



Stochastic modelling
of spatially
consistent daily
precipitation
time-series

D. E. Keller et al.

Stochastic modelling of spatially and temporally consistent daily precipitation time-series over complex topography

D. E. Keller^{1,2,3}, A. M. Fischer¹, C. Frei¹, M. A. Liniger^{1,3}, C. Appenzeller^{1,3}, and R. Knutti^{2,3}

¹Federal Office of Meteorology and Climatology MeteoSwiss, Operation Center 1, 8085 Zürich-Flughafen, Switzerland

²Climate Physics, Institute for Atmospheric and Climate Science, ETH Zürich, Universitaetstrasse 16, 8092 Zurich, Switzerland

³Center for Climate Systems Modeling (C2SM), ETH Zurich, Universitaetstrasse 16, 8092 Zurich, Switzerland

Received: 27 June 2014 – Accepted: 15 July 2014 – Published: 28 July 2014

Correspondence to: D. E. Keller (denise.keller@meteoswiss.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Many climate impact assessments over topographically complex terrain require high-resolution precipitation time-series that have a spatio-temporal correlation structure consistent with observations. This consistency is essential for spatially distributed modelling of processes with non-linear responses to precipitation input (e.g. soil water and river runoff modelling). In this regard, weather generators (WGs) designed and calibrated for multiple sites are an appealing technique to stochastically simulate time-series that approximate the observed temporal and spatial dependencies. In this study, we present a stochastic multi-site precipitation generator and validate it over the hydrological catchment *Thur* in the Swiss Alps. The model consists of several Richardson-type WGs that are run with correlated random number streams reflecting the observed correlation structure among all possible station pairs. A first-order two-state Markov process simulates intermittence of daily precipitation, while precipitation amounts are simulated from a mixture model of two exponential distributions. The model is calibrated separately for each month over the time-period 1961–2011.

The WG is skilful at individual sites in representing the annual cycle of the precipitation statistics, such as mean wet day frequency and intensity as well as monthly precipitation sums. It reproduces realistically the multi-day statistics such as the frequencies of dry and wet spell lengths and precipitation sums over consecutive wet days. Substantial added value is demonstrated in simulating daily areal precipitation sums in comparison to multiple WGs that lack the spatial dependency in the stochastic process: the multi-site WG is capable to capture about 95 % of the observed variability in daily area sums, while the summed time-series from multiple single-site WGs only explains about 13 %. Limitation of the WG have been detected in reproducing observed variability from year to year, a component that has not been considered in the WG calibration. Given the obtained performance, the presented stochastic model is expected to be a useful tool to re-sample the observed record and valuable to be used as a statistical downscaling method in a climate change context.

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

1 Introduction

In Switzerland, precipitation is a key weather variable with high relevance for sectors such as energy production, infrastructure, tourism, security, agriculture and ecosystems. Owing to a complex topography daily precipitation varies strongly in space and in time (Frei and Schär, 1998; Isotta et al., 2013). The spatial distribution of daily precipitation frequency and intensity clearly depends on the topography, with higher frequencies and intensities along the North-Alpine ridge during summer, and a strong north–south gradient with heavier intensities in southern Switzerland from spring until autumn. The most prominent weather situations causing these precipitation patterns are shallow pressure systems favouring convective precipitation, orographically induced precipitation (e.g. Föhn situations), and frontal passages. Precipitation amounts and frequencies are typically largest in summer, mainly due to convective processes (Frei and Schär, 1998).

To effectively assess the impacts related to precipitation events, often highly localized daily data are needed that are ideally both consistent in time and in space (Köplin et al., 2010). The observational record, however, is rather short for a broad impact assessment. Different sequences of daily precipitation would have been well possible arising from the chaotic nature of the weather system and its large spatio-temporal variability over Switzerland. Statistical models can help to explore many such evolutions by mimicking observed daily precipitation. Generating synthetic data is also highly relevant in the context of downscaling climate predictions and climate change projections. This is because daily data from climate models are not sufficiently reliable to be used without bias correction and hence the information often comes in form of aggregated quantities (e.g. regional and seasonal averages).

To accommodate these needs, stochastic simulations by weather generators (WGs) are an appealing technique, due to their computational cheapness and their simplicity relative to full climate models. WGs simulate synthetic weather time-series of practically unlimited length that statistically resemble the observed weather records in terms

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



of aggregated quantities (monthly means), frequency distribution, and temporal correlations. The individual time-series are hence multiple plausible realizations of daily weather sequences without a concrete day-by-day match to observations (Wilks and Wilby, 1999).

To date, a vast number of multi-variate WGs exist that simulate precipitation as a primary variable at single locations. These single-site WGs can be grouped into four main classes: (a) Richardson-type WG (Dubrovsky, 1997; Richardson and Wright, 1984; Richardson, 1981), that is the basis for this study here, (b) serial model using semi-empirical distributions to simulate the lengths of wet and dry day series and daily precipitation (Racsco et al., 1991; Semenov and Barrow, 1997), (c) non-parametric resampling from observed weather variables at the day of interest (Lall and Sharma, 1996; Rajagopalan and Lall, 1999), and (d) Neymann–Scott models that simulate precipitation as a sequence of storms with associated rainfall cell numbers based on a Poisson-cluster process (Rodríguez-Iturbe et al., 1987). All single-site WGs have in common that they do not incorporate the spatial dependencies when simulating concurrently at multiple sites. As a result, the areal-mean time-series have too small variance. This is a major drawback for many of the above mentioned applications.

Therefore, multi-site WGs have been developed that explicitly take into account the spatial-temporal correlation structure between different sites. Several attempts have been suggested to incorporate the spatial-dependency structure in a stochastic simulation process: (a) multi-site extension of single-site Richardson-type WGs by using correlated random number streams (Baigorria and Jones, 2010; Brissette et al., 2007; Srikanthan and Pegram, 2009; Wilks, 1998) that is also the basis for this study here, (b) non-homogeneous hidden Markov models that are commonly based on transition probabilities between pre-defined precipitation states over a domain (Mehrotra and Sharma, 2006; Zucchini and Guttorp, 1991), (c) nearest-neighbour resampling which implicitly take into account the spatial dimension by weather resampling at multiple sites simultaneously (Buishand and Brandsma, 2001; Leander and Buishand, 2009) and (d) the spatial extension of the rainfall models based on Poisson-Cluster processes

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



(Cowpertwait, 1995). A conceptually different space-time model for rainfall is the string-of-beads model developed by Pegram et al. (2001).

Each of these multi-site WGs comes with method-specific benefits and limitations for the reproduction of the daily precipitation statistics and consequently the driving of impact models. For instance, some of them do better in simulating more realistically longer-term variability (e.g. GLM-based multi-site WGs Chandler, 2014), some are explicitly adapted to deal with extreme precipitation (e.g. Maraun et al., 2009). Depending on the envisaged application, some of these aspects are more important than others. For instance, for agricultural studies a correct reproduction of dry and wet spell lengths is crucial (e.g. Calanca, 2007), while areal sums of precipitation are especially relevant for hydrological impacts. Another selection-criteria is the degree of complexity among the WG models and associated calibration requirements for the setup. For an exhaustive review on stochastic generation of precipitation we refer to Wilks and Wilby (1999) or Maraun et al. (2010).

In this study here, we develop, implement and evaluate a multi-site precipitation generator, that is based on a Richardson-type WG and run with correlated random number streams. The basic concept has been described in Wilks (1998). We deliberately choose a rather simple multi-site precipitation model, in order to be flexible in various aspects: (a) the WG should be easily adjustable for different catchment sizes and catchment locations regarding calibration setup, (b) the WG should be built so that it can be easily perturbed in a climate change context (as e.g. in Hirschi et al., 2012), (c) the output of the WG should be the base for a wide range of subsequent impact studies, which is why we choose a rather general method that shows on average, for several aspects of the daily precipitation statistics, a reasonable performance. Finally, we keep it as a simple toolbox, as one of our aims is to better understand uncertainties on the local scale, for which we treat the WG as toy model.

To date, only a few multi-site WGs have been tested over complex topographies. The multi-site WG of Wilks (1998) was employed over part of the Rocky Mountains (Wilks, 1999b), but never over the Alps which is the focus here. With its very high

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

spatio-temporal variability on local scales as well as the strong separation between southern and northern slopes, it is not guaranteed that the concept of Wilks (1998) is also applicable over the Alps. To test this, we apply and validate the generator over the alpine river catchment *Thur*. This catchment serves as an ideal testbed comprising 5 different precipitation regimes and steep elevation gradients. It also serves as a good test-domain to investigate, if and under what circumstances an added value of a multi-site WG against multiple single-site WGs is to be expected.

The structure of this paper is as follows: Sect. 2 introduces the hydrological catchment of the river *Thur* together with the used station data. In Sect. 3 we first describe 10 the statistical models for simulating precipitation occurrence and amount and show how these models are combined to multi-site simulation. The validity of our generated multi-site precipitation series and the comparison to single-site generators is presented in Sect. 4. We conclude the article by a summary and an outlook (Sect. 5).

2 Data

15 This study focuses on the hydrological catchment of the river *Thur*, which is located in the north-eastern part of Switzerland (Fig. 1a). The river *Thur* is a feeder river of the Rhine with a length of about 135 km and a catchment area of approximately 1696 km². It represents the largest Swiss river without a natural or artificial reservoir and therefore exhibits discharge fluctuations similar to unregulated Alpine rivers. Its flow regime is nivo-pluvial that is heavily influenced by snowmelt (BAFU, 2007). Owing to the complex 20 topography over this catchment area, precipitation exhibits a large variability both in space and in time. This is illustrated in Fig. 1b based on gridded observational data from Frei and Schär (1998). Over 1961–2011 and for a winter and summer month, the data clearly show larger precipitation frequencies and intensities over higher-elevated 25 regions compared to the lowlands. Additionally, this catchment lies in one of the Swiss regions featuring well above-average precipitation. A large portion of these precipitation

characteristics can be explained by a north-east to south-west lying mountain range (*Alpstein*) extracting precipitation from westerly flows and triggering convective storms.

For the purpose of our study here, we selected eight evenly distributed measurement stations (Fig. 1a) of MeteoSwiss that meet several requirements: (a) they all provide homogenized time-series covering a 51 year period from 1961–2011 (Begert et al., 2003), (b) they sufficiently reflect the elevation profile of the catchment area from *Andelfingen* lying at 382 m a.s.l. to *Saentis* lying at 2502 m a.s.l.

3 Method

Core of our multi-site WG is a Richardson-type precipitation generator (Richardson, 1981), that relies on the concept of modelling two processes at one single station: an occurrence and an amount process. Based on earlier work by Wilks (1998), this single-site WG is then extended in order to simultaneously generate precipitation at several sites taking into account the complex spatio-temporal correlation structure.

In the following, we explain the setup of our multi-site generator step by step: Sects. 3.1 and 3.2 present the concepts of statistically characterizing occurrence and amount at single-sites. The simulation procedure of new synthetic time-series is detailed in Sect. 3.3. In Sect. 3.4 we give a description of how we implemented the multi-site WG over the *Thur* catchment.

3.1 Precipitation occurrence process

Our procedure to model occurrence at a single station is based on the concept of a first-order two-state Markov chain (Gabriel and Neumann, 1962; Richardson, 1981). The first-order two-state Markov chain is a statistical model describing the probability to stay in the same state or switch to the other state. In this context, first-order implies the state at a given day depends only on the state at the previous day. The use of a first-order model in our WG was justified by inspecting the Bayesian information criterion

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[⏪](#)
[⏩](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

(Schwarz, 1978) revealing a better match compared to second- or higher order Markov models (not shown). We used a specific wet-day threshold of 1 mm day^{-1} to discretize a given daily precipitation time-series $\mathbf{X}(t)$ at a given site into the two states “dry” ($\mathbf{X}(t) < 1 \text{ mm day}^{-1}$) and “wet” ($\mathbf{X}(t) \geq 1 \text{ mm day}^{-1}$) and to dichotomise subsequently into a binary series (i.e. J_t with $J_t = 0$ for a dry state and $J_t = 1$ for a wet state). Four transitions are to be distinguished: a dry day following a dry day (00), a wet day following a dry day (01), a dry day following a wet day (10) and a wet day following a wet day (11).

Mathematically, the first-order two-state Markov chain model can be specified by formulating the probabilities (ρ) of these state-transitions:

$$\begin{aligned} \rho_{11} &= P\{J_t = 1 | J_{t-1} = 1\} \\ \rho_{01} &= P\{J_t = 1 | J_{t-1} = 0\} \end{aligned} \quad (1)$$

The corresponding counterparts of transition probabilities (ρ_{00} and ρ_{10}) can then be easily derived, since the sum of two probabilities conditioned on the same state at the previous day equals one:

$$\begin{aligned} \rho_{11} + \rho_{10} &= 1 \\ \rho_{00} + \rho_{01} &= 1 \end{aligned} \quad (2)$$

The two transition probabilities of Eq. (1) suffice to fully specify the first-order two-state Markov chain model. For the remaining part of this study, we therefore concentrate on these two parameters when addressing state transitions. For an estimate we rely on their conditional relative frequencies (Wilks, 2011):

$$\begin{aligned} \hat{\rho}_{01} &= \frac{n_{01}}{n_{0\bullet}} \\ \hat{\rho}_{11} &= \frac{n_{11}}{n_{1\bullet}} \end{aligned} \quad (3)$$

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



where n_{01} and n_{11} are the number of transitions from dry to wet and wet to wet in the binary series and $n_{0\bullet}$ and $n_{1\bullet}$ are the total number of zero's and one's in the series followed by any of the two states. From the transition probabilities of Eq. (3) other important precipitation indices can be inferred. The wet day frequency (wdf) is defined as the ratio of the number of wet days to the total number of days over a given time period. It can be expressed in terms of the two transition probabilities (Wilks, 2011):

$$\pi = \frac{\rho_{01}}{1 + \rho_{01} - \rho_{11}} \quad (4)$$

Similarly, the lag-1 autocorrelation r_1 is defined as the difference between the transition probabilities (Wilks, 2011):

$$r_1 = \rho_{11} - \rho_{01} \quad (5)$$

Since day-to-day precipitation generally exhibits positive serial correlation (i.e. r_1 greater than 0), ρ_{11} is usually larger than ρ_{01} and the wdf is between the two. Note, that a first-order two-state Markov chain does not imply independence for lags greater than one. The autocorrelation r_L decays exponentially with larger lags L :

$$r_L = (\rho_{11} - \rho_{01})^L \quad (6)$$

3.2 Precipitation amount process

As will be detailed in Sect. 3.3, at wet days, precipitation amounts are drawn from probability density functions (PDFs) fitted at single stations. Many studies use either an exponential (Richardson, 1981) or a gamma distribution (Buishand, 1978; Katz, 1977) to model non-zero precipitation amounts ($\mathbf{X}(t) \geq 1 \text{ mm day}^{-1}$). Both distribution types, however, do not appropriately characterize the frequency of the heavily right skewed precipitation amounts: they underestimate either light precipitation (exponential distribution) or heavy precipitation (gamma distribution). As an alternative, a mixture

model of two exponential distributions has been proposed to provide better overall fits and to better represent precipitation extremes (Wilks, 1999a). The PDF can be formulated as:

$$f(x) = \frac{w}{\beta_1} \exp\left(-\frac{x}{\beta_1}\right) + \frac{1-w}{\beta_2} \exp\left(-\frac{x}{\beta_2}\right) \quad (7)$$

- 5 $f(x)$ is a weighted average (weight w) of two exponential distributions with means β_1 and β_2 . Its quantile function exists in a closed form. Consequently, random samples from this distribution can easily be obtained by inversion (Wilks, 2011). The parameters w , β_1 and β_2 are estimated by using the concept of maximum-likelihood (Tallis and Light, 1968). Note that the estimation of PDF parameters is subject to sampling
- 10 uncertainty from the available number of wet days in a given calendar month.

3.3 Stochastic modelling of daily precipitation

3.3.1 Single-site

In this section, we demonstrate how the occurrence (Sect. 3.1) and amount model (Sect. 3.2) are applied to stochastically simulate daily precipitation at a single site.

15 The simulation process is based on Richardson (1981) with the five above-introduced parameters serving as input in Fig. 2: i.e. the transition probabilities p_{11} and p_{01} as well as w , β_1 and β_2 . The simulation of precipitation at a given day and a given station (say A) is accomplished in four main steps (see yellow circles in Fig. 2):

1. A random number $u_{t,A}$ is drawn from a standard normal distribution.
- 20 2. The conditional wet day probability p_A is determined depending on the state of the previous day. It is set to $p_{11,A}$ or $p_{01,A}$, depending on whether the previous simulated day was wet or dry, respectively.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

3. The random number $u_{t,A}$ is compared to the standard normal quantile function Q , evaluated at p_A : if $u_{t,A}$ is larger than $Q[p_A]$, a dry day ($J'_{t,A} = 0$) is simulated and else a wet day ($J'_{t,A} = 1$) is set.

4.1 In case of a dry day, the simulated amount $X'_{t,A}$ is set to zero.

5 4.2.1 In case of a wet day, a second random number $v_{t,A}$ (independent from $u_{t,A}$) is drawn from a standard normal distribution.

4.2.2 This random number is then substituted by the corresponding quantile (x_A) of the cumulative distribution function of the mixture model.

10 Note that this simulation procedure could be simplified by taking random uniform $[0, 1]$ numbers instead of Gaussian random numbers. We use the latter here in order to be consistent with the later introduced multi-site extension (Sect. 3.3.2).

15 Steps 1–4 are repeated over all remaining days within a certain simulation period. Based on this procedure time-series of arbitrary length can be generated that resemble observed climatological precipitation statistics, both in terms of frequency and intensity. For a more realistic reproduction of the annual cycle of precipitation the WG is calibrated on a monthly basis (see Sect. 3.4).

3.3.2 Multi-site

20 So far, the procedure to generate precipitation consists of multiple single-site WGs only. Specifically, no spatial dependence in the simultaneous simulation of precipitation at several sites was taken into account. To close this gap several single-site WGs are driven simultaneously with spatially correlated but serially independent random numbers (Wilks, 1998). For simplicity, the concept is illustrated in Fig. 2 for the example of two fictitious sites (A and B) only. The extension to several sites is straightforward. One of the main hurdles in simultaneously generating precipitation at several sites is the description of the spatial correlation matrices such that the dependence is also preserved

25

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



in the final generated time-series (Wilks and Wilby, 1999; Wilks, 1998). This difficulty mainly arises from the stochastic process that partly destroys the initially imposed correlation structure again (Wilks, 1998). We will come back to this problem later. For the moment, let us assume that the optimal correlation matrices for both, occurrence and amount (i.e. $\phi_{AB, \text{optim}}$ and $r_{AB, \text{optim}}$), are known. In this case, the main extensions to single-site WGs are two spatially correlated but serially independent random number streams (dashed boxes in Fig. 2): one for the occurrence (u) and the other for the amount (v) process. They are determined prior to the simulation process (see below) and contain the same number of days as the simulation period. Once these correlated random number streams are generated, the simulation proceeds as in Sect. 3.3.1 for all stations simultaneously. In practice, the multi-site WG implies the handling of three main methodological hurdles that are the following:

1. Calculating spatial correlation coefficients ϕ_{AB} and r_{AB}

Spatial dependence in binary series at site A and B is inferred by the phi-coefficient (ϕ_{AB}). Similarly as the Pearson correlation coefficient, the phi-coefficient ϕ_{AB} is bounded by -1 and 1 . For the precipitation amounts, the spatial correlation coefficient (r_{AB}) is determined by the conventional Pearson product-moment correlation coefficient. The correlation is calculated over the whole precipitation series that also include time-steps with zero amounts. From a statistical point of view, this is not an optimal procedure, since the correlation coefficients could be strongly affected by the number of zeros in the time-series. However, the purpose here is to use this spatial similarity measure rather as a tool to compare the observed spatial dependencies with those in artificial data. It is assumed that the statistical limitations in the calculation apply similarly to observations and generated data. The spatial correlations between different sites are determined pair-wise.

algorithm. The iteration is repeated until the generated correlation equals the one of observations with a precision of 0.005 (see Fig. S2). Note that this estimation procedure is done prior to the simulation and has to be repeated for each station pair and month.

3. Generation of correlated random number streams

There are several approaches to generate spatially correlated random numbers streams (e.g. Monahan, 2011). For the study at hand we used the concept of the Cholesky decomposition (e.g. Higham, 2009):

- (a) Sample for each station a random number stream from a standard Gaussian distribution.
- (b) Apply a Cholesky decomposition to the optimized correlation matrix to get a lower triangular matrix and its transposed.
- (c) Multiply the resulting lower triangular matrix with the matrix of random number streams.

Cholesky decomposition requires matrices that are positive definite, i.e. that contain no negative eigenvalues. However, in case of inter-station correlations this is not always fulfilled and depends on the number of stations with incomplete records. In absence of positive definite matrices, a fall-back solution based on the nearest positive correlation matrix was chosen. This problem occurred in our study only a few times. Note, that the temporal correlation structure of the precipitation time-series at one specific site is not altered by the imposed spatial correlation, since the spatially correlated random number streams exhibit no serial correlation.

**Stochastic modelling
of spatially
consistent daily
precipitation
time-series**

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



3.4 Implementation

3.4.1 Implementation of the multi-site WG over the Thur catchment

Our developed precipitation generator is calibrated on a monthly basis. First, all the single-site input parameters (p_{11} , p_{01} , β_1 , β_2 and w) were estimated for each of the 8 stations within the catchment and for each month separately using a time-window of 51 years (1961–2011). For the two transition probabilities in a given month, the climatological mean over the 51 yearly values of p_{11} and p_{01} was taken. In case of fitting a PDF to non-zero precipitation amounts and the estimation of β_1 , β_2 and w , we used the daily data over all 51 years together. In addition, a three-month window centred at the month of interest was chosen, in order to increase sample size and the robustness. The distributional parameters were derived based on maximum-likelihood (Tallis and Light, 1968). Despite our three-month time-window, cases occurred when the maximum-likelihood algorithm did not converge. For such cases, a fall back solution was applied where the parameter estimates from the previous month were adopted. With the monthly parameters from all the calibrated single-site WGs and the monthly observed inter-station correlations (symmetric correlation matrices), the optimized correlation matrices had to be found for each month based on the procedure described in Sect. 3.3.2. Note, that by calibrating the multi-site WG on a monthly instead of a seasonal basis, additional sampling uncertainty is introduced due to the rather small time-window to estimate our parameters. This is the downside of prescribing an improved annual cycle in the WG parameters.

Once the multi-site WG was calibrated, we generated 100 ensembles of daily time-series, of 51 year length. All the results presented in Sect. 4 are calculated over the time-period 1961–2011.

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

3.4.2 Reproduction and uncertainty of WG model parameters

To test whether our developed WG is properly implemented, we evaluated the reproduction of WG input parameters extracted from the generated time-series. A correct reproduction in parameters such as wet day intensity, frequency and transition probabilities is a prerequisite for all the subsequent analyses presented in Sect. 4. The evaluation was performed for four subjectively-defined climatic regimes: a very dry, a dry, a wet and a very wet climate. The corresponding model parameters are indicated in Fig. 3 with dashed vertical lines. For each of these precipitation regimes, 100 synthetic daily time-series were generated. To test the effect of sample-size, different sizes of time-windows were used: (a) 10 000 days, (b) 1000 days, (c) 100 days and (d) 30 days. The latter corresponds to the same sample-size as used to simulate monthly precipitation over the *Thur* catchment. For each of the generated time-series the WG parameters were re-estimated and the 95 % interquantile range was computed across the set of 100 realizations (Fig. 3). Three main results can be inferred: (a) our precipitation generator is able to correctly reproduce the key WG parameters implying that the chances for substantial coding errors are small, (b) as expected the estimate of the input parameters becomes more uncertain the smaller the sample size is; in fact, the uncertainty range enlarges by a factor of roughly 19 from a sample size of 10 000 down to 30. At a sample size of 1000 the uncertainty range stays at around ± 0.03 , that only marginally lowers when going to a sample of 10 000. (c) The different pre-defined climate regimes affect the uncertainty, particularly in the estimated transition probabilities. In a very dry (wet) climate, the wet–wet (dry–wet) transition probability exhibits large uncertainties in the estimate. This again is mainly related to a sample size problem due to very few wet–wet (dry–wet) pairs. Thus, we expect that the weather generator does not work optimally in arid climates.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



4 Results

An in-depth evaluation of the generated time-series with our multi-site WG is now undertaken with real observations. First, the reproduction of the daily and longer-term precipitation statistics at individual sites is analysed (Sect. 4.1). In a second step, the performance of the multi-site model is investigated regarding spatially aggregated precipitation indices in comparison to WGs without incorporating spatial dependencies (Sect. 4.2).

4.1 Validation of the precipitation generator at individual sites

Based on our ensemble of synthetic time-series, each containing 51 years, we analyse the reproduction of key precipitation characteristics. This validation goes beyond the reproduction of pure model parameters used to calibrate the WG (Sect. 3.4.2), as it includes precipitation statistics that are not directly used in the specification and calibration of the model. Note, that we present this analysis for the same time-period as used for calibrating our WG. This is justified for the study here, as long as we treat and use our WG to simulate long-term monthly precipitation statistics. In such a setup the stationarity of the model is given by definition. However, in a climate prediction or projection context, this stationarity assumption would have to be tested and hence separate calibration and validation periods are needed.

4.1.1 Long-term mean and inter-annual variance of monthly precipitation sums

In a first step of validating our WG, we focus on the reproduction of the long-term mean in monthly precipitation sums. Figure 4 shows both the modelled (blue) and observed (black) long-term monthly precipitation sum for each of the eight investigated stations. In general, the annual cycle of precipitation sums is well reproduced. Consistently, this is also true for the long-term seasonal as well as for the annual precipitation sums (not shown). But the WG tends to slightly underestimate precipitation sums in June

and August, and overestimate them in October. In addition, the two stations *Bischofszell* (BIZ) and *Herisau* (HES) show rather large positive deviations from the observed record during the winter months. In order to explain part of these deviations, we decomposed the long-term mean of monthly ($T = 30$ days) precipitation sums ($E[S(T)]$) into the product of the mean monthly wet day frequency (wdf) and intensity (wdi) (Fig. 5):

$$E[S(T)] = T \cdot \text{wdf} \cdot \text{wdi} \quad (8)$$

Since these two climatological quantities are indirectly forced (Sect. 3.4.2), we expect from the results in Fig. 3 a good match on average. As shown in Fig. 5, this is especially true for the wet day frequency, where the deviations between generated (red) and observed (black) values are relatively small. The differences, however, are more pronounced in case of mean wet day intensities. In fact, it is the wet day intensities that explain the mismatches in precipitation sums. In case of the winter performance over *Bischofszell* and *Herisau* the deviations can be attributed to the failure of converging in case of fitting the non-zero precipitation amount model. For those instances, the fallback solution had to be used (see Sect. 3.4.1).

Let us now focus on the inter-annual variability of monthly precipitation sums, which is often more difficult to realistically model than the long-term mean (Wilks and Wilby, 1999). The shaded areas in Fig. 4 represent the inter-quartile range of the observed (grey) and modelled (blue) monthly precipitation sums. From Fig. 4 it is obvious that the variability of the WG is smaller than in observations for all of the analysed stations. This implies that the stochastic model only explains part of the observed total variability. This reduced variability is expected, as observations are subject to additional sources of variability, which our comparable simple WG is not trained for. The WG is forced with mean observed values, varying between months but not between different years. The annual cycle is assumed to be stationary, and hence interannual variability, e.g. related to the North Atlantic Oscillation (Hurrell et al., 2003) is missing. Consequently, the ratio of simulated over observed variance accounts for approximately 33 % on average. The magnitude of this result is consistent with other studies (e.g. Gregory et al., 1993).

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Further insights can be gained from a decomposition of the variance of monthly ($T = 30$ days) precipitation sums ($\text{Var}[S(T)]$) into the variance of non-zero amount ($\text{Var}[X \geq 1 \text{ mm day}^{-1}]$) and the variance of the number of wet days ($\text{Var}[N(T)]$) as proposed by Wilks and Wilby (1999):

$$\text{Var}[S(T)] = T \cdot \text{wdf} \cdot \text{Var}[X \geq 1 \text{ mm d}^{-1}] + \text{Var}[N(T)] \cdot \text{wdi}^2 \quad (9)$$

Since the mean wet day frequency (wdf) and intensity (wdi) are reasonably reproduced, we expect that the reduced variability of monthly precipitation sums originate from deficiencies in correctly reproducing the inter-annual variability of the number of wet days and/or of the non-zero amount. One likely reason is the neglect of low-frequency variability in the WG parameters. It has been shown that physically based models that include large-scale circulation as a predictor could alleviate this problem (Chandler and Wheeler, 2002; Furrer and Katz, 2007; Wheeler et al., 2005; Yang et al., 2005).

4.1.2 Reproduction of PDF of daily non-zero amount

The adequate reproduction of the mean wet day intensity and frequency is a necessary but not sufficient precondition of a WG to be used for subsequent (impact) studies. Due to a large variability of precipitation amounts, it strongly matters how its frequency distribution is reproduced. For this, we compared simulated and observed quantiles of the daily non-zero precipitation distribution at each station (Fig. S3). Generally, the mixture model of two exponential distributions captures the frequencies of the intensities reasonably well, even at the high-Alpine station *Saentis* (SAE). This is at least the case up to the 80th percentile, above which intensities are systematically underestimated at all stations. This issue could be overcome by more sophisticated amount models combining e.g. a Gamma with a Generalized Pareto distribution (Vrac and Naveau, 2007). However, this comes at the expense of fitting many parameters with a limited sample size.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



4.1.3 Reproduction of multi-day statistics

While the frequencies of precipitation amounts and the frequencies of wet and dry days are realistically simulated, it remains unclear how the WG performs for multi-day spells. For many application studies, this is an essential information that requires a specific analysis. Figure 6 displays observed and modelled cumulative frequencies of dry and wet spells lengths at the example of two months and two stations. The two stations *Saentis* and *Andelfingen* are selected for display since they represent the stations with the highest and lowest elevation in the catchment. For both stations a clear seasonal difference in the probability of dry spells toward more short and less long dry spells during summer compared to winter is found. A plausible explanation are the more intermittent (convective) precipitation systems during summer. In contrast to dry spells, no seasonal differences in wet spell length probabilities can be inferred. This is likely related to the fact that the dry–dry transition probability p_{00} exhibits a more distinct annual cycle than the wet–wet transition probability p_{11} . Figure 6 also shows that the frequency at shorter spell lengths (up to 3 days) is more realistically reproduced by the model than the frequency at longer spell lengths. Generally, a better reproduction of wet spell probabilities is seen compared to the dry spell counterpart. Long dry spell lengths are more frequently underestimated by the model than longer wet spell lengths. The underestimation of long wet and dry spells is a common shortcoming of the Richardson-type weather generator and has been reported by many studies before (e.g. Racsko et al., 1991). This deficiency mainly arises due to the fast exponential decay of the autocorrelation function with larger lags (see Eq. 5). Similar to the underestimation of variability in precipitation sums, higher-order Markov chains (Wilks, 1999a) or GLMs with additional predictors might improve this aspect, which is out of scope in this study here.

Given that the frequency of wet spell lengths is realistically simulated, the question arises whether this also holds for multi-day precipitation sums. Multi-day periods of rain is a common phenomenon over Switzerland, especially during prevailing weather

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



situations that favour orographic uplift. We compared observed and simulated cumulative distribution functions (CDFs) of precipitation sums over multiple consecutive wet days (Fig. 7). Overall, we found that the differences between generated and observed time-series are largest for the higher quantiles and for long lasting wet spells (5 day wet spells) where the WG tends to underestimate large multi-day sums. This reduced skill in simulating longer wet spell sums can be explained by the fact that our WG is only prescribed with the temporal structure of precipitation occurrence but not in amount. In other words, the WG has memory to realistically reproduce multi-day wet spell lengths (Fig. 6), while the combined analysis of multi-day occurrence and accumulated amount loses somewhat this memory again. Two further noticeable features in Fig. 7 are that intense one-day precipitation sums are often overestimated by the model compared to the observations, while a relatively good match is obtained for three-day sums. Although the deficiency in correctly simulating multi-day sums of consecutive wet days is to be expected by construction of the WG, it could be improved by more sophisticated precipitation models, such as multi-states Markov-chains with different probability density distributions at each state (Buishand, 1978; Katz, 1977). This, however, comes at the expense of fitting many additional parameters with a limited sample size.

4.2 Performance of spatial precipitation indices

Up to this point we evaluated the generator at individual sites only. The key issue of this study though is the potential added value of incorporating inter-station dependencies. Similarly as in the previous section, we analyse the performance first in terms of occurrence-related statistics and second in terms of the combined occurrence and amount statistics.

4.2.1 Dry and wet spell statistics for the whole catchment

Based on the eight stations in our catchment with each being either in a wet or dry state at a given day, theoretically 2^8 (= 256) different dry-wet patterns in space are

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

possible. In observations, though, it turns out that 70 % of the investigated days over 1961–2011 are in fact either completely dry (45 %) or completely wet (25 %) and the remaining 254 dry–wet-patterns are subject to far smaller frequencies (around 10^{-5} – 10^{-3} %). The pre-dominance of a dry or a wet catchment makes sense given that the catchment is relatively small and given that precipitation is to a large degree circulation-triggered. Analysing the synthetic time-series from our multi-site WG reveals an almost perfect match with observations (Table 1), a consequence of prescribing the spatial dependency structure in the occurrence process. Indeed, when re-doing the same experiments with multiple single-site WGs without inter-site dependencies, only about 2 % of all days are completely dry in the catchment and none of the days are simulated as completely wet (Table 1). In a single-site WG setup, the chances for all stations being dry or wet ultimately depend on the calibrated wet day frequencies at the eight stations that remain below 0.5 in almost all months (see Fig. 5). This implies that the likelihood for dry conditions over the catchment is higher than for wet conditions.

Those days with complete dry or wet catchment conditions were further investigated in terms of the temporal structure. Table 1 presents observed and multi-site simulated spell length statistics for the catchment. In general, remarkably good agreement between observations and the multi-site model is found. This is also true for longer spell lengths, where the spatio-temporal correlation structure is only indirectly given as input to the WG. All of these results imply that our multi-site WG not only captures the frequencies of spatially aggregated binary series very well, it also does a surprisingly good job in reproducing multi-day dry/wet spells of the *Thur* catchment.

4.2.2 Daily non-zero precipitation sums over the catchment

The above findings on the spatio-temporal correlation structure in the occurrence process also give confidence that daily precipitation sums aggregated over the catchment are reasonably simulated. To answer this user-relevant question, we first analyse seasonal distributions of single-day precipitation area sums over the time-period 1961–2011 (Fig. 8). Area sums are defined as the precipitation sum over the eight

stations. Note, that days with an area sum of zero were excluded from this analysis and are not shown. The observations (grey boxplots) show in the median only a weak inter-seasonal variability with somewhat higher sums during summer. The spread in daily precipitation is smallest for winter and spring and largest for summer owing to the higher extreme precipitation values observed. Common to all seasons is a distribution that is heavily right-skewed ranging from nearly dry conditions up to about 220 mm day⁻¹. Note, that the spread shown here includes variability from year-to-year but also within the season of the same year.

Compared to observations, the multi-site generator reproduces well the median of the observed daily areal sums. The relative deviations remain rather small, ranging from -8.5 % in summer to +1.6 % in autumn. Moreover, the multi-site model is able to capture about 95 % of the observed variability in the daily sums, while the single-site WG only explains about 13 %. Even for extreme areal precipitation, the deficiencies are rather small. Contrary to a multi-site model, the areal sum derived from several single-site WGs over the catchment (red) systematically underestimates median, variability and consequently the magnitude of extreme precipitation amounts (Fig. 8). The relative deviations from observations in the median range from -28 % in autumn to -18 % in spring. The underestimation may be explained by the fact that the single-site model rarely simulates days where all stations are wet (Sect. 4.2.1). Also, the spatial structure of the precipitation amount is not accounted for.

4.2.3 Annual maximum precipitation sums of consecutive days over the catchment

The previous analysis has revealed a pronounced added value when incorporating spatial dependencies in the stochastic simulation of daily areal precipitation sums over the *Thur*.

Similarly to Sect. 4.2.1, we want to go a step beyond and additionally include the temporal structure. Note that by investigating spatial precipitation sums over multi-days, we explore the limits of our WG. We analyse in Fig. 9 annual maxima of observed (grey),

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



and modelled (blue and red for multi-site and single-site, respectively) precipitation sums over several consecutive days (2, 5, and 10 days). This means that out of the aggregated catchment-time-series we compute temporal sums over consecutive days and take the maximum in each year.

Regarding the performance of our WG in multi-site and single-site mode, Fig. 8 shows that both are clearly underestimating the observed sums. Yet, the multi-site model exhibits much smaller deviations from the observed distribution than the single-site model, and hence the added value of the multi-site WG is clearly evident. In fact, the sums simulated with the multi-site WG are larger by a factor of around 1.8 than those generated with the single-site WG. Overall, deviations from observations are reduced from about -53% (single-site WG) to about -17% (multi-site WG). The added value of the multi-site model is not constant for different consecutive sums. Differences are larger at shorter multi-day sums and decrease toward longer time-windows. This is related to the fact that the spatio-temporal correlation structure at longer lags is not prescribed in the model as already seen in Sect. 4.2.1 and Table 1. The benefit of a multi-site WG in terms of maximum daily areal precipitation sums is therefore restricted to consecutive sums over a few days only. And as a consequence for time-windows of 30 days (or monthly sums), a single-site WG performs equally good as a multi-site WG (not shown), as both models are calibrated for monthly sums at the eight stations and consequently at the catchment.

4.2.4 Discussion

The incorporation of inter-station dependencies in the stochastic model brings substantial added value over multiple single-site models regarding daily and multi-day areal precipitation sums over the *Thur* catchment. Similar benefits from the multi-site WG would be expected for other Alpine catchments and regions with complex topography, where correlations between sites are significant but well below unity. For very homogeneous regimes (inter-station correlation near unity) one single-site WG would be

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

sufficient for the catchment-area, whereas for low spatial correlations several independent single-site WGs can be used.

A stochastic simulation with multi-site correlation structure comes with additional uncertainty from parameter estimations, additional implementation complexity and additional computational costs. The decision for incorporating spatial dependencies must therefore be balanced with the benefit. A careful inspection of the observed precipitation regime and its spatial structure over the catchment prior to the simulation is necessary to decide in favour or against multi-site simulation.

5 Summary and outlook

A multi-site daily precipitation generator has been successfully developed, implemented and tested over the Swiss alpine river catchment *Thur*. The generator is built after suggestions by Wilks (1998). Core of our multi-site precipitation generator is a Richardson-type WG with simulation of daily precipitation occurrence as a chain dependent process and simulation of non-zero daily precipitation amounts from a mixture model of two exponential distributions. The spatial dependencies between the stations are imposed by running the precipitation models with spatially correlated but serially independent random numbers. The model was calibrated on a monthly basis by using daily station data over a 51 year long time-period from 1961–2011 and extensively inter-compared to the observed record and to simulations based on multiple (independent) single-site WGs.

Our main findings of this study are:

- Our developed multi-site precipitation generator realistically reproduces key precipitation statistics at single stations, including the annual cycle, quantiles of non-zero precipitation amounts, multi-day spells and multi-day amount statistics.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

– Based on its good performance in a range of spatio-temporal precipitation aspects, our weather generator is expected to serve as a helpful data provision tool for multiple applications including climate change assessments.

– The precipitation generator is able to generate relatively large stochastic variability. Nevertheless, it is rather low compared to observed inter-annual variability where it underestimates inter-annual variability by a factor of 3.

– The incorporation of inter-station dependencies in the stochastic process brings substantial added value over multiple single-site WGs over heterogeneous catchment areas such as the *Thur* catchment:

a. The median of daily area sums are by about a factor of 1.3 higher than those from independent single-site models. In addition, the multi-site WG is able to capture about 95 % of the observed variability, while the single-site WG only explains about 13 %.

b. Annual maxima of multi-day sums over the catchment increase by about a factor of 1.8 by incorporating the inter-site dependence in the stochastic simulations.

– The added value is expected to become most distinct when the precipitation regime is subject to a large spatial and temporal heterogeneity as is the case over the *Thur* catchment.

These results give us confidence that the developed precipitation generator is a very helpful tool to simulate current climate. Nevertheless, from an end-user perspective, some relevant limitations remain: the synthetically generated time-series do not capture the day-to-day and multi-day variability of precipitation to a full extent. Extreme values are hence underestimated and should not be the focus of any such analysis with the data at hand. Furthermore, our generator underestimates the year-to-year variability in monthly precipitation sums. This problem could be alleviated by sampling the input WG



parameters from the observed distribution instead of solely taking the best estimate. A more sophisticated way would be to use a model that incorporates large-scale atmospheric variables as predictors to estimate the WG parameters, such as for instance demonstrated by Furrer and Katz (2007) using Generalized linear models (GLMs).

In light of these inherent limitations, care should be taken when using the generated time-series as basis for a comprehensive risk assessment of different climatic impacts. To increase robustness in our results here, the generator should be ideally applied to further catchments of different sizes and in different time-periods. This would entail a better quantification of the benefits and limitations. In any case, the presented generator is subject to further developments, including the extension to a multi-variate weather generator and its adaptation for climate change studies. If proven skilful, it is planned to use the weather generator as a downscaling technique to simulate spatially and temporally consistent daily precipitation time-series at the local scale consistent with large-scale climate model projections of a future climate.

The Supplement related to this article is available online at doi:10.5194/hessd-11-8737-2014-supplement.

Acknowledgements. This work is supported by the ETH Research Grant CH2-01 11-1. We would like to thank the Center for Climate Systems Modeling (C2SM) at ETH Zurich for providing technical and scientific support.

References

- BAFU: Hydrologischer Atlas der Schweiz HADES, Hydrol. Atlas der Schweiz HADES, available at: <http://www.hydrologie.unibe.ch/hades/index.html> (last access: 10 April 2014), 2007.
- Baigorria, G. A. and Jones, J. W.: GiST: a stochastic model for generating spatially and temporally correlated daily rainfall data, *J. Climate*, 23, 5990–6008, doi:10.1175/2010JCLI3537.1, 2010.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Begert, M., Seiz, G., Schlegel, T., Musa, M., Baudraz, G., and Moesch, M.: Homogenisierung von Klimamessreihen der Schweiz und Bestimmung der Normwerte 1961–1990, Zurich, Me-teoSwiss, 2003.
- 5 Brissette, F. P., Khalili, M., and Leconte, R.: Efficient stochastic generation of multi-site synthetic precipitation data, *J. Hydrol.*, 345, 121–133, doi:10.1016/j.jhydrol.2007.06.035, 2007.
- Buishand, T. A.: Some remarks on the use of daily rainfall models, *J. Hydrol.*, 36, 295–308, 1978.
- Buishand, T. A. and Brandsma, T.: Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling distribution, *Water Resour. Res.*, 37, 2761–2776, doi:10.1029/2001WR000291, 2001.
- 10 Calanca, P.: Climate change and drought occurrence in the Alpine region: how severe are becoming the extremes?, *Global Planet. Change*, 57, 151–160, doi:10.1016/j.gloplacha.2006.11.001, 2007.
- Chandler, R. E.: Rglimclim, available at: <http://www.homepages.ucl.ac.uk/~ucakarc/work/glimclim.html> (last access: 10 April 2014), 2014.
- 15 Chandler, R. E. and Wheeler, H. S.: Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland, *Water Resour. Res.*, 38, 1192, doi:10.1029/2001WR000906, 2002.
- Cowpertwait, P. S. P.: A generalized spatial–temporal model of rainfall based on a clustered point process, *P. Roy. Soc. A-Math. Phys.*, 450, 163–175, doi:10.1098/rspa.1995.0077, 1995.
- 20 Dubrovsky, M.: Creating daily weather series with use of the weather generator, *Environmetrics*, 8, 409–424, 1997.
- Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rainn-gauge observations, *Int. J. Climatol.*, 18, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.
- 25 Furrer, E. M. and Katz, R. W.: Generalized linear modeling approach to stochastic weather generators, *Clim. Res.*, 34, 129–144, doi:10.3354/cr034129, 2007.
- Gabriel, K. R. R. and Neumann, Y.: A Markov chain model for daily rainfall occurrence at Tel Aviv, *Q. J. Roy. Meteor. Soc.*, 88, 90–95, 1962.
- 30 Gregory, J., Wigley, T. M., and Jones, P.: Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals, *Clim. Dynam.*, 8, 299–310, doi:10.1007/BF00209669, 1993.

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Monahan, J. F.: Numerical Methods of Statistics, Cambridge University Press, Cambridge, 2011.
- Pegram, G. G. S. and Clothier, A. N.: High resolution space–time modelling of rainfall: the “String of Beads” model, *J. Hydrol.*, 241, 26–41, doi:10.1016/S0022-1694(00)00373-5, 2001.
- 5 Racsco, P., Szeidl, L., and Semenov, M. A.: A serial approach to local stochastic weather models, *Ecol. Model.*, 57, 27–41, doi:10.1016/0304-3800(91)90053-4, 1991.
- Rajagopalan, B. and Lall, U.: A k -nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resour. Res.*, 35, 3089–3101, doi:10.1029/1999WR900028, 1999.
- Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resour. Res.*, 17, 182–190, doi:10.1029/WR017i001p00182, 1981.
- 10 Richardson, C. W. and Wright, D. A.: WGEN: a Model for Generating Daily Weather Variables, USDA-ARS, Washington, D.C., 1984.
- Rodríguez-Iturbe, I., de Power, B. F., and Valdés, J. B.: Rectangular pulses point process models for rainfall: analysis of empirical data, *J. Geophys. Res.*, 92, 9645, doi:10.1029/JD092iD08p09645, 1987.
- 15 Schwarz, G.: Estimating the dimension of a model, *Ann. Stat.*, 6, 461–464, doi:10.1214/aos/1176344136, 1978.
- Semenov, M. A. and Barrow, E. M.: Use of a stochastic weather generator in the development of climate change scenarios, *Clim. Change*, 35, 397–414, doi:10.1023/A:1005342632279, 1997.
- 20 Srikanthan, R. and Pegram, G. G. S.: A nested multisite daily rainfall stochastic generation model, *J. Hydrol.*, 371, 142–153, doi:10.1016/j.jhydrol.2009.03.025, 2009.
- Tallis, G. M. and Light, R.: The use of fractional moments for estimating the parameters of a mixed exponential distribution, *Technometrics*, 10, 161–175, doi:10.1080/00401706.1968.10490543, 1968.
- 25 Vrac, M. and Naveau, P.: Stochastic downscaling of precipitation: from dry events to heavy rainfalls, *Water Resour. Res.*, 43, W07402, doi:10.1029/2006WR005308, 2007.
- Wheater, H. S., Chandler, R. E., Onof, C. J., Isham, V. S., Bellone, E., Yang, C., Lekkas, D., Lourmas, G., and Segond, M.-L.: Spatial-temporal rainfall modelling for flood risk estimation, *Stoch. Env. Res. Risk A.*, 19, 403–416, doi:10.1007/s00477-005-0011-8, 2005.
- 30 Wilks, D. S.: Multisite generalization of a daily stochastic precipitation generation model, *J. Hydrol.*, 210, 178–191, doi:10.1016/S0022-1694(98)00186-3, 1998.

HESSD

11, 8737–8777, 2014

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Wilks, D. S.: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models, *Agr. Forest Meteorol.*, 93, 153–169, doi:10.1016/S0168-1923(98)00125-7, 1999a.

5 Wilks, D. S.: Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain, *Agr. Forest Meteorol.*, 96, 85–101, doi:10.1016/S0168-1923(99)00037-4, 1999b.

Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, Academic Press, 2011.

Wilks, D. S. and Wilby, R. L.: The weather generation game: a review of stochastic weather models, *Prog. Phys. Geogr.*, 23, 329–357, doi:10.1177/030913339902300302, 1999.

10 Yang, C., Chandler, R. E., Isham, V. S., and Wheeler, H. S.: Spatial–temporal rainfall simulation using generalized linear models, *Water Resour. Res.*, 41, W11415, doi:10.1029/2004WR003739, 2005.

Zucchini, W. and Guttorp, P.: A hidden Markov model for space–time precipitation, *Water Resour. Res.*, 27, 1917–1923, doi:10.1029/91WR01403, 1991.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Table 1. Frequencies (given in percent) of a completely wet or dry catchment together with the frequencies of its spell lengths. The observed (OBS) frequencies are calculated over 1961–2011. The multi-site simulated frequencies are given by the mean of 100 runs over 51 years (1961–2011).

		Wet catchment			Dry catchment		
		OBS	multi-site	single-site	OBS	multi-site	single-site
Overall frequency		25	25	0	45	44	2
Frequencies of spell lengths	1	34.8	34.4	0.0	14.1	17.3	2
	2	27.3	29.4	0.0	16.2	20.7	0.0
	3	16.7	18.2	0.0	13.0	18.2	0.0
	4	11.5	9.7	0.0	10.8	14.1	0.0
	5	4.1	4.7	0.0	9.1	10.3	0.0
	6	2.7	2.1	0.0	5.9	7.0	0.0
	7	0.9	0.9	0.0	7.2	4.7	0.0
	8	0.7	0.4	0.0	5.1	3.0	0.0
	9	0.6	0.2	0.0	3.5	1.9	0.0
	10	0.2	0.0	0.0	3.5	1.2	0.0

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[⏪](#)
[⏩](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

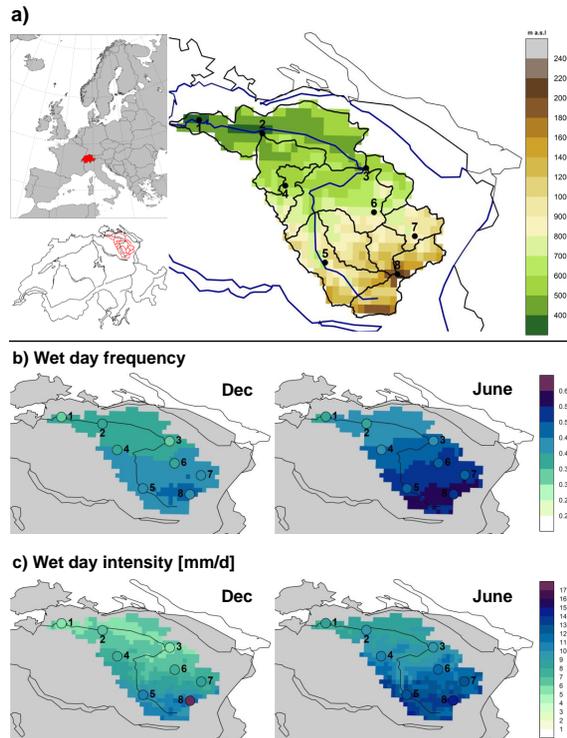



Figure 1. (a) The catchment of the river *Thur*, located in north-eastern Switzerland, together with the underlying topography (in m.a.s.l.). The dots indicate the locations of the investigated stations. 1: *Andelfingen* (AFI), 2: *Frauenfeld* (FRF), 3: *Bischofszell* (BIZ), 4: *Eschlikon* (EKO), 5: *Ebnat-Kappel* (EBK), 6: *Herisau* (HES), 7: *Appenzell* (APP), 8: *Saentis* (SAE). (b) Observed precipitation climatology of the wet day frequency (1961–2011) derived from a $2\text{ km} \times 2\text{ km}$ gridded daily precipitation dataset (Frei and Schär, 1998) for December and June. (c) The same as in (b), but for wet day intensity (in mm day^{-1}). The filled circle symbols point to the station locations (as in a) together with the observed station measurements.

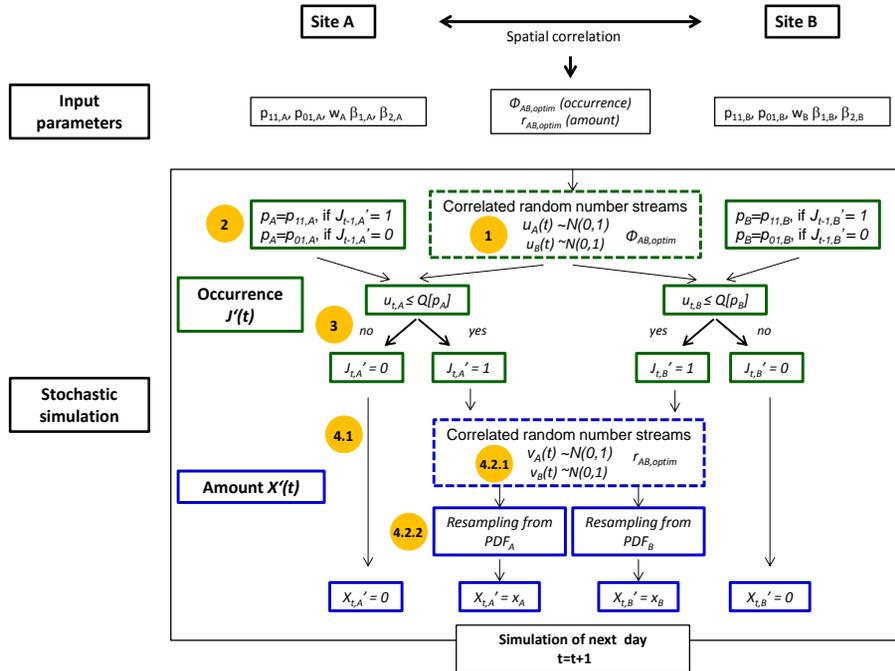


Figure 2. Technical workflow of a multi-site precipitation generator after Wilks (1998) at the example of two fictitious sites A and B. In general, it is a combination of multiple single-site precipitation generators that are calibrated at each site individually (see input parameters) and run simultaneously with spatially correlated random number streams (dashed boxes). The correlated random number streams (of similar length as the simulation period) are determined beforehand (see Sect. 3.3.2). The orange-labelled numbers indicate the steps for single-site precipitation simulation (see Sect. 3.3.1).

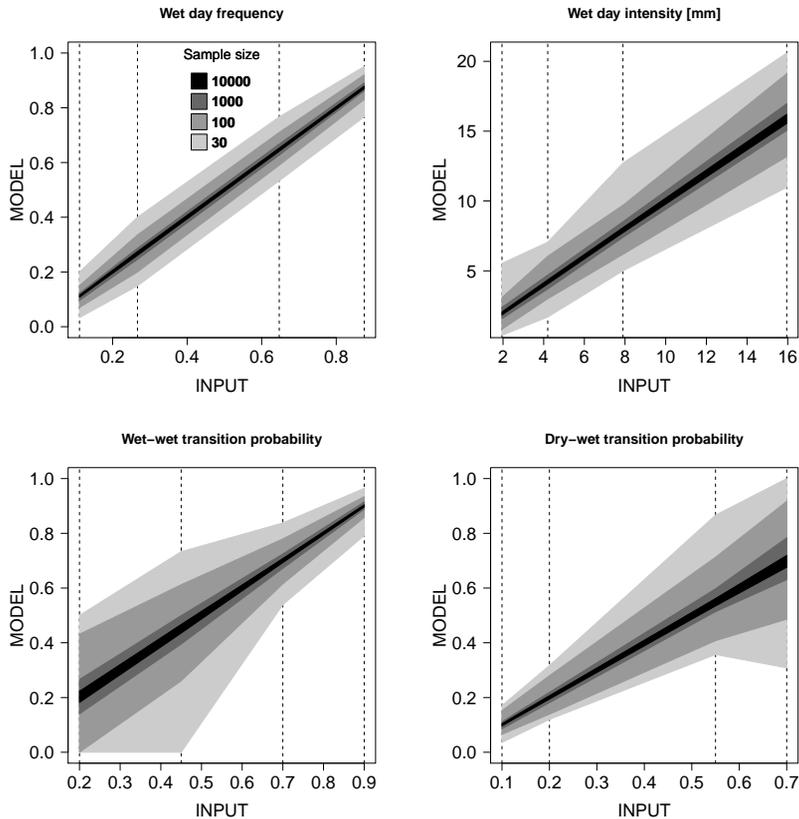


Figure 3. Reproduction of average wet day frequency (wdf), mean wet day intensity (wdi), wet–wet transition probability (p_{11}) and dry–wet transition probability (p_{01}) for the four idealized climate regime ranging from very dry (left) to very wet (right) as indicated by dashed lines. The shaded areas correspond to the range between the 2.5 % and the 97.5 % empirical quantiles of 100 realizations. Results are shown for sample sizes of 10 000, 1 000, 100 and 30 (grey shading).

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

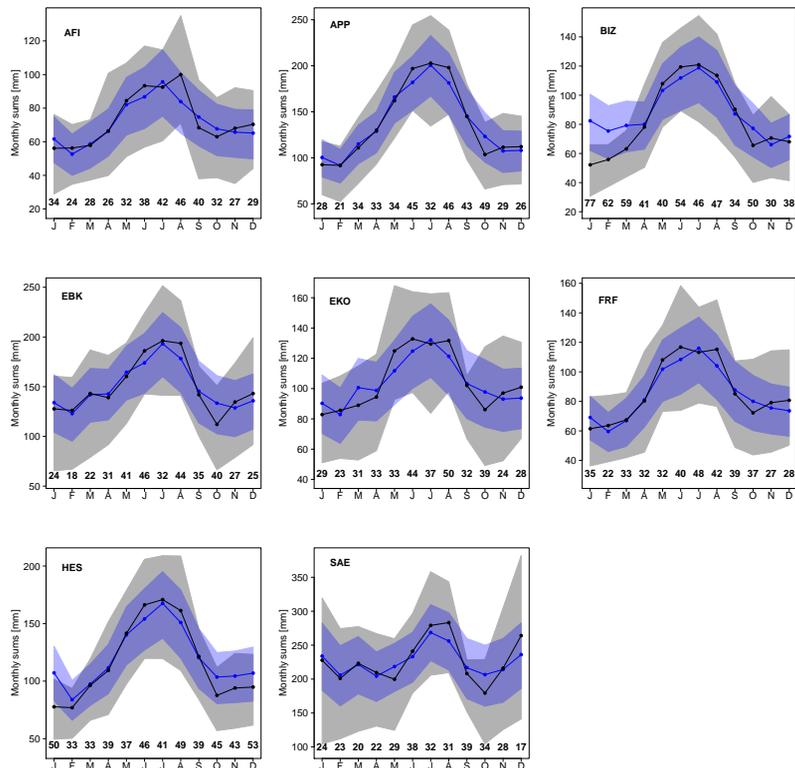


Figure 4. Long-term mean and variability of monthly precipitation sums during the period 1961–2011 for eight stations in the *Thur* catchment. The black (blue) lines refer to the mean annual cycle of observed (modelled) precipitation sums. The grey (blue) shaded areas represent the inter-quartile ranges of observed (simulated) monthly precipitation sums. The simulation comprises 100 realizations covering each 51 years. The numbers at the bottom indicate for each month the percentage of variance explained by the precipitation generator. Note that the scale of the y-axis differ between different stations.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

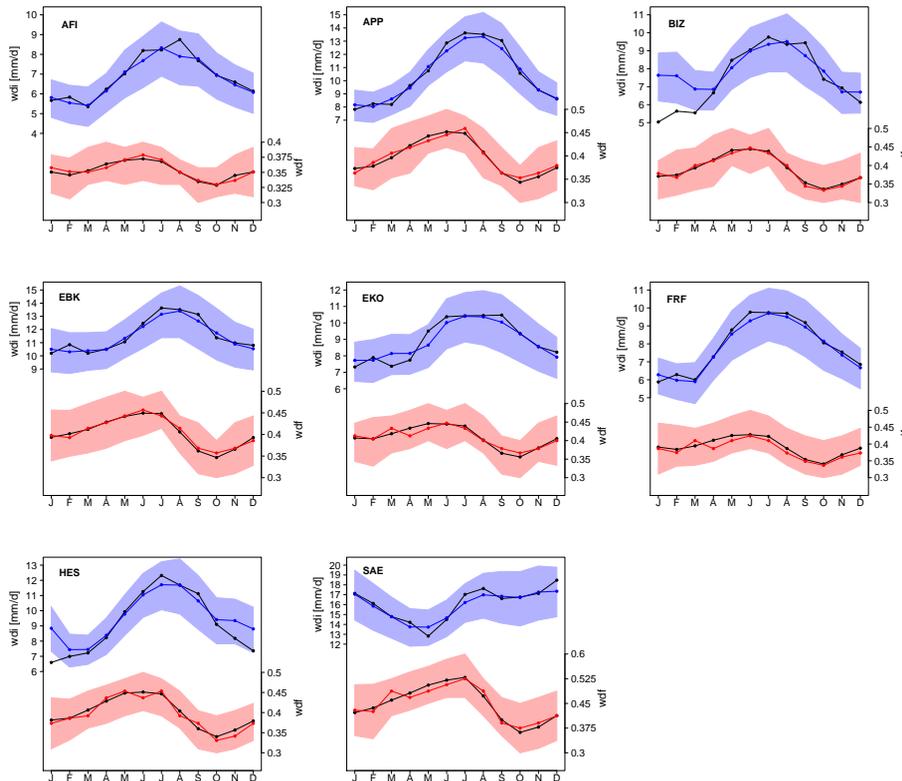


Figure 5. Observed and modelled monthly mean wet day intensity (blue) and frequency (red) at eight stations during 1961–2011. The black (coloured) lines indicate the observed (modelled) values. The blue (red) shaded areas correspond to the inter-quartile range across the set of synthetic daily time-series. They comprise 100 runs covering each 51 years.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[⏪](#)
[⏩](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

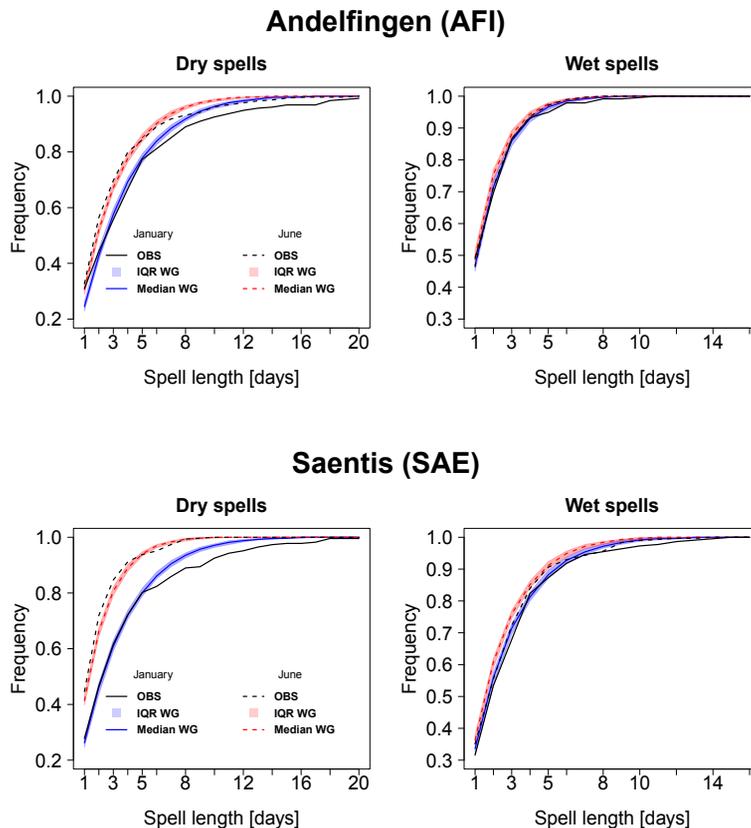


Figure 6. Cumulative distribution of the observed and simulated dry (left) and wet (right) spell length frequencies for the lowland station *Andelfingen* (top) and the mountain station *Saentis* (bottom). Results are for January and June during the time period of 1961–2011. The coloured area (line) represents the inter-quartile range (median) of the 100 realizations covering each 51 year-long daily time-series.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

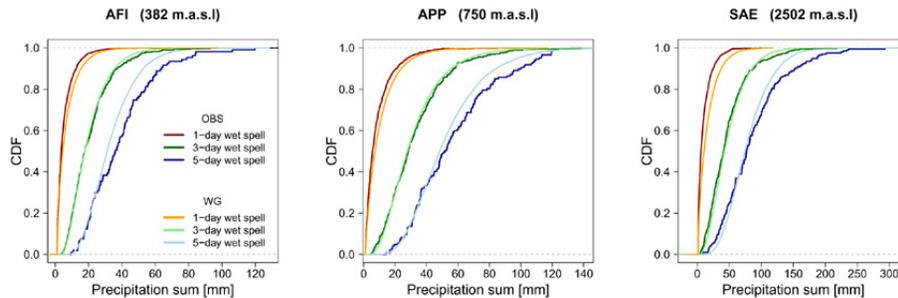


Figure 7. Cumulative distribution functions (CDFs) of multi-day precipitation sums for the three stations *Andelfingen* (AFI), *Appenzell* (APP) and *Saentis* (SAE). The lines represent the CDFs of non-zero precipitation amounts over one wet day (red), over three consecutive wet days (green) and over five consecutive wet days (blue). Darker and lighter colours refer to observations and simulations, respectively. The observed CDFs have been derived from a 51 year long daily time-series between 1961 and 2011, those of the weather generator from 100 realizations of 51 year long daily simulations. Note that the scaling of the horizontal axis differs between different stations.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[⏪](#)
[⏩](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

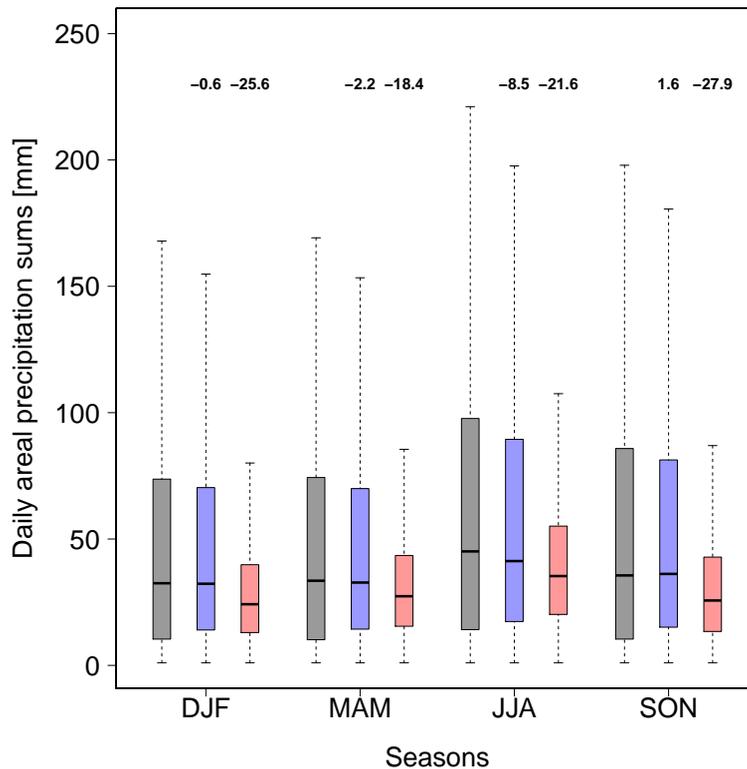


Figure 8. Daily non-zero precipitation sums over the catchment for the four seasons during 1961–2011. Daily Precipitation intensity of the eight stations are summed and days with an area sum of zero are excluded. Boxplots of observed daily sums (grey), of multi-site simulated time-series (blue) and of single-site simulated time-series (red) are shown. The WG models were run 100 times over a 51 year time-period. The numbers (in percentage) indicated above the corresponding model represent the relative deviation of the simulated median from the observed.

Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

[Title Page](#)

[Abstract](#) | [Introduction](#)

[Conclusions](#) | [References](#)

[Tables](#) | [Figures](#)

[⏪](#) | [⏩](#)

[◀](#) | [▶](#)

[Back](#) | [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Stochastic modelling of spatially consistent daily precipitation time-series

D. E. Keller et al.

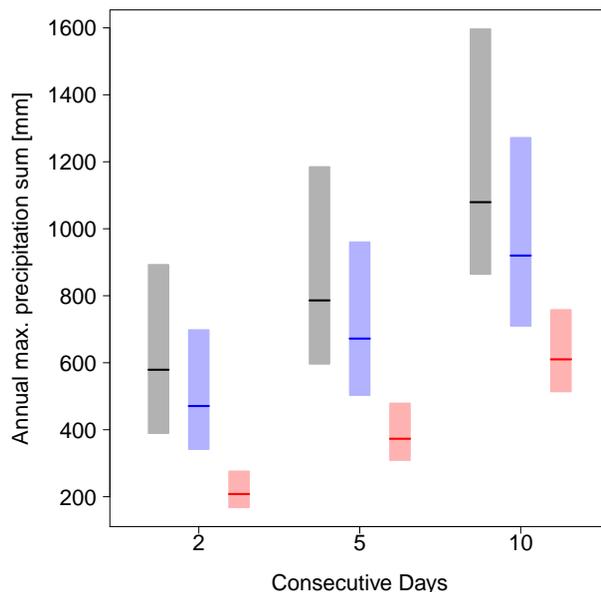


Figure 9. Annual maximum precipitation summed over all eight stations and over consecutive days. The analysis is done for all days of year. The bars indicate the range between the 2.5 % and the 97.5 % empirical quantiles of the yearly maximum area sums during 1961–2011. The observations are plotted in grey, the multi-site simulations in blue and the single-site simulations in red. The observations comprise 51 years, the models were run 100 times over a 51 year time-period.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)