

Future multivariate weather generation by combining Bartlett-Lewis and vine copula models

J. Van de Velde^{a,b} and M. Demuzere^c and B. De Baets^b and N.E.C Verhoest^a

^aHydro-Climatic Extremes Lab, Ghent University, Ghent, Belgium;

^bKERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium;

^cDepartment of Geography, Ruhr-University Bochum, Bochum, Germany

ARTICLE HISTORY

Compiled April 12, 2023

ABSTRACT

The assessment of future extremes is hindered by the lack of long time series. Weather generators can alleviate this problem, but easily become more complex when generating multiple variables. In this study, a weather generator combining Bartlett-Lewis models and vine copulas is presented. The combination of these models allows for the stochastic and physically coherent generation of longer time series with statistics similar to those of the time series used as input. This model chain has already been assessed on the basis of historical observations, but never on the basis of future simulations. However, the model chain could have practical value for extending climate simulations, which should be investigated. Combining recent versions of the Bartlett-Lewis model (for the generation of precipitation) and vine copulas (for the generation of temperature and evaporation), the model was applied for two time series of historical observations and one time series simulated by the RCA4 RCM for the years 2071-2100. For the future simulations, the weather generator performed comparably as for the historical observations for the statistical moments and the correlation. The results for the extremes were more ambiguous, but still provided valuable information. The adequate performance for the statistical moments and the correlation, combined with the continuous development of both Bartlett-Lewis models and vine copulas, indicates that the weather generator might be of use for the characterization of extreme events under climate change.

KEYWORDS

Weather generation; Copulas; Bartlett-Lewis models; Climate change

1. Introduction

Under climate change, it is important to correctly design hydraulic structures. The simulation of future statistical moments by climate models is nowadays considered standard practice for this purpose (François, Schlef, Wi, & Brown, 2019). Although climate models, whether they are Global Climate Models (GCMs) or Regional Climate Models (RCMs), run detailed simulations for the full 21st century, often only the end of the century is considered for long-term planning purposes. Data from e.g., the years 2071 until 2100, present an assumed stable climate and the extreme values can be compared with recent (e.g., 1970-2000) or current (e.g., 1990-2020) climate

(e.g., Hirabayashi et al. (2013)). However, with only 30 years of data, there are few extreme events and the uncertainty becomes very large (Brunner, Slater, Tallaksen, & Clark, 2021). This lack of data can be overcome with weather generators (Maraun et al., 2010; Wilks & Wilby, 1999), which can extend time series while preserving the statistical moments. As such, this extension allows for inferring more information from a climatological time series.

Recently, Pham, Vernieuwe, De Baets, and Verhoest (2018) developed a weather generator combining Bartlett-Lewis models for precipitation generation (Rodriguez-Iturbe, Cox, & Isham, 1987) and vine copula models (Aas, Czado, Frigessi, & Bakken, 2009) for the generation of temperature and evaporation. The combination of both models allows to stochastically generate the three variables conditionally on each other, which is valuable for many impact models (Zscheischler et al., 2018). Bartlett-Lewis models combine rectangular rainfall blocks, which resemble small mesoscale precipitation systems (Burlando & Rosso, 1993; Onof & Wang, 2020). Vine copula models allow for the consistent generation of both temperature and evaporation by exploiting the relationship between multiple variables (Pham, Vernieuwe, De Baets, Willems, & Verhoest, 2016; Pham et al., 2018). The stochastic element is obtained by random sampling: a combination of Poisson processes and sampling from gamma distributions in the Bartlett-Lewis model, and sampling from a uniform distribution in the vine copula models.

Although the weather generator developed by Pham et al. (2018) might thus be a powerful tool for climate change impact assessment, it has only been calibrated and assessed on historical conditions. Nonetheless, the set-up should in theory be transferable to future conditions. By conditioning on future time series simulated by climate models and hence generating longer time series, the weather generator might improve the assessment of extreme meteorological values. There are only a few examples of such a direct application of Bartlett-Lewis models on climate simulations (e.g., Onof and Arnbjerg-Nielsen (2009) or Cross, Onof, and Winter (2020)). In contrast, most studies applying Bartlett-Lewis or related models in a climate change context rescale the weather generator model parameters on the basis of the difference between historical and future climate model simulations (e.g., Kilsby et al. (2007), Burton, Fowler, Blenkinsop, and Kilsby (2010), Fatichi, Ivanov, and Caporali (2011)). This ‘change factor’ method has been criticised by Cross et al. (2020) on the basis of two arguments. First, it assumes that the scaling relationships of rainfall statistics are stationary between the present and the future. The second problematic assumption is the relationship between mesoscale circulation and point precipitation as simulated by the Bartlett-Lewis models. Both relationships could change in the future, and evidence is indeed mounting that precipitation dynamics will change in north-western Europe, the area under study here (Bevacqua et al., 2021; Bevacqua, Zappa, & Shepherd, 2020; Kahraman, Kendon, Chan, & Fowler, 2021). Thus, change factor methods might become less applicable and new ways to apply the Bartlett-Lewis models under climate change conditions have to be explored, such as the weather generator studied here. Similarly, although the application of standard bivariate copulas in climate science has grown steadily over the past decade (Schölzel & Friederichs, 2008; Tootoonchi et al., 2022), the application of the multivariate vine copulas in weather generator or similar set-ups is limited (e.g. Vernieuwe, Vandenbergh, De Baets, and Verhoest (2015)). Yet, they are gaining popularity in climate science (e.g., Hobaek Haff, Frigessi, and Maraun (2015), Bevacqua, Maraun, Hobaek Haff, Widmann, and Vrac (2017), Sun et al. (2021)) because of their ability to model the joint distribution of any number of variables. Thus, implementations within the field of weather generation might prove

fruitful.

Given the sparse application of both Bartlett-Lewis models and vine copulas within weather generation, the knowledge of the transferability of such models under climate change conditions is limited: it is possible that this model chain cannot completely deal with the nonstationarity induced by climate change. Consequently, its performance has to be assessed. If the quality of generated time series based on future simulations is similar to the quality of the time series generated on the basis of historical data, it can be assumed that the weather generator is transferable to future simulations, where it could be of use not only for hydrological impact analysis, but also many other types of impact. For example, the study of multivariate compound events (Zscheischler et al., 2020) could profit from a conceptually relatively simple multivariate weather generator.

In summary, in this paper we will assess for the first time the direct calibration of a model chain consisting of a Bartlett-Lewis model and vine copula models on climate model output, with the goal of extending the climate model time series. This should allow for a better investigation of climate extremes. The Bartlett-Lewis model under study is point-based and hence only observed and simulated data for Uccle, Belgium, will be used. To assess whether biases in the model chain affect climate change impact analysis, the generated data will be used as input for a hydrological model.

2. Methods

2.1. Data

Two sets of time series are used in the present paper. For the historical calibration of the weather generator, the meteorological time series of Uccle, collected by the Royal Meteorological Institute of Belgium, is used. This time series consists of 100 years (1901-2000) of daily temperature, 120 years of 10-min precipitation (limited to a series from 1901 to 2000 and aggregated to the daily level to be consistent with the temperature) and daily potential evaporation, calculated using the Penman-Monteith method on the basis of other measured variables. The precipitation time series has been analysed frequently (Bertrand, Ingels, & Journée, 2021; De Jongh, Verhoest, & De Troch, 2006; Demarée, 2003) and was already used for the calibration of Bartlett-Lewis models in earlier studies, based on a part of the complete series (Onof & Wang, 2020; Vandenbergh, Verhoest, Onof, & De Baets, 2011; Verhoest, Troch, & De Troch, 1997).

For the future calibration, end-of-the-century data (2071-2100) from the Rossby Centre regional climate model RCA4 were used (Strandberg et al., 2015), as it is one of the few models with potential evaporation as an output variable. This model is part of the EURO-CORDEX project (Jacob et al., 2014) and was forced with boundary conditions from the MPI-ESM-LR GCM (Popke, Stevens, & Voigt, 2013). As climate model data, and especially climate model precipitation data, are often biased (Maraun et al., 2010), the Quantile Delta Mapping (QDM, Cannon, Sobie, and Murdock (2015)) bias-adjusting method was applied on all variables before the calibration of different models, similar to Van de Velde, Demuzere, De Baets, and Verhoest (2022). Although multiple variables are considered in this set-up, we have chosen for the univariate bias adjustment by QDM, as this method was shown to be robust under climate change for the Uccle dataset (Van de Velde et al., 2022). Given the limited spatial resolution of only 0.11° , or 12.5 km, and the limited topography around Uccle (Brussels area), we assumed that the meteorological conditions were similar enough to only

apply bias adjustment. Nevertheless, a bias-adjusting method applied to RCM data will always work as a de facto downscaling method. This should be avoided, as both bias adjustment and statistical downscaling have different goals (Lange, 2019). In addition, the large-scale variability will be transposed to the point level, inflating the local variability, unless a stochastic method is used (Maraun, 2013). As QDM is not stochastic, this should be taken into account, although earlier work illustrated the robust performance of this method (Van de Velde et al., 2022). However, while important for impact assessment, this does not influence the argumentation in this paper. In fact, the application of QDM transforms the RCM data into point-based data, which is the most logical choice for Bartlett-Lewis calibration.

2.2. Stochastic generator chain

2.2.1. Bartlett-Lewis model

The first step in the model chain consists of a Bartlett-Lewis model (Rodriguez-Iturbe, De Power, & Valdes, 1987). Bartlett-Lewis models generate a precipitation time series by generating storms using a Poisson process. Within storms, multiple overlapping rain cells are simulated according to another Poisson process: the final precipitation amount at a time step t is based on the superposition of the rain cells. The strength of Bartlett-Lewis models lies in their relative simplicity and a physical interpretation as small mesoscale precipitation systems (Burlando & Rosso, 1993). Although BL models performed adequately for most statistical moments at many locations (Onof et al., 2000), they have been further improved, e.g., for precipitation extremes (Kaczmarek et al., 2014), sub-hourly rainfall simulation (Onof & Wang, 2020) and for low-frequency variability (Kim & Onof, 2020; Park, Onof, & Kim, 2019).

In the original version of the model chain (Pham et al., 2018), the ‘Randomized Bartlett-Lewis model’ (RBL) (Rodriguez-Iturbe, Cox, & Isham, 1988) was used. However, the RBL model underestimates precipitation extremes, which can propagate through the model chain and in impact models (Verhoest et al., 2010). Hence, Kaczmarek et al. (2014) proposed a new model version, the ‘Randomized Bartlett-Lewis model 2’ (RBL2), which performed well in recent studies (Onof & Wang, 2020). In this paper, this more recent version (Onof & Wang, 2020) is used for the generation of precipitation time series.

Seven parameters describe the RBL2 model (Fig. 1). First, a series of rainstorm arrivals is generated according to a Poisson process with parameter λ [1/T]. For each of the corresponding rainstorms, a temporal scaling factor η [1/T] is generated from a gamma distribution with shape and scale parameters α [–] and $1/\nu$ [1/T], respectively. The duration of the rainstorms is exponentially distributed with parameter $\eta\phi$, with ϕ [–] a model parameter. The rain cells in a storm arrive according to a Poisson process with parameter $\eta\kappa$, with κ [–] an additional model parameter. The rain cell duration is exponentially distributed with parameter η and the intensity is gamma distributed with shape and scale parameters ω [–] and $\nu\eta/\omega$, where the scaling factor ν [L] determines the relation between intensity and duration. Thus, the seven parameters of the model are λ , α , ν , ϕ , κ , ω and ν .

2.2.2. Vine copula models

For the simulation of evaporation and temperature, the vine copula models by Pham et al. (2016, 2018) are used. These models apply the vine copula theory initiated by

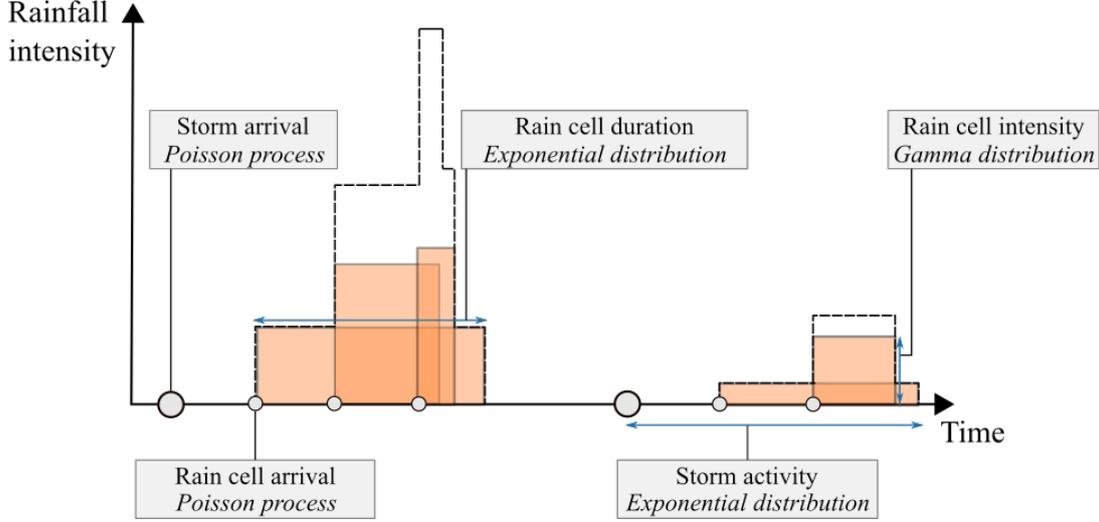


Figure 1.: Illustration of the rain intensity (dashed line) generation in the RBL2 model

Bedford and Cooke (2001, 2002); Joe (1996). Using bivariate copulas (Nelsen, 2006), the joint distribution of any number of variables can modeled using vine copulas. As a short reminder, Sklar’s theorem (Sklar, 1959) states that for every bivariate distribution H with marginals $F(x), G(y)$, a copula C exists such that for all x, y in $\bar{\mathbb{R}}$

$$H(x, y) = C(F(x), G(y)),$$

with $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. If F and G are continuous, then C is unique.

Although other methods for modelling multivariate dependencies also exist (Joe, 2014), vine copulas have the advantage of offering a flexible choice in bivariate copulas (Aas, 2016; Aas & Berg, 2009). Trees, as defined in graph theory (Diestel, 2018), form the basis of the vine copula construction, and the combination of multiple trees resembles a vine, hence the name. Each tree consists of nodes and edges. In the first tree, the nodes represent the variables used and each bivariate copula is associated with an edge. These bivariate copulas are transformed into the nodes of the next tree according to the following formula:

$$F(x | \mathbf{v}) = \frac{\partial C_{x, v_j | \mathbf{v}_{-j}}(F(x | \mathbf{v}_{-j}) F(v_j | \mathbf{v}_{-j}))}{\partial F(v_j | \mathbf{v}_{-j})}, \quad (1)$$

where $F(\cdot | \mathbf{v})$ is the conditional cumulative distribution function, \mathbf{v} is a vector of conditioning variables, v_j is an arbitrarily chosen component of \mathbf{v} and \mathbf{v}_{-j} is the vector without this component. The choice of the conditioned variable x and the vector of conditioning variables \mathbf{v} determine the structure of the vine copula, where the number of possible constructions rapidly increases with an increasing number of variables. The regular vine was introduced by Bedford and Cooke (2001, 2002) to guide the construction, but still allows for many possible vine copulas. The subclasses of C- and D-vines further constrain the choice. In C-vines, there is one variable in every tree that is joined with every other variable. In contrast, in D-vines, variables are joined with no more than two other variables. In the present paper, we only apply C-vines, as

only they allow for a driving meteorological variable in the vine copula construction.

2.3. Set-up

2.3.1. Ensemble overview

In the present paper, three different set-ups were run to investigate whether the stochastic generator chain can be easily used in a future climate change setting. In the first set-up, 100 years of the daily Uccle time series of P , T and E are used to calibrate the models and stochastically generate 100 years of daily data for the three variables. Care should be taken with this time series, as it is much longer than what can be considered as a climatological time series according to the WMO definition (Trewin, 2007). However, in the context of hydrological extreme-value analysis, a 100-year time series is valuable and a comparison of models calibrated on shorter time series with models calibrated on this time series could be informative. To account for nonstationarity probably present in the time series, a detrending is performed, as detailed in Section 2.3.2. In the second set-up, only the last 30 years (1971-2000) of the Uccle dataset are used to calibrate, but once again 100 years of data are generated. If the statistical moments of the 30-year calibrated time series are similar to those of the 100-year calibrated time series, this provides strong evidence that calibrating the models on 30 years of future data can provide valuable long time series of climatological data. Finally, 30 years of future RCA4 simulations are used to calibrate the model chain and subsequently generate 100 years of P , T and E data. The three resulting 100-year time series will be discussed in what follows as ‘Hist100’ (calibration on 100 years of historical data), ‘Hist30’ (calibration on 30 years of historical data) and ‘Fut30’ (calibration on 30 years of future data). The details of the calibration and simulation are discussed in Sections 2.3.2 and 2.3.3.

To account for the stochasticity in the models, an ensemble is created in each of the aforementioned set-ups. In every set-up, the RBL2 is run 20 times. Temperature is then generated 20 times conditioned on each RBL2 run. For each of the resulting 400 runs, evaporation is generated 20 times conditioned on the 400 temperature series and the corresponding 20 precipitation series. Thus, in the end we obtain 8000 100-year daily time series of evaporation, 400 100-year daily time series of temperature and 20 100-year daily time series of precipitation. This ensemble allows for an adequate assessment of the statistical moments of each of these variables.

2.3.2. Calibration and model fit

The Bartlett-Lewis model is calibrated with the generalized method of moments (GMM) (see Jesus and Chandler (2011); Vanhaute, Vandenberghe, Scheerlinck, De Baets, and Verhoest (2012) for a thorough overview). In this calibration method, k theoretical and observed statistics are compared according to the objective function

$$S(\theta \mid \mathbf{T}) = \sum_{i=1}^k w_i (T_i(y) - \tau_i(\theta))^2, \quad (2)$$

with $T(y)$ the observed statistic, τ the expected value under the given model with parameter vector θ and w a weight for the given statistic. Based on the work of Jesus and Chandler (2011), the weight matrix is the diagonal matrix of inverse variances. The

chosen statistics are the mean, the variance, the lag-1 autocovariance, the zero depth probability and the skewness at the aggregation levels of 24 h, 48 h and 72 h; each month is calibrated separately. The objective function is optimized with the Shuffled Complex Evolution algorithm (Duan, Gupta, & Sorooshian, 1993), as suggested by Vanhaute et al. (2012). The optimization algorithm is repeated 20 times, and for each month, the best-scoring result is used to determine the parameter values.

For the simulation of temperature, a 3-dimensional vine copula is fit to the temperature (T) and precipitation (P) of the same day and temperature of the previous day (T_p). In what follows, this vine copula will be denoted as V_{T_pPT} . For the simulation of evaporation (E), this variable is combined with temperature (T), precipitation (P) and evaporation on the previous day (E_p), fitting a 4-dimensional vine copula V_{TPE_pE} . Before the data are fit, they should be independent and identically distributed. To ensure this, the time series is split into monthly series and a vine copula model is fit to every month. Within-month trends are assessed by an ANOVA test when distributions are homoscedastic, a Welch ANOVA test when distributions are heteroscedastic (Welch, 1947), or a Kruskal-Wallis test (Kruskal & Wallis, 1952) when distributions are non-normal and heteroscedastic, at a significance level of 0.001. The within-month trends are then removed to standardize temperature and evaporation before fitting the copula models. The fit itself is realized using the VineCopula R package (Nagler et al., 2021). The choice was limited to C-vines, and only Gaussian, Clayton, Gumbel and Frank copula models were considered for the fit. The C-vine structure was defined before the fit, with T_p and T the central variables of respectively V_{T_pPT} and V_{TPE_pE} . This choice was based on the high correlation of these variables with respectively T and E in the original time series. The copula families were fit at the bivariate copula level, with the AIC (Akaike, 1973) as criterion. Once all bivariate copulas were fit, the global goodness-of-fit was assessed using the White goodness-of-fit test (Schepsmeier, 2015, 2019).

The calibrated RBL2 and vine copula parameter values can be found in the Appendix, respectively in Table A1 and Tables B1–B6. Although slightly varying, most RBL2 parameter values have the same order of magnitude across the three calibration time series. The only exception is ω , which varies more when calibrated for both 30-year time series in comparison with calibration for the 100-year time series. It thus seems that the longer time series constrains this parameter. For the vine copulas, Tables B1–B6 show that varying configurations are obtained depending on the month. As can be seen from these tables, the configurations all resemble each other and no set-up has remarkable differences with an other. Yet, larger differences could occur with a broader selection of copula families. The White goodness-of-fit test results indicate that the global fit of the V_{T_pPT} vine copula is poor for all three set-ups. This is especially pronounced in the Hist30 and Fut30 set-ups; in the Hist100 set-up, only a few months have a globally poor fit. In addition, the Hist30 set-up has a poor global fit for the V_{TPE_pE} vine copula. The poor fits do not necessarily imply that the estimation and fitting procedures are futile. The bivariate copula families and their parameters are determined tree by tree with the sequential method (Aas, 2016), whereas the goodness-of-fit is determined globally. This causes a trade-off between the global and local fit in vine copula fitting (Dißmann, Brechmann, Czado, & Kurowicka, 2013).

2.3.3. Sampling

The generation of new time series by RBL2 is relatively straightforward. Based on the monthly parameters, rain cells are generated according to the principles described in

Section 2.2.1. The precipitation of these cells is summed to obtain a daily precipitation amount.

The sampling of the vine copulas is done according to Section 2.2.2. If the conditioning vector v is univariate, Eq. (1) can be simplified as the h -function (Aas et al., 2009):

$$h(x, v, \Theta) = F(x | v) = \frac{\partial C_{x,v}(x, v, \Theta)}{\partial v}, \quad (3)$$

where the second variable always corresponds to the conditioning variable and Θ denotes the vector of parameters of the copula. For the sampling, either Eq. (3) or its inverse is used. As the structure of the vine copulas is defined such that the unknown variable (T or E) is the fourth variable, the sampling becomes easy to understand. First, the h -function is applied for all copulas joining known variables, such as C_{TP} in the first or $C_{PE_p|T}$ in the second tree of the V_{TPE_pE} vine copula. This application thus passes through all trees sequentially, using the values for conditional cumulative probabilities calculated on the basis of the former tree as input for the following bivariate copula. Upon reaching the last tree, a random value is uniformly drawn from $[0, 1]$ to describe the cumulative probability of the vine. From this point on, the inverse of the h -function is used to calculate the unknown conditional cumulative probabilities, starting from the second-to-last tree. This is an implementation of the standard vine sampling algorithm; Czado (2019) provides a full overview.

2.4. Hydrological model

In this study, we use the Probability Distributed Model (PDM, Moore (2007)) to analyze the impact of potential biases in the stochastic generator chain. The PDM is a lumped conceptual rainfall-runoff model that uses precipitation and evaporation time series as inputs to generate a discharge time series. The PDM as used here was calibrated for the Grote Nete watershed in Belgium (RMSE = 0.9 m³, see Pham et al. (2018) for more details). The distance between the Uccle measurement station and the watershed is ± 50 km. The Uccle grid cell thus does not contain the watershed, but given the limited distance and the flat topography, the meteorological conditions can be assumed to be similar. Besides, the intended application of the hydrological model is not to make predictions, but to illustrate the sensitivity of an impact model to biases generated by the stochastic generator chain. Hence, no observed discharge simulations will be used for comparison, but the original Uccle observations or climate model output, i.e., the input for the stochastic generator chain, will be used as input for the hydrological model.

3. Results

3.1. Distribution

The probability density functions (PDFs) of the variables allow us to determine whether the yearly distribution is well reproduced by the stochastic generators. The PDFs are shown in Fig. 2 and all indicate that the model chain is able to correctly reproduce the distributions, apart from some small deviations. This is confirmed by the Perkins's Skill Score (PSS) values (Perkins, Pitman, Holbrook, & McAneney, 2007),

which are all higher than 0.95 (on a maximum of 1).

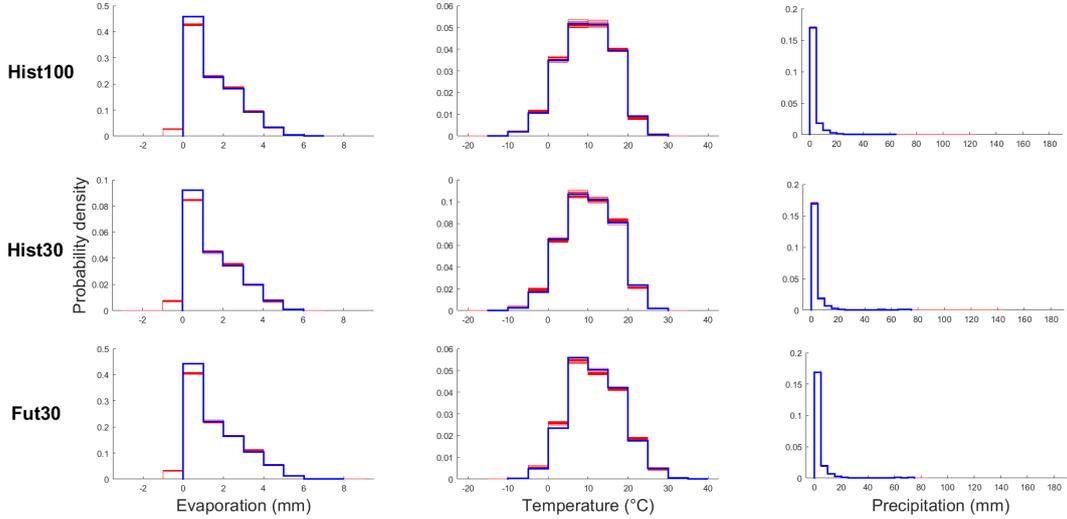


Figure 2.: Original (blue lines) and ensemble of generated (red lines) PDFs of evaporation, temperature and precipitation for the three set-ups.

Although the results for the PDFs indicate that the distributions are well reproduced, the statistical moments should be studied more in depth. As all models were calibrated monthly, studying the moments at a monthly level allows for a better assessment of the calibration quality. The moments for P are shown in Fig. 3 for the 24 h aggregation level. In the Hist100 set-up, all moments are generally simulated adequately, apart from a slight underestimation of the variance in June and August and an overestimation of the third central moment in June and August. The performance of the Hist30 set-up is similar to that of the Hist100 set-up. Only the results for the mean are slightly worse, but the results for the third central moment are better, with less pronounced outliers. The good performance of the Hist30 set-up implies that the time series length has a limited effect on the calibration. This conclusion is corroborated by the results for the Fut30 set-up, although some months display large biases. For the mean, February and November are consistently underestimated by the model. For May and June, and to some extent September and November, the variance, autocovariance and third central moment are poorly simulated. As the results for these months are generally better in the historical set-ups, this could be an effect of climate change, and especially the effect of climate change on the extremes. For example, the precipitation amount in one September month was an extreme outlier, which had a large influence on the calibration for September. This month was consequently removed from the data, leading to an improved calibration. The results for this September month imply that months with single high precipitation extremes can exacerbate the results. However, removal of more months might create unrepresentative and unrealistic results. Alternatively, one could consider to work with longer time series through which single events have less impact on the calibration. Yet, this discards the meteorological definition of ‘climate’, which is relevant for communication and comparison purposes.

The moments for temperature (Fig. 4) are adequately simulated for every set-up. For variance and autocovariance, there are some limited biases, but the original values

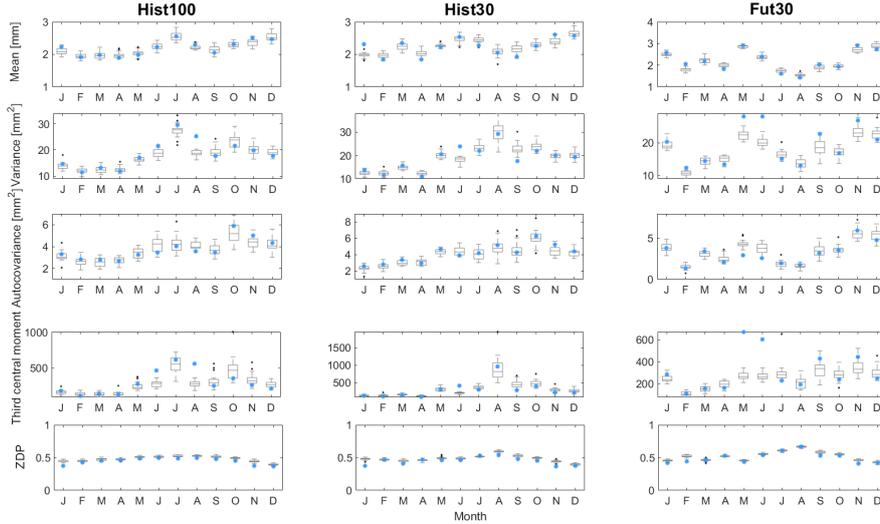


Figure 3.: Original (blue dots) and generated (box-plots) moments of the precipitation time series for the 24 h aggregation level.

are still included in the interquartile range. Although the statistical moments for precipitation were poorly reproduced in the Hist30 time series, the bias for temperature is much smaller, which illustrates the robustness of the weather generator: biases are not simply recreated or enlarged in subsequent steps. For evaporation, Fig. 5 shows that the moments of original time series are satisfactorily reproduced by the Hist100 set-up, although the autocovariance is slightly underestimated from May until August. The Hist30 and Fut30 set-ups perform similarly. Nonetheless, for all three time series, the underestimation is small in absolute values and the reproduction of the mean and variance is the most important aspect.

3.2. Extremes

One of the main goals of the weather generator is the assessment of extreme weather conditions. As the precipitation variable is most skewed, we will focus on the precipitation extremes in this section. From Fig. 6, it can be seen that the results are far more consistent for the historical set-ups than for the Fut30 set-up. While the Hist100 and Hist30 set-ups respectively adequately reproduce and slightly underestimate the highest return period, the future set-up underestimates all return periods larger than 5 years. Given the results for the calibration (Fig. 3), it seems that the limited capability of the Bartlett-Lewis model to correctly calibrate the third central moment in the Fut30 set-up is the reason for the poor reproduction of the extremes. This becomes more clear when analyzing the extremes on a monthly basis (not shown): the months with poor calibration results for variance, autocovariance and third central moment also show a poor performance for the reproduction of the extremes. In the context of the application of the Bartlett-Lewis model on future data, it is important to understand where the extremes in the original climate data originate from. Climate variability is the most obvious explanation, e.g., that these are extremes with a higher return period than 30 years which are included in the limited 30-year time series. There are two arguments in favour of this hypothesis. First, the extremes are appropriately

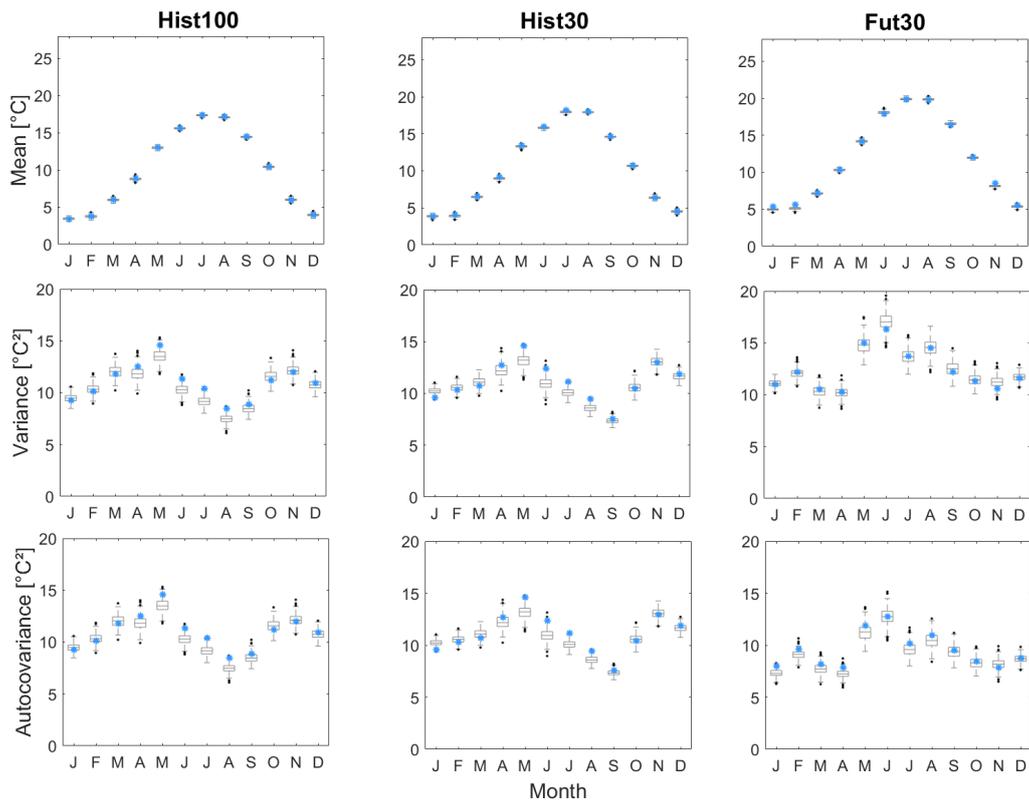


Figure 4.: Original (blue) and generated (box-plots) statistical moments of temperature for the three set-ups.

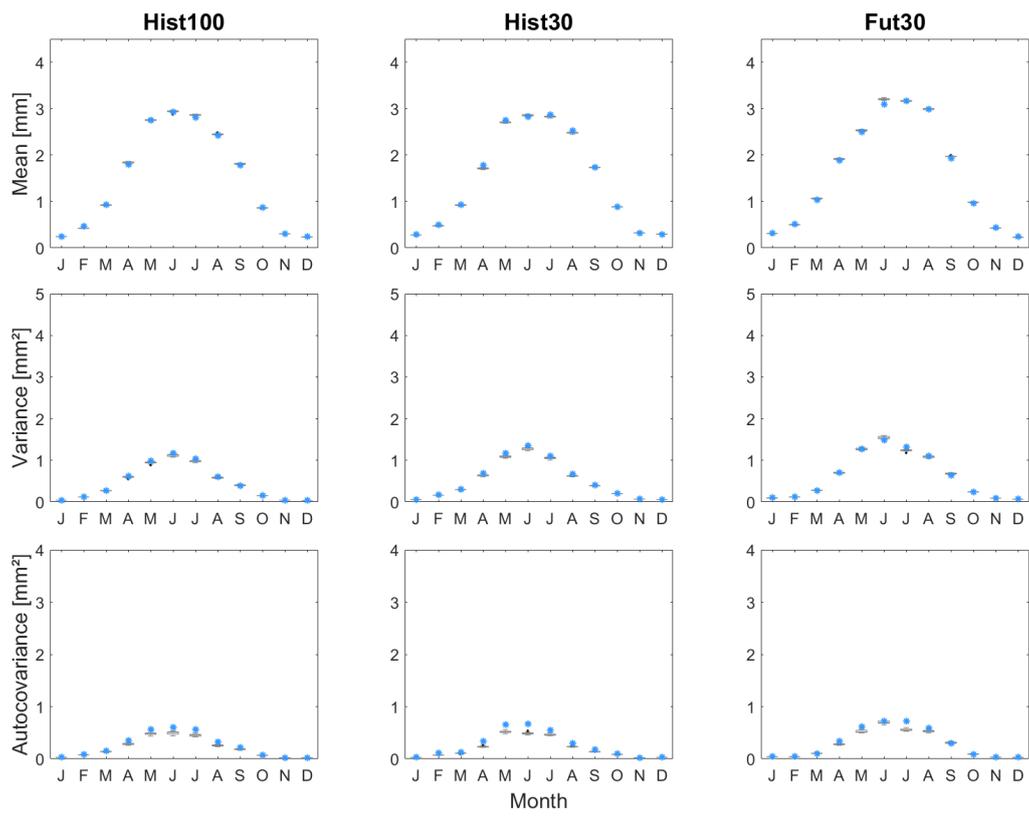


Figure 5.: Original (blue) and generated (box-plots) statistical moments of evaporation for the three set-ups

reproduced in the historical time series. Second, when running the Bartlett-Lewis simulations for 1000 years (not shown), the highest precipitation values in the original time series correspond with return periods of ± 100 years, which is reasonable. Thus, the combined results for all three set-ups suggest that the RBL2 model is able to appropriately simulate extremes, with the length of the input time series as the largest source of uncertainty.

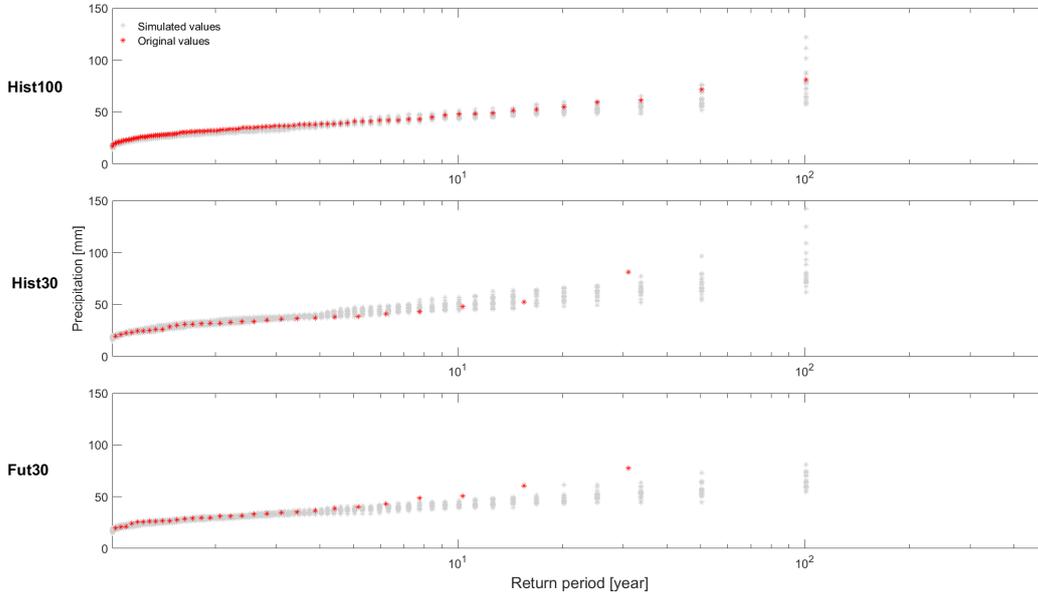


Figure 6.: Original (red) and ensemble of generated (grey) rainfall extremes for all three set-ups.

3.3. Correlation

As the preservation of the correlations is an important goal of the application of vine copulas, these should be properly assessed. In Fig. 7, the Kendall tau values for the original and generated time series are shown. For every plot, only 20 randomly chosen pairs were used. For all three set-ups, the same patterns emerge: the P - T and P - E relationships are slightly underestimated, but the E - T correlation is adequately reproduced. Despite the underestimation for P - T and P - E , the seasonal pattern is represented quite well, implying that with some small modifications, the correlation could be even better reproduced. The easiest modification would be the inclusion of more copula families, although this would increase fitting complexity. In addition, copula modelling is continuously evolving (Aas, 2016; Gröber & Okhrin, 2021) and recent advances in parameter estimation and goodness-of-fit testing could also improve the results. Less straightforward would be the inclusion of more advanced models, such as dynamic vine copulas (Acar, Czado, & Lysy, 2019).

3.4. Discharge

The application of a hydrological model allows for understanding the impact of biases in the stochastic generator chain on a subsequent impact model. The difference be-

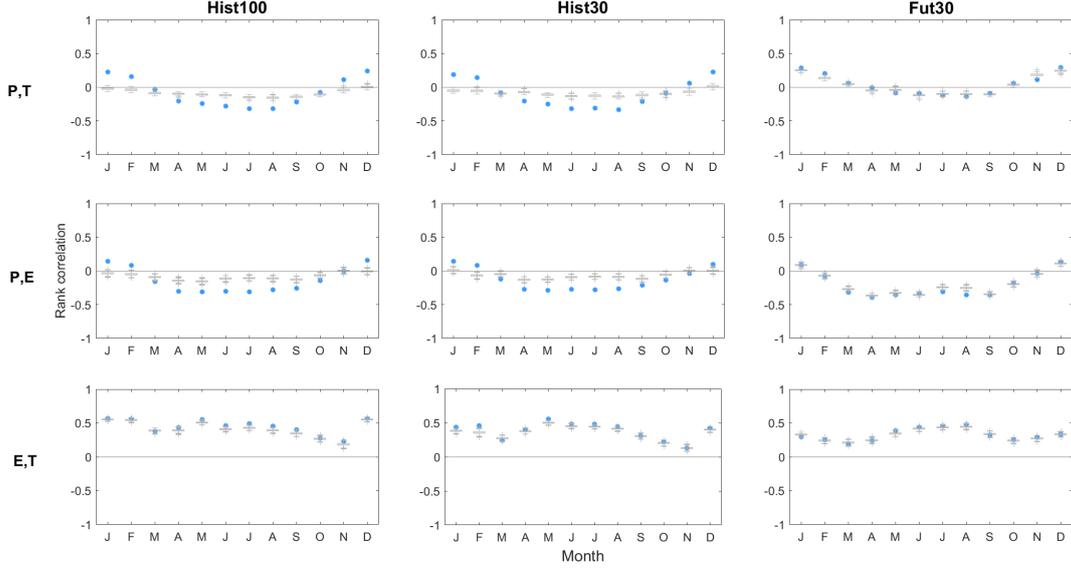


Figure 7.: Original (blue) and generated (box-plots) correlations for the three set-ups

tween Hist100 and Hist30 on the one hand, and Fut30 on the other hand, was clearest for the moments of precipitation (Fig. 3) and the precipitation extremes (Fig. 6). As both have a large impact on the discharge, we can also expect a difference between the discharge calculated on the basis of historical or future set-ups. The discharge moments (Fig. 8) and extremes (Fig. 9) illustrate the propagation of the biases in the Fut30 set-up. For the moments, the biases are largest in January, March and November. Except for November, these months do not correspond with biased months for the precipitation moments (Fig. 3), i.e., May, June and September. This illustrates how a hydrological model integrates P and E over a longer time period and thus makes it more complex to link biases in the discharge with biases in precipitation. This integration and complex link is a principle that can be extended to most impact models. The discharge biases are clearest for variance and autocovariance. As was noticed for precipitation, a bias in variability often corresponds with biases in the extremes. A comparison of the historical set-ups and Fut30 illustrates this. Hist100 and Hist30 are only slightly biased: they mostly overestimate the original variance and autocovariance (except for January). This overestimation leads to a small overestimation of the extremes. However, the biases for Fut30 are larger. For January, March and November, the variance is clearly underestimated, leading to an underestimation of the extremes. Yet, as for precipitation, the biases can to some extent be attributed to the limited length of the time series. This implies that with either longer time series or longer simulations, the model chain can yield relevant results for impact models.

4. Discussion and conclusion

In this paper, we studied the reproduction of different statistical aspects of precipitation, temperature and evaporation by a multivariate stochastic weather generator. In general, the results for the three set-ups considered show that most aspects are as well reproduced in the future time series as in the historical time series: the PDFs (Fig. 2), statistical moments (Figs. 3–5) and correlation (Fig. 7) were all fairly well

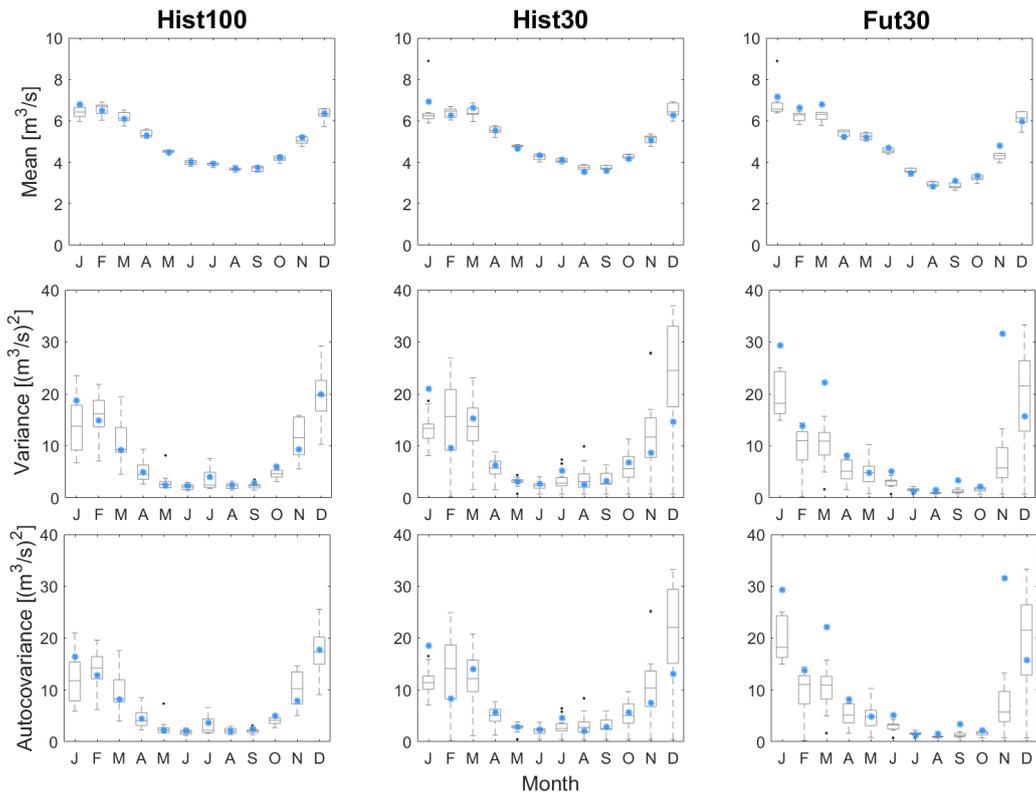


Figure 8.: Original (blue) and generated (box-plots) statistical moments of discharge for the three set-ups.

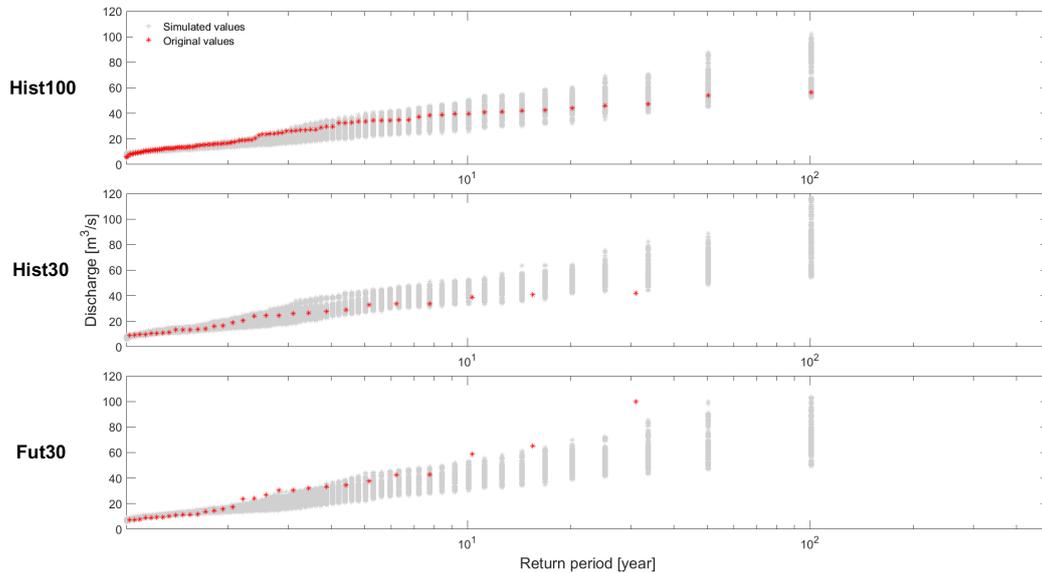


Figure 9.: Original (red) and ensemble of generated (grey) discharge extremes for all three set-ups.

reproduced, although there are some considerable biases in the Fut30 set-up depending on the month or the statistical aspect. Yet, the largest biases were present in the reproduction of the extremes (Fig. 6). In contrast with the historical set-ups, the extremes were clearly underestimated in the Fut30 set-up. As a consequence, discharge extremes (Fig. 9) were also underestimated.

The results for the extreme values for precipitation are linked to the biases in the statistics. The highest extreme values present in the original data are underestimated by the Bartlett-Lewis model (Fig. 6). On a yearly basis, it is hard to pinpoint the origin of these biases. However, on a monthly basis, there is a clear correspondence between a poor calibration and poor results for the extremes. As it is clear that anomalously high extremes exacerbate the calibration and hence the reproduction of extremes, they should be better understood. There is always a chance that precipitation extremes with an effective return period larger than 30 years occur in the limited 30-year time series. This seems to be the major reason for difficulties in the calibration of the Bartlett-Lewis model. Despite these difficulties, the model chain can yield relevant results. The good performance of the Hist100 and Hist30 set-ups indicates that the model chain is preferably applied to generate long simulations of at least a few 100 years. These long-term results could provide extremes with a more realistic occurrence probability than the original 30-year time series, which was the goal of the paper. If only the original time series would be used, the occurrence probability of the highest extremes would be inflated, leading to possible over-design. In addition, impact models often integrate meteorological aspects over longer time frames and may be non-linear, as was shown for the discharge discussed in this paper. This further complicates results and may lead to relatively large biases on the extremes when using only limited time series. In addition to long simulations, the recent advances in the simulation of low-frequency variations with Bartlett-Lewis models (e.g., Kim and Onof (2020); Park et al. (2019)) should further improve the extremes. The problem of representation of extreme values with Bartlett-Lewis models, either of precipitation or discharge, has often been discussed (Verhoest et al., 2010). For a long time, many studies have focused on improving the calibration (see e.g. Onof and Wang (2020) for a recent example). However, during the past few years, the extension of the Bartlett-Lewis model with additional steps for monthly to seasonal precipitation simulation has also proven to be fruitful (Kim & Onof, 2020; Park et al., 2019). While the performance of the Bartlett-Lewis models is good at the hourly to daily scale, these additional steps allow for more variation on longer time scales and thus more varied extremes. Although it is a seemingly obvious suggestion, simulation of longer time series and extending the model chain might be better solutions for climate change impact assessment than using longer input time series of e.g., 50 years instead of only 30 years. Under climate change, precipitation statistics might change too much if a too long time series is considered. Introducing longer time series for climate change impact assessment might thus actually negatively affect the calibration, as they can no longer be considered representative of the climate in the strict meteorological sense.

The results in this study are promising and could open up many possibilities, but it is important to remember that they are all conditional on the input data, i.e., the observations and the bias-adjusted climate simulations. If not done properly, bias adjustment can introduce additional uncertainties and biases (Ehret, Zehe, Wulfmeyer, Warrach-Sagi, & Liebert, 2012). For example, although Van de Velde et al. (2022) warned that non-stationarity can influence the bias adjustment of the simulated future time series, this is not considered in this paper: only the reproduction of the values is of importance. In addition, inflation could cause an unrealistic level of variability

at the station level (Maraun, 2013). Thus, when using this weather generator set-up for effective climate change impact assessment studies, these uncertainties should be taken into account. Both the reproduction of the seasonal cycle and the correct adjustment of the variables could have a large impact on the final result of the generator. Besides the role of the bias adjustment, only one climate model and location was used in this paper, whereas for impact assessment a model ensemble should be considered and multiple locations should be assessed, although this ultimately depends on the goal of the impact assessment. In this situation, statistical downscaling (Maraun, Widmann, & Gutiérrez, 2019) should also be applied to bridge the gap between the climate model output and the station level preferred for Bartlett-Lewis calibration. However, it could also be worthwhile to integrate the approach considered here with already existing spatio-temporal Bartlett-Lewis models (Aryal & Jones, 2021), which could be calibrated more directly on climate model output. Finally, in the context of climate impact assessment, one should take into account that nonstationarity of the variables might be more problematic than anticipated here. The traditional approach of calibrating a Bartlett-Lewis model (and subsequent models) on calendar months might thus not suffice, and other approaches, such as calibration on temperature bins Cross et al. (2020), could be further investigated.

In summary, the results showed that the generator performs almost as well for the future simulations as for the historical observations for most statistical aspects. The largest biases were present for the third central moment of precipitation, leading to underestimated extremes. This illustrates the limitations of a 30-year time series. As using longer time series might have a negative impact, it will be more advantageous to apply other recent advances in Bartlett-Lewis modelling (e.g. Kim and Onof (2020)) and use these to generate sufficiently long time series. Nevertheless, the adequate results for most statistical moments imply that the weather generator can already be transferred to future conditions.

Acknowledgements

We thank the two anonymous referees for their insightful suggestions.

Disclosure statement

The authors declare that they have no conflict of interest.

Funding

This work was funded by FWO, grant number G.0039.18N. M. Demuzere is supported by the ENLIGHT project, funded by the German Research Foundation (DFG) under grant No. 437467569.

Data availability

The authors are grateful to the RMI for allowing the use of 117-year Uccle dataset. The code used to write this paper is available at github (<https://github.com/h->

cel/StochGen/). For the vine copula fitting, the R packages (R Core Team, 2021) VineCopula (Nagler et al., 2021) and R.Matlab (Bengtsson, 2018) were used.

Appendices

A: Bartlett-Lewis calibration results

Table A1.: Calibrated RBL2 parameter values.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
λ												
	Hist100	0.017	0.011	0.011	0.014	0.015	0.015	0.016	0.014	0.013	0.012	0.016
	Hist30	0.013	0.010	0.011	0.013	0.012	0.013	0.010	0.010	0.010	0.010	0.013
	Fut30	0.020	0.010	0.011	0.015	0.020	0.011	0.010	0.013	0.012	0.010	0.016
ν												
	Hist100	0.357	1.766	0	0	4.561	3.164	0	3.350	0	5.380	3.382
	Hist30	0	0	0	6.454	9.575	3.483	0	2.459	1.265	5.785	1.001
	Fut30	3.245	0	4.520	0	0	2.200	0	4.040	4.515	0	0.001
α												
	Hist100	1.989	2.020	2.029	2.021	1.907	1.727	1.962	2.056	2.017	2.018	2.050
	Hist30	2.047	2.030	2.027	1.967	1.841	1.769	2.119	2.064	2.051	1.967	2.0312
	Fut30	1.917	2.018	1.920	2.126	1.567	1.543	2.218	2.129	1.714	1.930	1.999
ι												
	Hist100	1.057	1.581	1.658	1.053	1.870	2.297	1.870	2.871	1.279	2.028	1.932
	Hist30	1.762	1.457	1.836	1.982	2.037	2.833	2.216	2.059	1.543	1.894	1.664
	Fut30	2.027	2.263	2.522	1.695	1.313	3.531	2.242	2.705	2.979	2.462	1.968
ϕ												
	Hist100	0.006	0.011	0	0	0.130	0.075	0	0.034	0	0.076	0.022
	Hist30	0	0	0	0.119	0.897	0.038	0	0.030	0.012	0.060	0.005
	Fut30	0.092	0	0.031	0	0	0.032	0	0	0.07	0.077	0
κ												
	Hist100	0.025	0.042	0	0	0.259	0.152	0	0.073	0	0.076	0.022
	Hist30	0	0	0	0.224	1.654	0.080	0	0.096	0.059	0.241	0.027
	Fut30	0.153	0	0.068	0	0	0.047	0	0	0.096	0.135	0
ω												
	Hist100	1.072	1.865	2.393	1.404	1.433	2.698	1.037	3.202	0.741	0.838	1.174
	Hist30	3.027	1.969	2.758	5.076	1.445	20.000	1.857	0.451	0.670	0.750	1.182
	Fut30	2.530	8.715	4.787	1.801	4.934	18.188	1.430	3.380	1.704	1.971	1.966

B: Vine copula calibration results

Table B1.: Fitted copulas for $V_{T_p,PT}$ for 100 years of historical data. F: Frank copula, N: Gaussian copula, C: Clayton copula.

Month	Copula		
	T_p, P	T_p, T	$P, T; T_p$
Jan	F	N	F
Feb	F	N	F
Mar	C	N	F
Apr	F	N	F
May	F	N	F
Jun	F	N	F
Jul	F	N	F
Aug	F	N	F
Sep	N	N	F
Oct	C	N	F
Nov	F	N	N
Dec	F	N	N

Table B2.: Fitted copulas for $V_{T_p,PT}$ for 30 years of historical data. F: Frank copula, N: Gaussian copula, C: Clayton copula.

Month	Copula		
	T_p, P	T_p, T	$P, T; T_p$
Jan	F	C	N
Feb	F	F	F
Mar	N	F	F
Apr	F	N	F
May	F	N	F
Jun	F	N	F
Jul	F	F	F
Aug	F	F	F
Sep	N	F	N
Oct	N	N	N
Nov	C	F	N
Dec	F	F	N

Table B3.: Fitted copulas for $V_{T_p,PT}$ for 30 years of future data. F: Frank copula, N: Gaussian copula, C: Clayton copula, G: Gumbel copula.

Month	Copula		
	T_p, P	T_p, T	$P, T; T_p$
Jan	C	C	F
Feb	C	F	F
Mar	C	N	G
Apr	G	F	F
May	F	N	F
Jun	F	N	F
Jul	F	N	F
Aug	F	N	F
Sep	F	N	F
Oct	C	N	G
Nov	N	N	F
Dec	F	F	N

-

Table B4.: Fitted copulas for V_{TPE_pE} for 100 years of historical data. F: Frank copula, N: Gaussian copula, C: Clayton copula, G: Gumbel copula.

Month	Copula					
	T, P	T, E_p	T, E	$P, E_p; T$	$P, E; T$	$E_p, E; T, P$
Jan	F	F	F	N	F	F
Feb	F	F	F	F	F	C
Mar	F	N	N	F	F	G
Apr	F	G	G	F	F	G
May	F	N	N	F	F	G
Jun	F	N	N	F	F	G
Jul	F	N	N	F	F	G
Aug	F	G	G	F	F	G
Sep	F	N	G	N	F	G
Oct	N	G	G	F	F	F
Nov	C	F	G	N	N	F
Dec	F	F	F	N	N	F

Table B5.: Fitted copulas for V_{TPE_pE} for 30 years of historical data. F: Frank copula, N: Gaussian copula, C: Clayton copula, G: Gumbel copula.

Month	Copula					
	T, P	T, E_p	T, E	$P, E_p; T$	$P, E; T$	$E_p, E; T, P$
Jan	F	F	F	F	F	F
Feb	F	F	C	N	N	C
Mar	F	N	N	F	N	N
Apr	F	G	N	F	F	N
May	F	G	N	F	N	G
Jun	F	F	F	F	N	F
Jul	F	F	F	F	N	N
Aug	F	F	F	F	N	N
Sep	N	F	N	F	F	F
Oct	N	N	N	F	F	N
Nov	C	C	C	F	F	F
Dec	F	F	F	N	N	F

Table B6.: Fitted copulas for V_{TPE_pE} for 30 years of future data. F: Frank copula, N: Gaussian copula, C: Clayton copula, G: Gumbel copula.

Month	Copula					
	T, P	T, E_p	T, E	$P, E_p; T$	$P, E; T$	$E_p, E; T, P$
Jan	F	F	F	N	N	G
Feb	F	N	G	G	N	G
Mar	C	G	G	F	F	N
Apr	F	G	G	F	F	N
May	F	F	G	F	F	N
Jun	F	N	G	N	F	N
Jul	F	F	G	F	N	N
Aug	F	F	G	F	N	N
Sep	F	F	G	F	F	N
Oct	G	G	G	N	F	F
Nov	N	F	F	F	N	F
Dec	F	F	F	F	F	F

References

- Aas, K. (2016). Pair-copula constructions for financial applications: A review. *Econometrics*.
- Aas, K., & Berg, D. (2009). Models for construction of multivariate dependence: A comparison study. *The European Journal of Finance*.
- Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2), 182–198.
- Acar, E. F., Czado, C., & Lysy, M. (2019). Flexible dynamic vine copula models for multivariate time series data. *Econometrics and Statistics*, 12, 181–197.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd Intr. Symp. on Information Theory, Budapest*. Akademiai Kiado.
- Aryal, N. R., & Jones, O. D. (2021). Spatial-temporal rainfall models based on poisson cluster processes. *Stochastic Environmental Research and Risk Assessment*.
- Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32(1-4), 245–268.
- Bedford, T., & Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*.
- Bengtsson, H. (2018). R.matlab: Read and Write MAT Files and Call MATLAB from Within R [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=R.matlab> (R package version 3.6.2)
- Bertrand, C., Ingels, R., & Journée, M. (2021). Homogenization and trends analysis of the Belgian historical precipitation time series. *International Journal of Climatology*.
- Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*.
- Bevacqua, E., Shepherd, T. G., Watson, P. A. G., Sparrow, S., Wallom, D., & Mitchell, D. (2021). Larger spatial footprint of wintertime total precipitation extremes in a warmer climate. *Geophysical Research Letters*, 48(8), e2020GL091990.
- Bevacqua, E., Zappa, G., & Shepherd, T. G. (2020). Shorter cyclone clusters modulate changes in European wintertime precipitation extremes. *Environmental Research Letters*, 15(12), 124005.
- Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. P. (2021). Challenges in modeling and predicting floods and droughts: A review. *Wiley Interdisciplinary Reviews: Water*.
- Burlando, P., & Rosso, R. (1993). Stochastic models of temporal rainfall: reproducibility, estimation and prediction of extreme events. In *Stochastic hydrology and its use in water resources systems simulation and optimization* (pp. 137–173). Springer.
- Burton, A., Fowler, H. J., Blenkinsop, S., & Kilsby, C. G. (2010). Downscaling transient climate change using a Neyman–Scott Rectangular Pulses stochastic rainfall model. *Journal of Hydrology*.
- Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *Journal of Climate*, 28(17), 6938–6959.
- Cross, D., Onof, C., & Winter, H. (2020). Ensemble estimation of future rainfall extremes with temperature dependent censored simulation. *Advances in Water Resources*, 136, 103479.
- Czado, C. (2019). *Analyzing dependent data with vine copulas*. Springer.
- De Jongh, I. L. M., Verhoest, N. E. C., & De Troch, F. P. (2006). Analysis of a 105-year time series of precipitation observed at Uccle, Belgium. *International Journal of Climatology*, 26(14), 2023–2039.
- Demarée, G. R. (2003). The centennial recording raingauge of the Uccle Plateau: Its history, its data and its applications. *Houille Blanche*, 4, 95–102.
- Diestel, R. (2018). *Graph Theory. Fifth. vol. 173*. Springer, Berlin.
- Dißmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data*

Analysis.

- Duan, Q. Y., Gupta, V. K., & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76(3), 501–521.
- Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., & Liebert, J. (2012). HESS Opinions” Should we apply bias correction to global and regional climate model data?”. *Hydrology and Earth System Sciences*, 16(9), 3391–3404.
- Fatichi, S., Ivanov, V. Y., & Caporali, E. (2011). Simulation of future climate scenarios with a weather generator. *Advances in Water Resources*, 34(4), 448–467.
- François, B., Schlef, K. E., Wi, S., & Brown, C. M. (2019). Design considerations for riverine floods in a changing climate—a review. *Journal of Hydrology*, 574, 557–573.
- Größer, J., & Okhrin, O. (2021). Copulae: An overview and recent developments. *Wiley Interdisciplinary Reviews: Computational Statistics*.
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., . . . Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9), 816.
- Hobaek Haff, I., Frigessi, A., & Maraun, D. (2015). How well do regional climate models simulate the spatial dependence of precipitation? An application of pair-copula constructions. *Journal of Geophysical Research: Atmospheres*, 120(7), 2624–2646.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., . . . Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional Environmental Change*.
- Jesus, J., & Chandler, R. E. (2011). Estimating functions and the generalized method of moments. *Interface Focus*.
- Joe, H. (1996). Families of m -variate distributions with given margins and $m(m - 1)/2$ bivariate dependence parameters. In *Distributions with Fixed Marginals and Related Topics* (pp. 120–141). Institute of Mathematical Statistics.
- Joe, H. (2014). *Dependence Modeling with Copulas*. CRC Press.
- Kaczmarek, J., Isham, V., & Onof, C. (2014). Point process models for fine-resolution rainfall. *Hydrological Sciences Journal*, 59(11), 1972–1991.
- Kahraman, A., Kendon, E. J., Chan, S., & Fowler, H. J. (2021). Quasi-stationary intense rainstorms spread across Europe under climate change. *Geophysical Research Letters*.
- Kilsby, C. G., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpham, C., . . . Wilby, R. L. (2007). A daily weather generator for use in climate change studies. *Environmental Modelling & Software*.
- Kim, D., & Onof, C. (2020). A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade. *Journal of Hydrology*.
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*.
- Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1. 0). *Geoscientific Model Development*, 12(7), 3055–3070.
- Maraun, D. (2013). Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate*, 26(6), 2137–2143.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., . . . Thiele-Eich, I. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*.
- Maraun, D., Widmann, M., & Gutiérrez, J. M. (2019). Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. *International Journal of Climatology*, 39(9), 3692–3703. Retrieved from <https://rmetsonline.wiley.com/doi/abs/10.1002/joc.5877>
- Moore, R. J. (2007). The PDM rainfall-runoff model. *Hydrology and Earth System Sciences*.
- Nagler, T., Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., & Erhardt, T. (2021). VineCopula: Statistical Inference of Vine Copulas [Computer software manual]. (R package

- version 2.4.2)
- Nelsen, R. B. (2006). *An Introduction to Copulas, Second Edition*. Springer Science+Business Media, New York.
- Onof, C., & Arnbjerg-Nielsen, K. (2009). Quantification of anticipated future changes in high resolution design rainfall for urban areas. *Atmospheric Research*, *92*(3), 350–363.
- Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheeler, H. S., & Isham, V. (2000). Rainfall modelling using Poisson-cluster processes: a review of developments. *Stochastic Environmental Research and Risk Assessment*, *14*(6), 384–411.
- Onof, C., & Wang, L.-P. (2020). Modelling rainfall with a Bartlett–Lewis process: new developments. *Hydrology and Earth System Sciences*, *24*(5), 2791–2791.
- Park, J., Onof, C., & Kim, D. (2019). A hybrid stochastic rainfall model that reproduces some important rainfall characteristics at hourly to yearly timescales. *Hydrology and Earth System Sciences*, *23*(2), 989–1014.
- Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007). Evaluation of the AR4 climate models’ simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of climate*, *20*(17), 4356–4376.
- Pham, M. T., Vernieuwe, H., De Baets, B., Willems, P., & Verhoest, N. E. C. (2016). Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas. *Stochastic Environmental Research and Risk Assessment*, *30*(8), 2197–2214.
- Pham, M. T., Vernieuwe, H., De Baets, B., & Verhoest, N. E. C. (2018). A coupled stochastic rainfall–evapotranspiration model for hydrological impact analysis. *Hydrology and Earth System Sciences*, *22*(2), 1263–1283.
- Popke, D., Stevens, B., & Voigt, A. (2013). Climate and climate change in a radiative-convective equilibrium version of ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*(1), 1–14.
- R Core Team. (2021). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond. A*.
- Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1988). A point process model for rainfall: further developments. *Proc. R. Soc. Lond. A*, *417*(1853), 283–298.
- Rodriguez-Iturbe, I., De Power, B. F., & Valdes, J. B. (1987). Rectangular pulses point process models for rainfall: analysis of empirical data. *Journal of Geophysical Research: Atmospheres*, *92*(D8), 9645–9656.
- Schepsmeier, U. (2015). Efficient information based goodness-of-fit tests for vine copula models with fixed margins: a comprehensive review. *Journal of Multivariate Analysis*, *138*, 34–52.
- Schepsmeier, U. (2019). A goodness-of-fit test for regular vine copula models. *Econometric Reviews*, *38*(1), 25–46.
- Schölzel, C., & Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research-introduction to the copula approach. *Nonlinear Processes in Geophysics*, *15*, 761–772.
- Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, *8*, 229–231.
- Strandberg, G., Barring, L., Hansson, U., Jansson, C., Jones, C., Kjellström, E., . . . Ullerstig, A. (2015). *CORDEX scenarios for Europe from the Rossby Centre regional climate model RCA4* (Tech. Rep.). SMHI.
- Sun, C., Huang, G., Fan, Y., Zhou, X., Lu, C., & Wang, X. (2021). Vine copula ensemble downscaling for precipitation projection over the Loess Plateau based on high-resolution multi-RCM outputs. *Water Resources Research*, *57*(1), 2020WR027698.
- Tootoonchi, F., Sadegh, M., Haerter, J. O., Rätty, O., Grabs, T., & Teutschbein, C. (2022). Copulas for hydroclimatic analysis: A practice-oriented overview. *Wiley Interdisciplinary Reviews: Water*.
- Trewin, B. C. (2007). *The role of climatological normals in a changing climate*. World

- Meteorological Organization.
- Vandenbergh, S., Verhoest, N. E. C., Onof, C., & De Baets, B. (2011). A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall. *Water Resources Research*, *47*(7), W07529.
- Van de Velde, J., Demuzere, M., De Baets, B., & Verhoest, N. E. C. (2022). Impact of bias nonstationarity on the performance of uni- and multivariate bias-adjusting methods: a case study on data from Uccle, Belgium. *Hydrology and Earth System Sciences*, *26*(9), 2319–2344.
- Vanhaute, W., Vandenbergh, S., Scheerlinck, K., De Baets, B., & Verhoest, N. E. C. (2012). Calibration of the modified Bartlett-Lewis model using global optimization techniques and alternative objective functions. *Hydrology and Earth System Sciences*.
- Verhoest, N. E. C., Troch, P. A., & De Troch, F. P. (1997). On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point. *Journal of Hydrology*, *202*(1-4), 108–120.
- Verhoest, N. E. C., Vandenbergh, S., Cabus, P., Onof, C., Meca-Figueras, T., & Jameleddine, S. (2010). Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrological Processes*, *24*(23), 3439–3445.
- Vernieuwe, H., Vandenbergh, S., De Baets, B., & Verhoest, N. E. C. (2015). A continuous rainfall model based on vine copulas. *Hydrology and Earth System Sciences*, *19*(6), 2685–2699.
- Welch, B. L. (1947). The generalization of student's problem when several different population variances are involved. *Biometrika*, *34*(1/2), 28–35.
- Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, *23*(3), 329–357.
- Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., ... Vignotto, E. (2020). A typology of compound weather and climate events. *Nature Reviews Earth & Environment*.
- Zscheischler, J., Westra, S., Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., ... Zhang, X. (2018). Future climate risk from compound events. *Nature Climate Change*, *1*.