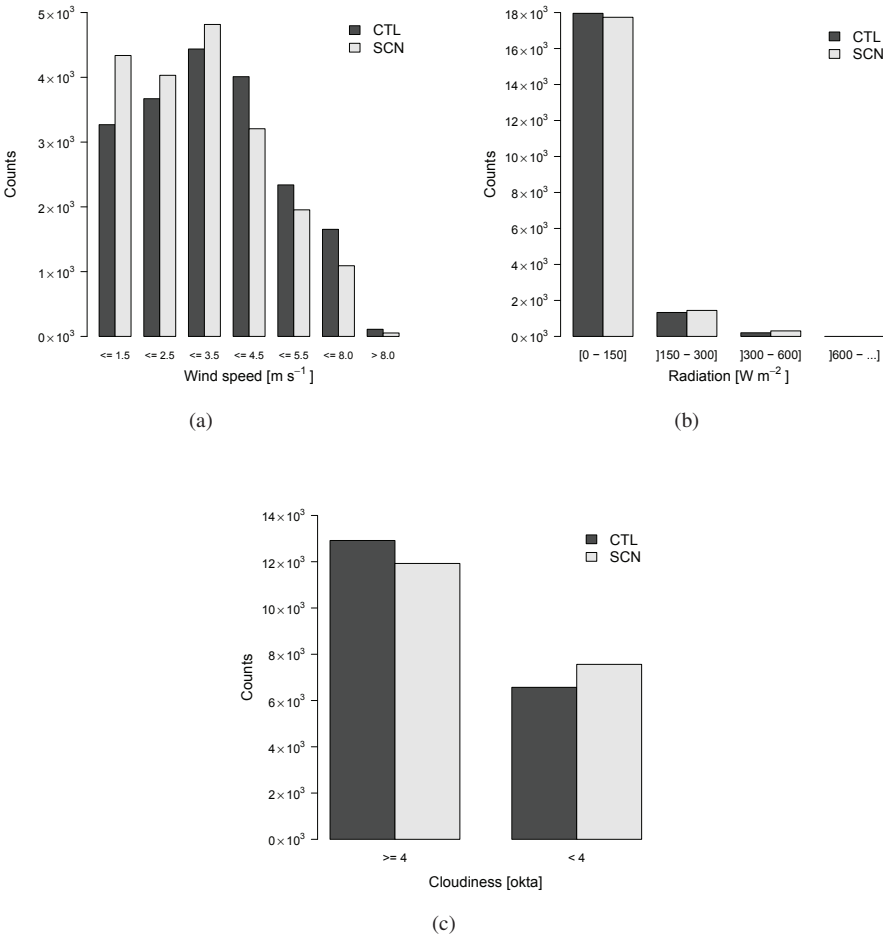
Figure D.2: *Frequency distribution of (a) wind speed at 10 m (m s^{-1}), (b) global solar radiation (W m^{-2}), and (c) cloudiness (okta) for the 9-yr DJF control (CTL) period 1990/91-1998/99 and scenario (SCN) period 2046/47-2054/55. Frequencies are calculated for CTL and SCN values for the closest model grid point to Uccle, according to the thresholds given by the stability scheme that is used to determine the Pasquill indices (Table 6.2).*

E
CURRICULUM VITAE

Rozemien De Troch

CONTACT
INFORMATION

Rozemien De Troch
Segherslaan 56/3
1081 Brussels
Belgium
rozemien.detroch@meteo.be

RESEARCH
INTERESTS

Regional Climate Modeling
Atmospheric Sciences
Numerical Weather Prediction
Extreme Value Analysis

EDUCATION

Ghent University, Ghent

Postgraduate Studies in Weather and Climate Modeling, June 2011

Free University of Brussels, Brussels

Master in Geography, profile Terrestric Ecosystems and Global Change, June 2010

- Topic: *The relationship between meteorological conditions and forest fires in Greece, Portugal and the Canary Islands: A comparative study*
- Supervisor: Prof. Dr. van Lipzig Nicole

RESEARCH
EXPERIENCE

Research Assistant   January 2012 to present
Royal Meteorological Institute of Belgium,
Research Department, in association with Ghent University
Supervisor: Dr. Piet Termonia

REFEREED
JOURNAL
PUBLICATIONS
AND BOOK
CHAPTERS

1. **De Troch, R.**, R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia, 2013: Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. *J. Climate*, 26, 8895-8915, doi:10.1175/JCLI-D-12-00844.1.

2. Hamdi, R., H. Van de Vyver, **R. De Troch**, and P. Termonia, 2014: Assessment of three dynamical urban climate downscaling methods: Brussels's future urban heat island under an A1B emission scenario. *Int. J. Climatol.*, 34, 978-999, doi:10.1002/joc.3734.

3. Hamdi, R., H. Van de Vyver, **R. De Troch**, P. Termonia, and A. Delcloo, 2013: Assessment of Three Dynamical Urban Climate Downscaling Methods. In book: *Air Pollution Modeling and its Application XXIII*, Publisher: Springer International Publishing, Editors: D. Steyn and R. Mathur, ISBN 978-3-319-04378-4, pp. 379-383.

4. Hamdi, R., O. Giot, **R. De Troch**, A. Deckmyn, and P. Termonia, 2015: Future climate of Brussels and Paris for the 2050s under the A1B scenario. *Urban Climate*, 12, 160-182, doi:10.1016/j.uclim.2015.03.003.

5. Giot, O., P. Termonia, D. Degrauwe, **R. De Troch**, S. Caluwaerts, G. Smet, J. Berckmans, A. Deckmyn, L. De Cruz, P. De Meutter, A. Duerinckx, L. Gerard, R. Hamdi, J. Van den Bergh, M. Van Ginderachter, and B. Van Schaeybroeck, 2016. Validation of the ALARO-0 model within the EURO-CORDEX framework. *Geosci. Model Dev.*, 9, 1143-1152, doi:10.5194/gmd-9-1143-2016.

SUBMITTED JOURNAL PUBLICATIONS	1. Tabari, H., **R. De Troch**, O. Giot, R. Hamdi, P. Termonia, S. Saeed, S., E. Brisson, N. van Lipzig, and P. Willems, 2016: Local impact analysis of climate change on precipitation extremes: Are high-resolution climate models needed for realistic simulations? *Hydrol. Earth Syst. Sci. Discuss.*, in review, doi:10.5194/hess-2016-106.
PAPERS IN PREPARATION	1. **De Troch, R.**, P. Termonia, and R. Hamdi, 2015: Practical assessment of climate change impact on winter smog episodes: a case study for Brussels. To be submitted to *J. Appl. Meteor. Climatol.*
OTHER PUBLICATIONS	1. **De Troch, R.**, O. Giot, R. Hamdi, S. Saeed, H. Tabari, M. T. Taye, P. Termonia, N. van Lipzig, and P. Willems, 2014: Overview of a few regional climate models and climate scenarios for Belgium. *Scientific and technical publication*, 65, Publisher: Royal Meteorological Institute of Belgium, Editor: D. Gellens, pp. 1-32.
	2. Vannitsem, S., **R. De Troch**, and R. Hamdi, 2015: Hoofdstuk 3. Onderzoek over klimaatvoorspelling en klimaatmodellering. In: *Oog voor klimaat - 2015*, Publisher: Koninklijk Meteorologisch Instituut van België, Editor: D. Gellens, ISBN 02008/0224/049, pp. 52-70.
	3. Brouwers, J., B. Peeters, N. van Lipzig, H. Wouters, J. Beullens, M. Demuzere, P. Willems, **R. De Troch**, P. Termonia, K. De Ridder, B. Maiheu, K. Cauwenberghs, M. Craninx, W. Defloor, W. Maetens, and T. Vansteenkiste, 2015: MIRA Climate Report 2015: about observed and future climate changes in Flanders and Belgium. Publisher: Flemish Environment Agency - Vlaamse Milieumaatschappij (VMM), Editor: Flemish Environment Agency - Vlaamse Milieumaatschappij (VMM), ISBN: 9789491385469, pp. 1-152.
ORAL PRESENTATIONS	1. **De Troch, R.**, R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia: Multiscale behavior of ALARO-0 for extreme summer precipitation over Belgium. Seminar at Royal Meteorological Institute of Belgium, Brussels, Belgium, October 2012.
	2. **De Troch, R.**, R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia: Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. European Meteorological Society Annual Meeting, Reading, UK, September 2013.
	3. **De Troch, R.**, R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia: Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. Working Group Meeting 1+2 on Parameterization and Downscaling of COST Action ES0905: Basic Concepts for Convection Parameterization in Weather Forecast and Climate Models, Milan, Italy, January 2014.
	4. **De Troch, R.**, H. Van de Vyver, O. Giot, R. Hamdi, and P. Termonia: Assessment of subdaily precipitation from the ALARO-0 model at different spatial resolutions. European Geosciences Union General Assembly, Vienna, Austria, April 2014.
	5. De Meutter, P., S. Caluwaerts, A. Deckmyn, A. Duerinckx, D. Degrauwe, G. Smet, J. Berckmans, L. De Cruz, L. Gerard, M. Van Ginderachter, O. Giot, P. Termonia, R. Hamdi, and **R. De Troch**: ALARO-0 experience in Belgium. ALARO-1 working days, Vienna, Austria, May 2014.

6. **De Troch, R.**, H. Van de Vyver, O. Giot, R. Hamdi, and P. Termonia: Assessment of subdaily precipitation from the ALARO-0 model at different spatial resolutions. MeteoClim 2014, Antwerp, Belgium, June 2014.

7. De Meutter, P., A. Deckmyn, A. Duerinckx, D. Degrauwe, G. Smet, J. Berckmans, L. De Cruz, L. Gerard, M. Van Ginderachter, O. Giot, P. Termonia, R. Hamdi, **R. De Troch**, and S. Caluwaerts: ALARO Forecast Experience in Belgium. ALADIN Forecasters meeting, Ankara, Turkey, September 2014.

8. **De Troch, R.**: High resolution regional climate modelling. Part 1. Evaluation and projection for Belgium. Seminar at Royal Meteorological Institute of Belgium, Brussels, Belgium, November 2014.

9. Delcloo, A., **R. De Troch**, O. Giot, R. Hamdi, A. Deckmyn, and P. Termonia: Future climate and air quality of the Brussels Capital Region for the 2050s under the A1B scenario. 34th International Technical Meeting on Air Pollution Modelling and its Application (ITM), Montpellier, France, May 2015.

10. Hamdi, R., H. Van de Vyver, **R. De Troch**, and P. Termonia: Assessment of three dynamical urban climate downscaling methods. 9th International Conference on Urban Climate, Toulouse, France, July 2015.

11. **De Troch, R.**, P. Termonia, O. Giot, R. Hamdi, and J. Berckmans: Practical assessment of climate change impact on winter smog episodes: A case study for Brussels. Belgian Geography Days, Brussels, Belgium, November 2015.

Poster Presentations

1. **De Troch, R.**, R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia: Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. European Geosciences Union General Assembly, Vienna, Austria, April 2013.

2. Hamdi, R., H. Van de Vyver, **R. De Troch**, A. Delcloo, and P. Termonia: Assessment of three dynamical urban climate downscaling methods: Brussels's urban heat island under an A1B emission scenario. International Technical Meeting on Air Pollution Modelling and its Application, Miami, Florida USA, August 2013.

3. Hamdi, R., **R. De Troch**, O. Giot, and P. Termonia: Regional Climate Modelling at the RMI. Scientific conference on the occasion of the Centenary of the Royal Meteorological Institute of Belgium, Brussels, Belgium, September 2013.

4. Hamdi, R., **R. De Troch**, O. Giot, and P. Termonia: Regional Climate Modelling at the RMI. 35th EWGLAM and 20th SRNWP meeting, Antalya, Turkey, October 2013.

5. Hamdi, R., H. Van de Vyver, **R. De Troch**, A. Delcloo, and P. Termonia: Assessment of three dynamical urban climate downscaling methods: Brussels's urban heat island under an A1B emission scenario. International Conference on Regional Climate CORDEX 2013, Brussels, Belgium, November 2013.

6. Giot, O., **R. De Troch**, R. Hamdi, A. Deckmyn, and P. Termonia: Impact of using the surface scheme SURFEX within ALARO-0 for the ERA-Interim high resolution dynamical downscaling over Belgium. 3rd International Lund Regional-Scale Climate Modelling Workshop, Lund, Sweden, June 2014.

7. **De Troch, R.**, J. Berckmans, O. Giot, R. Hamdi, and P. Termonia: Future climate impact on meteorological conditions that are unfavorable for the dispersion of air pollution in Brussels. European Geosciences Union General Assembly, Vienna, Austria, April 2015.

8. Berckmans, J., R. Hamdi, **R. De Troch**, and O. Giot: Validation of the Regional Climate Model ALARO with different dynamical downscaling approaches and different horizontal resolutions. European Geosciences Union General Assembly, Vienna, Austria, April 2015.

9. Berckmans, J., S. Caluwaerts, L. De Cruz, **R. De Troch**, A. Deckmyn, D. Degrauwe, P. De Meutter, A. Duerinckx, L. Gerard, O. Giot, R. Hamdi, G. Smet, P. Termonia, J. Van Den Bergh, and M. Van Ginderachter: Belgium National poster. Joint 25th ALADIN Workshop & HIRLAM All Staff Meeting, Helsingor, Denmark, April 2015.

10. Giot, O., P. Termonia, **R. De Troch**, S. Caluwaerts, and G. Smet: Validation of ALARO-0 within the EURO-CORDEX framework. European Geosciences Union General Assembly, Vienna, Austria, April 2015.

11. **De Troch, R.**, A. Delcloo, R. Hamdi, O. Giot, A. Deckmyn, and P. Termonia: Future air quality of the Brussels Capital Region for the 2050s under the A1B emission scenario. 9th International Conference on Urban Climate, Toulouse, France, July 2015.

12. Hamdi, R., O. Giot, **R. De Troch**, A. Deckmyn, and P. Termonia: Future climate of Brussels and Paris for the 2050s under the A1B scenario. 9th International Conference on Urban Climate, Toulouse, France, July 2015.

13. **De Troch, R.**, B. Van Schaeybroeck, P. Termonia, P. Willems, N. van Lipzig, J.-P. Van Ypersele, X. Fettweis, K. De Ridder, A. Gobin, T. Stavrakou, P. Luyten, and E. Pottiaux: Combining the regional downscaling expertise in Belgium: CORDEX and beyond. Belgian Geography Days, Brussels, Belgium, November 2015.

REFERENCES   Dr. Piet Termonia
                   Guest Professor at Ghent University   Phone: +32 2 373 06 38
                   Head of the Research Department of the   E-mail: piet.termonia@meteo.be
                   Royal Meteorological Institute of Belgium