Guan, X., Nissen, K., Nguyen, D., Merz, B., Winter, B., Vorogushyn, S. (2023): Multisite temporal rainfall disaggregation using methods of fragments conditioned on circulation patterns. - Journal of Hydrology, 621, 129640.

https://doi.org/10.1016/j.jhydrol.2023.129640

Multisite temporal rainfall disaggregation using methods of fragments conditioned on circulation patterns

Xiaoxiang Guan[a], Katrin Nissen[b], Dung Viet Nguyen[a], Bruno Merz[a,c], Benjamin Winter[d], Sergiy Vorogushyn[a]

a. GFZ German Research Centre for Geosciences, Section Hydrology, Potsdam, Germany

b. Institute of Meteorology, Free University of Berlin, Berlin, Germany

c. University of Potsdam, Institute of Environmental Science and Geography, Potsdam, Germany

d. University of Innsbruck, Institute of Geography, alpS Research, Innsbruck, Austria

Correspondence: Xiaoxiang Guan (Email: guan.xiaoxiang@gfz-potsdam.de)

1. Introduction

A wide range of applications involving planning, design, and management of water resources systems at small spatial scales relies on rainfall data with high temporal, for instance hourly, resolution (Breinl and Di Baldassarre, 2019; Pui et al., 2012). However, high-resolution rainfall measurements are generally scarce (Gutierrez-Magness and McCuen, 2004), especially for rural and remote regions. The precipitation gauging network operated by the German Weather Service (DWD), for example, includes roughly 5,700 stations with daily rainfall observation history spanning back more than 100 years. Compared to this, only about 1,200 stations record rainfall at subdaily resolution and an even smaller number has publicly available data covering more than 30 years (Lisniak et al., 2013).

One way to overcome this shortcoming is to derive fine-resolution data from the widely available coarser-resolution (e.g. daily) data through data transformation procedures, generally referred to as rainfall disaggregation (Gutierrez-Magness and McCuen, 2004; Koutsoyiannis and Onof, 2001; Sharma and Srikanthan, 2006). Temporal disaggregation is often used for scenario-based hydrological simulations when combined with regional weather generators (Mezghani and Hingray, 2009; Winter et al., 2019). For instance, Winter et al. (2019) estimated 100-year design floods in 16 catchments in Vorarlberg, Austria with a continuous hydrologic modelling approach with hourly resolution, driven by a multi-site weather generator providing daily precipitation and air temperature in combination with a temporal disaggregation procedure.

Numerous disaggregation models have been proposed in the literature based on diverse concepts. These include: (1) Bartlett–Lewis/Neyman–Scott rectangular pulse models based on point process theory (Khaliq and Cunnane, 1996; Lu and Qin, 2014; Pui et al., 2012); (2) random cascade models based on scale-invariance theory (Anis and Rode, 2015; Müller and Haberlandt, 2018); and (3) the nonparametric method of fragments (MOF) inspired from analog principle (Carreau et al., 2019; Li et al., 2018; Lu et al., 2015;

Sharma and Srikanthan, 2006; Westra et al., 2012). Pui et al. (2012) compared these temporal rainfall disaggregation models at single sites and found that the nonparametric MOF outperformed the point process-based and cascade models in matching the observed hourly intensity-frequency relationship, in particular for extreme rainfall characteristics.

MOF achieves temporal disaggregation by employing the subdaily distribution of analog days, also known as fragments, to distribute daily precipitation totals into subdaily intervals. In the daily-to-hourly disaggregation procedure, the fragments are the rainfall amounts (or fractions of the daily rainfall amount) of the 24 hours of one day (Sharma and Srikanthan, 2006). Analog days are defined as days that are similar in terms of a number of features (Carreau et al., 2019). Often the only feature used to determine analog days is the daily rainfall total at a given location (Sharma and Srikanthan, 2006; Westra et al., 2012). Recently, modifications were introduced to MOF to improve the model's ability in reproducing the hourly rainfall. These modifications include the consideration of the wet-dry state of the days preceding and following the target day (Breinl and Di Baldassarre, 2019; Westra et al., 2012), other climate variables like temperature, relative humidity and air pressure (Rafatnejad et al., 2021), spatial information such as inter-site correlation or neighboring information (Carreau et al., 2019; Müller and Haberlandt, 2018), and the category of daily rainfall amount (classified in 5 mm bins) (Li et al., 2018). The features introduced in these modifications to better filter the analog days are generally daily variables. Hence, it is implicitly assumed that these daily features allow to condition the precipitation distribution on the subdaily scale (Carreau et al., 2019).

The basic assumption of the MOF rainfall temporal disaggregation is stationarity, i.e. that the daily-subdaily rainfall relationship from the analog period remains unchanged in the period for which the disaggregation is performed. This assumption is particularly questionable when the disaggregation is performed for future climate change projections. In order to account for dynamic changes in the atmosphere, circulation patterns (CPs) can be used to condition the selection of analog days for MOF.

CPs describe the atmospheric circulation at any given moment in time, and CP classifications are applied to simplify the physical reality by identifying a small number of representative patterns to which the instantaneous patterns are assigned (Huth et al., 2008). One of the goals of CP classification is to improve the description of effects the atmospheric circulation has on surface climate, like precipitation formation (Huth et al., 2016). Conditioning the rainfall temporal disaggregation on CPs allows to consider climate change effects related to changes in frequency, persistence, and seasonality of these weather patterns on the link between daily and hourly rainfall distributions. However, possible within-type changes, in particular due to increasing temperature and (Super-) Clausius-Clapeyron-scaling, may strongly determine the link between daily and hourly rainfall distributions (Lenderink et al., 2017). If such changes occur, they will not be represented by a CP-based disaggregation. Besides the capability to consider climatic changes – at least to some extent, a CP-based disaggregation may also be preferable as different CPs may be associated with different precipitation types (e.g., large-scale, long-lasting events vs. small-scale, short-duration events), which may have different subdaily distributions (Kronenberg et al., 2012; Lisniak et al., 2013). Hence, we hypothesize that MOF conditioned on CPs can better represent the rainfall generation mechanisms and provides a more robust link between daily and subdaily rainfall in the context of climate change than a standard MOF procedure.

In this study, we develop a rainfall temporal (daily to hourly) disaggregation procedure by conditioning the method of fragments on circulation patterns for multisite applications. The performance of this CP-based disaggregation procedure is evaluated against the standard MOF procedure in the Rhine river basin. The sensitivity of the MOF to the number of CPs and seasonal stratification is further analyzed and discussed.

2. Research area and data

2.1 Research area

The Rhine River basin with a drainage area of about 185,000 km^2 is situated in northwestern Europe, covering 9 countries (Figure 1). With a length of about 1,320 km, it flows from the Swiss Alps through middle mountain ranges and lowlands and drains into the North Sea in the Netherlands (see Ullrich et al. (2021) for a detailed description). The Rhine river and its tributaries are often affected by flooding caused by persistent precipitation, heavy rainfall or snowmelt. In July 2021, a record-breaking flood hit west Germany, Belgium, Luxembourg and the Netherlands and in particular the left-side Rhine tributaries Ahr, Erft and Ruhr, causing more than 200 fatalities and tremendous socio-economic impacts (Mohr et al., 2022). Hydrological analysis of such floods in small-scale catchments require hourly precipitation time series, which are rarely available. Long-term precipitation series are often available at daily time scale and need to be disaggregated. We select the rainfall gauges located in the German part of the Rhine basin to develop the rainfall disaggregation model and to compare it to existing methods given the availability and accessibility of hourly rainfall data in this region.

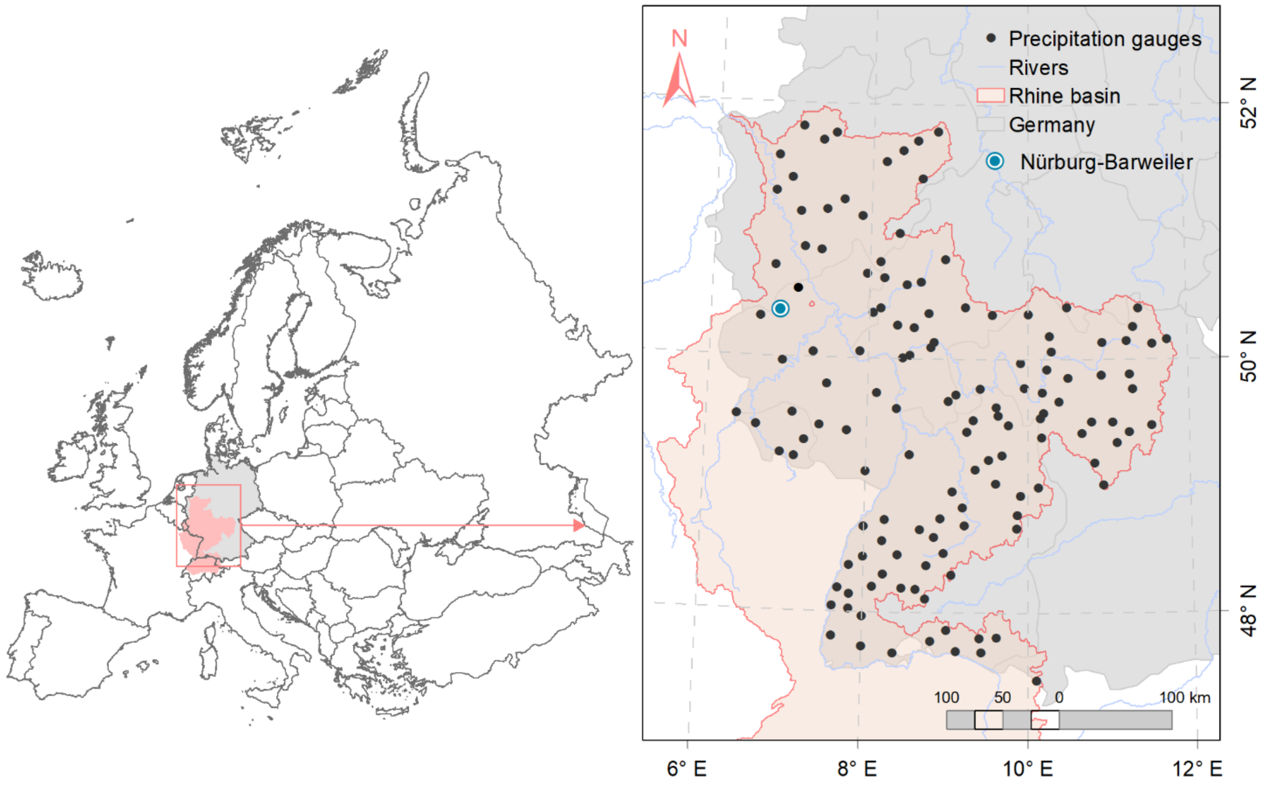

Figure 1 Research area (German part of the Rhine river basin) and the associated hourly rainfall observation network with 134 precipitation gauges and an exemplary gauge (Nürburg-Barweiler).

2.2 Hourly precipitation observations

The hourly rainfall observations are collected from the Climate Data Center of the German Weather Service (DWD, https://www.dwd.de/, last access: 26th of February, 2022). These observations are quality controlled measurements from DWD stations and legally and qualitatively equivalent partner stations. There are in total 1,038 stations across Germany, most of which provide hourly rainfall observations since 2000. 134 rainfall stations located in the German part of the Rhine basin are selected with at least 15 years (2006-2020) of hourly rainfall records and with no more than 2% missing data in each year (Figure 1).

2.3 Circulation patterns

We classify circulation patterns using daily values of the mean sea-level pressure field of the ERA5 data set for the period from January 1979 to July 2021. ERA5 is the fifth generation ECMWF atmospheric reanalysis of the global climate (Hersbach et al., 2020). We smoothed the original ERA5 data by interpolating it to a regular 1°x 1° grid. The classification is conducted using the objective classification algorithm SANDRA (**S**imulated **AN**nealing and **D**iversified **RA**ndomization) (Philipp et al., 2007). It is based on k-means and minimizes the within-cluster variance of the Euclidian distance between the cluster elements and the cluster centroid. The problem of a conventional k-means approach is that it often converges to a local optimum. SANDRA avoids this problem and searches for the global optimum by introducing random reassignments of cluster elements. The number of classes in CP classification affects its ability to stratify the surface climate variable of interest (Huth et al., 2016; Vallorani et al., 2018). Therefore, 5 different numbers of CP classes (4, 5, 6, 7, and 8) are generated and the sensitivity of the disaggregation performance to the number of classes is examined.

3. Methods

3.1 Multisite rainfall disaggregation with method of fragments

The method of fragments (MOF) is a non-parametric disaggregation technique. The idea is to resample a vector of fragments that represents the relative distribution of subdaily to daily rainfall (Pui et al., 2012). The number of fragments corresponds to the subdaily temporal resolution used, i.e. if the disaggregation is conducted from daily to hourly resolution, the relative distribution of subdaily values consists of 24 relative weights that sum up to 1. In the simulation, variability is introduced by a k-nearest neighbor algorithm. The proposed non-parametric multisite MOF model works as follows:

(i) Obtain the daily rainfall vector $\{R_{t,1}, R_{t,2}, \cdots, R_{t,n}\}$ to be disaggregated where t represents the date of the day, and n denotes the number of rainfall sites in the observation network. $R_{t,s}$ (s indicates individual rainfall site, with the value ranging from 1 to n) can be obtained from the observed daily rainfall or from other sources, such as daily weather generators (Nguyen et al., 2021; Winter et al., 2019) and General Circulation Models (GCMs) (Rafatnejad et al., 2021).

(ii) Use the observational hourly records, $R_{i,m,s}$, to build daily time vectors $R_{i,s}$, where *i* denotes the day, *m* is the hourly time step and *s* is a site of the rainfall observation network.

$$R_{i,s} = \sum_{m=1}^{24} R_{i,m,s} \tag{1}$$

Form a time series of vectors with hourly to daily ratios, which are so-called fragments.

$$f_{i,m,s} = R_{i,m,s}/R_{i,s} \tag{2}$$

(iii) To consider the seasonal variability, the standard MOF based disaggregation procedure builds a window with *l* days around the target day *t* from which to sample the fragments vectors (Westra et al., 2012; Pui et al., 2012). For example, if *t* represents the 15th of April and *l* = 14, all days between the first (1st) and last (29th) day of April from all available years are considered for disaggregation. This standard disaggregation procedure, usually with a time window width *l*=14, is further called monthly-based disaggregation procedure.

In this study, the MOF conditioned on circulation patterns is proposed and termed CP-based disaggregation procedure. To this end, the days with the same CP class as the target day *t* are selected into the candidate pool for fragments sampling instead of imposing a monthly window. Additionally, circulation patterns and candidate days are stratified into winter (November 1st – April 30th) and summer (May 1st – October 30th) seasons.

(iv) Filter the candidate pool obtained in Step (iii) based on the wet (rainy)-dry (non-rainy) status of the target day. Only days with the same wet-dry status as the target day *t* will be selected to obtain possible nearest neighbors; here N denotes the number of possible neighbors.

(v) Before calculating the distance between target day and candidate day, the daily rainfall amounts are standardized, which is a preferable operation for positively skewed random variables such as rainfall (Breinl and Di Baldassarre, 2019).

In this study, the Manhattan distance (the sum of absolute differences) is used as a measure to quantify the similarity of multisite daily rainfall vectors between target day *t* and candidate day *i*. The Manhattan distance has been shown to work well with nearest neighbor algorithms for rainfall disaggregation (Breinl and Di Baldassarre, 2019; Breinl et al., 2017). The distance $d(R_t, R_i)$ is calculated as:

$$d(R_t, R_i) = \sum_{s=1}^{n} \left| R_{t,s} - R_{i,s} \right| \tag{3}$$

(vi) Identify the k nearest neighbors, with $k = \sqrt{N}$. The distances are sorted for all *j* = 1,2, …, *k* and the highest probability is assigned to the neighbor with the smallest distance to the target day. The probability (p_j) is computed as:

$$p_j = \frac{1/j}{P_t}, P_t = \sum_{j=1}^{k} 1/j \tag{4}$$

(vii) Sample a candidate day by applying the inverse cumulative distribution function of Eq. (4) and using random numbers between 0 and 1 sampled from a uniform distribution. The date of the sampled day is used and the corresponding hourly fragments ($f_{i,m,s}$) are applied at each site in the disaggregation. The hourly time series $R_{i,m,s}$ of the target day t are then derived using daily rainfall series ($R_{t,s}$) as:

$$R_{t,m,s} = R_{t,s} \times f_{i,m,s} \tag{5}$$

(viii) Repeat Step (i) to Step (vii) for each day *t* until the entire daily records are disaggregated.

3.2 Experimental setup

To evaluate the CP-based disaggregation procedure developed in this study and to investigate its sensitivity towards the number of CP classes and towards the consideration of seasonality, three disaggregation experiments are performed (Table 1). To make full use of the limited hourly observations, we perform a leave-one-out cross-validation. Each single year of the hourly observations is selected and aggregated to the daily scale to be subsequently disaggregated using the remaining 14 years. The original hourly data for each year is then used to test the performance of the disaggregation procedure in the different experiments. A total of 30 Monte-Carlo (MC) runs, given the computational constraint, with the same length (15 years) as the observed precipitation records are generated to explore the sampling variability of the nearest neighbor candidate date.

Table 1 Overview of the disaggregation experiments. CLA4 to CLA8 and SW_CLA4 to SW_CLA8 abbreviations denote the experiments with 4 to 8

Experiment	Scheme	Classes	Conditional variable
Month	Standard monthly-based MOF disaggregation procedure	12 (months)	Month
CLA {4, 5, 6, 7, 8}	MOF disaggregation procedure conditioned on circulation patterns	Number of CPs {4, 5, 6, 7, 8}	CP
SW_CLA {4, 5, 6, 7, 8}	Disaggregation procedure conditioned on circulation pattern classification, considering two seasons (summer and winter)	Number of CPs 2×{4, 5, 6, 7, 8}	CP and season

3.3 Performance evaluation

The performance evaluation of the different rainfall disaggregation experiments employs the following indicators, as suggested by previous studies (Breinl and Di Baldassarre, 2019; Li et al., 2018; Pui et al., 2012). The performance in reconstructing standard rainfall statistics includes mean value, standard deviation, fraction of wet (rainy) hours, and lag-1 autocorrelation. The mean wet/dry spell length is selected to reflect the model performance in simulating the wet and dry features in the rainfall temporal disaggregation. The skill in maintaining the spatial correlation structure is indicated by inter-site Pearson correlation coefficients. Finally, the performance of the disaggregation procedures to meet the extreme percentiles (95^{th}, 97^{th}, 99^{th} and 99.5^{th}) of hourly rainfall is evaluated. The metrics for disaggregated precipitation relate to the median of the 30 MC runs and are compared against the indicators of the original records.

4. Results

4.1 Standard metrics

According to the different experiments (Table 1), the simulated hourly rainfall series at the 134 rainfall stations in the German part of Rhine river basin are compared with the observed counterparts, and the statistical metrics reflecting the disaggregation performance in reproducing key hourly rainfall characteristics are calculated. Figure 2 summarizes the model performance with regard to mean value, standard deviation (std), lag-1 autocorrelation coefficient (lag1auto) and fraction of wet hours. All disaggregation models tend to underestimate the mean hourly rainfall values, which results from the overestimation of the number of wet hours in disaggregation procedures as the same daily precipitation totals are distributed to more wet shares. The monthly-based MOF shows a slightly better performance in reproducing the mean value than the CP-based procedures, but the difference in mean value between monthly-based and CP-based approaches with seasonal stratification and high number of classes (like SW_CLA8) is not large. All model variants perform well with respect to standard deviation. Lag-1 autocorrelation is underestimated by the monthly-based MOF and overestimated by the CP-based MOF, whereas the seasonally stratified CP-based procedure performs best. Generally, the seasonally stratified CP-based procedure outperforms the one without stratification for all four indicators. With increasing number of CP classes, the performance improves.

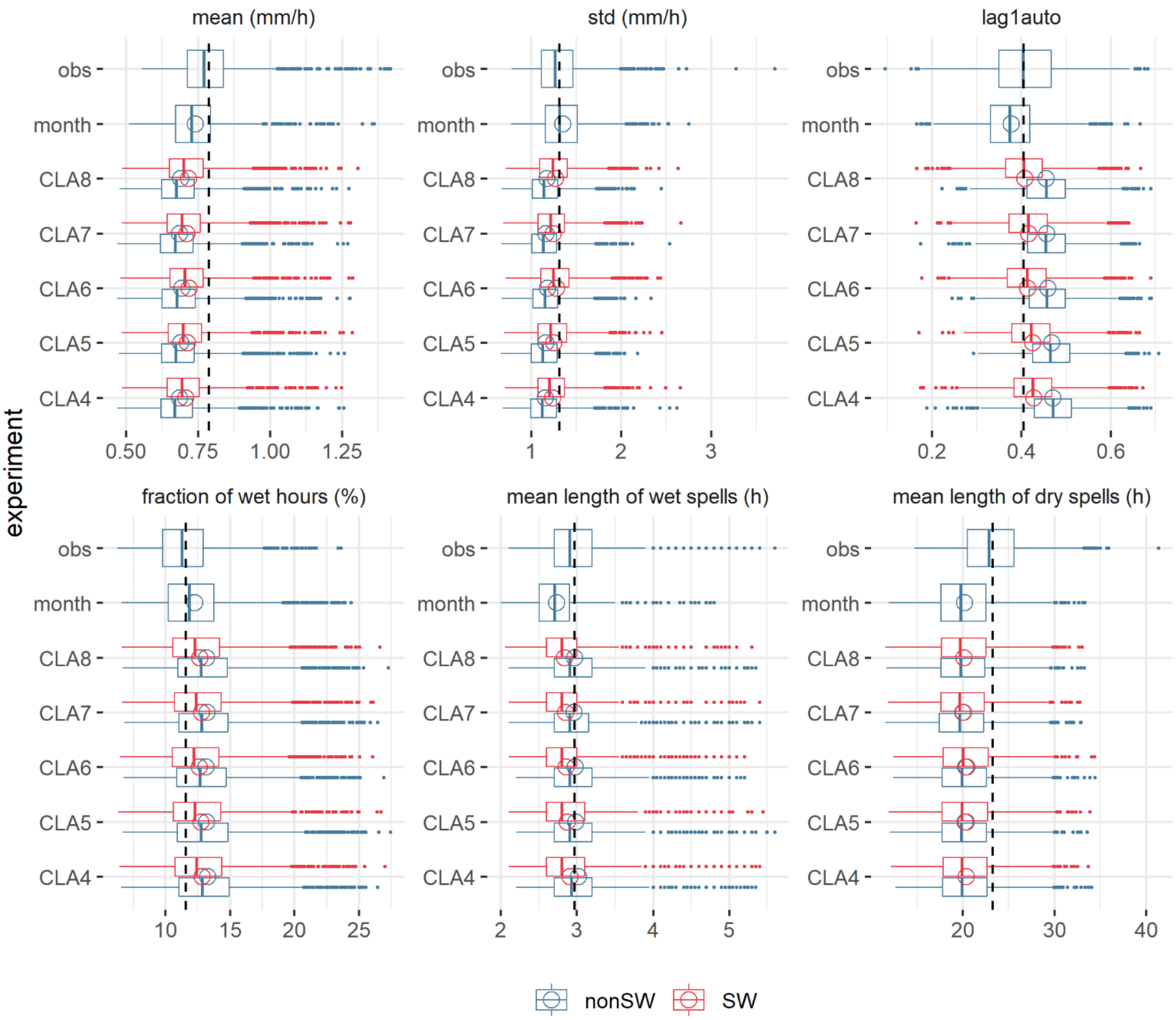

Figure 2: Comparison of mean, standard deviation (std), lag-1 autocorrelation (lag1augo), fraction of wet hours, mean length of wet and dry spells (average duration of precipitation events) between disaggregated hourly rainfall in different experiments (see Table 1) and observations (obs). SW and nonSW denote the results from CP-based disaggregation with and without considering season stratification respectively. The boxplots cover the median of metrics from 30 MC runs in 134 stations, where the middle solid line marks the median and the box represents the interquartile range. The black dashed lines represent the mean values of statistics for observed hourly rainfall, while the circles denote the mean values of the simulated counterparts.

All disaggregation procedures consistently underestimate the dry spell duration compared to observations (Figure 2). The mean length of the simulated dry spells is approximately 20 hours compared to the observed 23 hours. Such underestimation is also demonstrated in Figure 3 (b) and (d) and the difference among different disaggregation experiments is nearly indistinguishable. In terms of mean length of wet spells, the CP-based models perform better than monthly-based procedure, especially for longer wet spells (see Figure 3 a and c) which is largely underestimated by the monthly-based disaggregation procedure. We suggest that CP-based disaggregation pools days with long and short duration wet spells which may correspond to frontal and convective rainfall, respectively, in a better way than monthly-based disaggregation. The discontinuity of rainfall events between days could play a role as well, as the disaggregation procedure does not consider the wet-dry status in the preceding and following days of the target day.

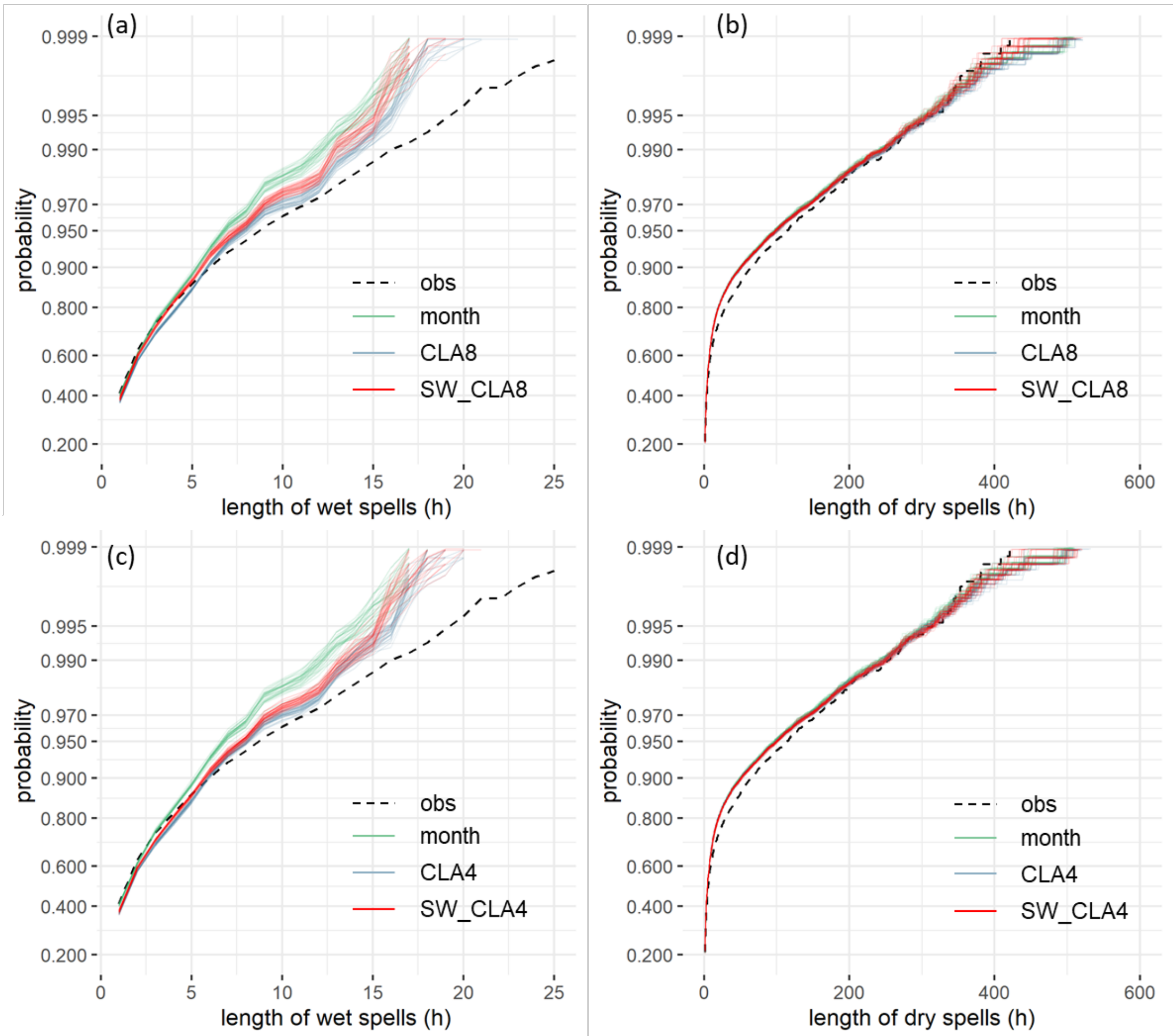

Figure 3 Comparisons of the cumulative distribution probability curves (shown as logit-scaled) of length of wet (a and c) and dry (b and d) spells for the example station Nürburg-Barweiler (see Figure 1) for five disaggregation procedures (month, CLA4, SW_CLA4, CLA8, and SW_CLA8) and observations

4.2 Spatial correlation

The pair-wise inter-site Pearson correlation coefficients of hourly rainfall series are given in Figure 4 to assess the spatial connections of precipitation across station locations. The deviations between simulated and observed correlation coefficients versus inter-site distance are given in Figure 5. All disaggregation procedures reproduce the spatial correlation well, with most absolute deviations between simulated and observed coefficients lower than 0.1. The monthly-based MOF performs best in terms of spatial correlation. The CP-based disaggregation procedure without seasonal stratification tends to slightly overestimate the inter-site correlations, whereas considering the seasonality (Figure 5) clearly reduces the bias. The number of CPs does not affect the performance for this characteristic.

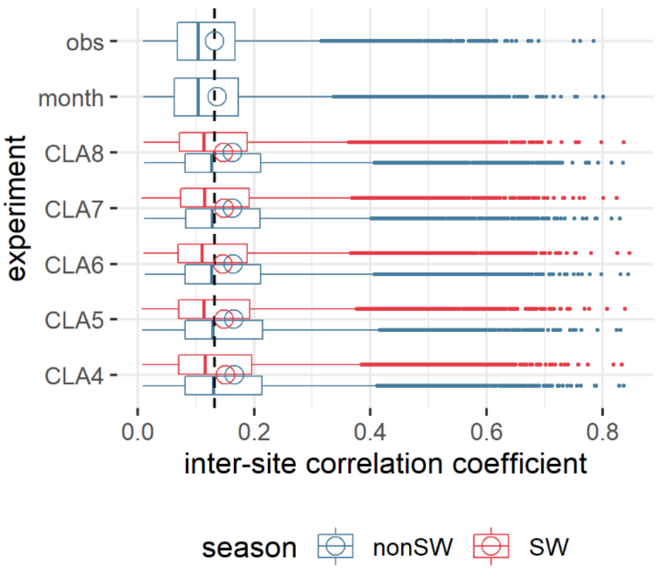

Figure 4 Comparison of inter-site correlation coefficients for different disaggregation procedures and observations

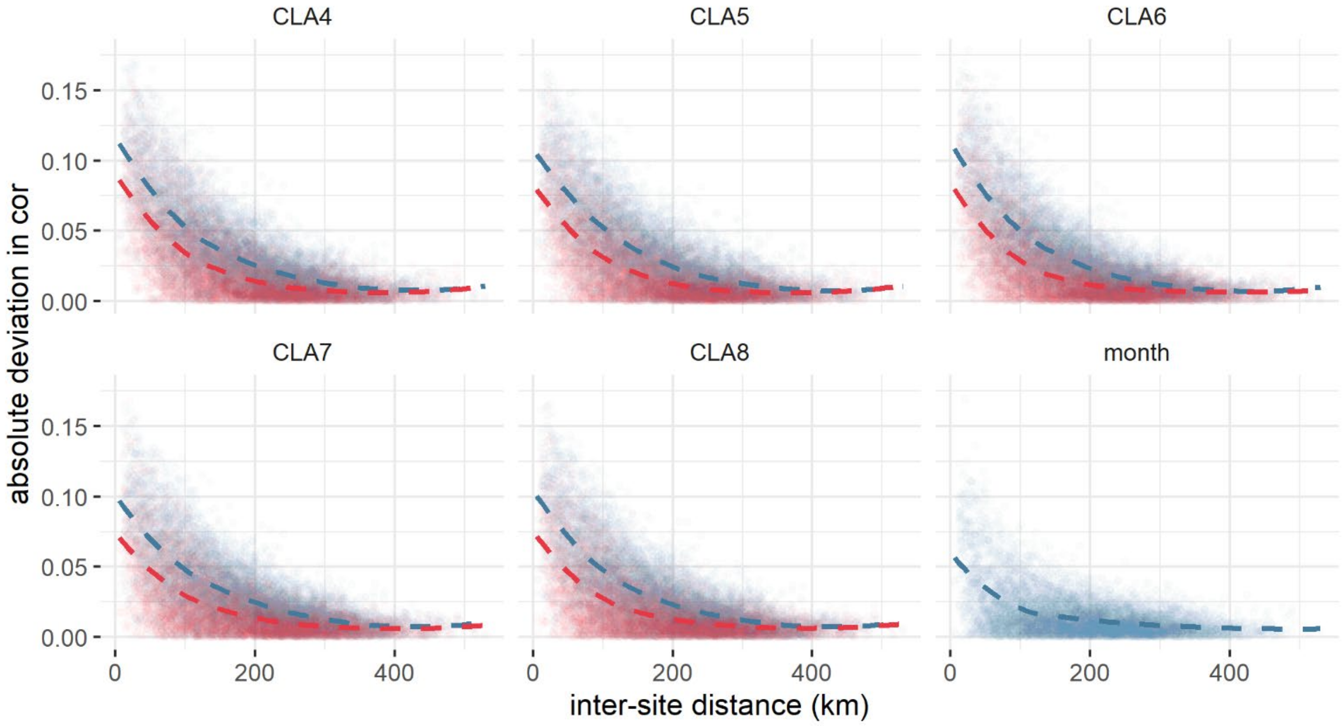

Figure 5: Deviation in inter-site correlation coefficients vs inter-site distance between simulated and observed hourly rainfall series. The distances between gauges are computed by using the Haversine formula. In the CP-based experiments (CLA4-8), the red and blue circles, together with the corresponding dash lines, denote the results from disaggregation with and without considering seasonal stratification, respectively.

4.3 Precipitation extremes

To examine the disaggregation performance in reproducing the hourly rainfall extremes, several high percentiles are calculated and compared with observations (Figure 6). In addition, the cumulative distribution functions for the example station Nürburg-Barweiler are given in Figure 7. For less extreme precipitation (95th percentile) the monthly-based MOF outperforms the CP-based disaggregation. For more extreme precipitation (97th, 99th, 99.5th percentiles), however, the monthly-based procedure overestimates, the CP-based without seasonal stratification underestimates and the CP-based with seasonal stratification largely matches the observed percentiles. This is also demonstrated in Figure 7 (for the example rainfall station) that the probability curves from 30 runs of SW_CLA8 disaggregation can better envelope the observation one (black dashed line). The number of CP classes does not play a strong role, though the classification with four classes performs worst.

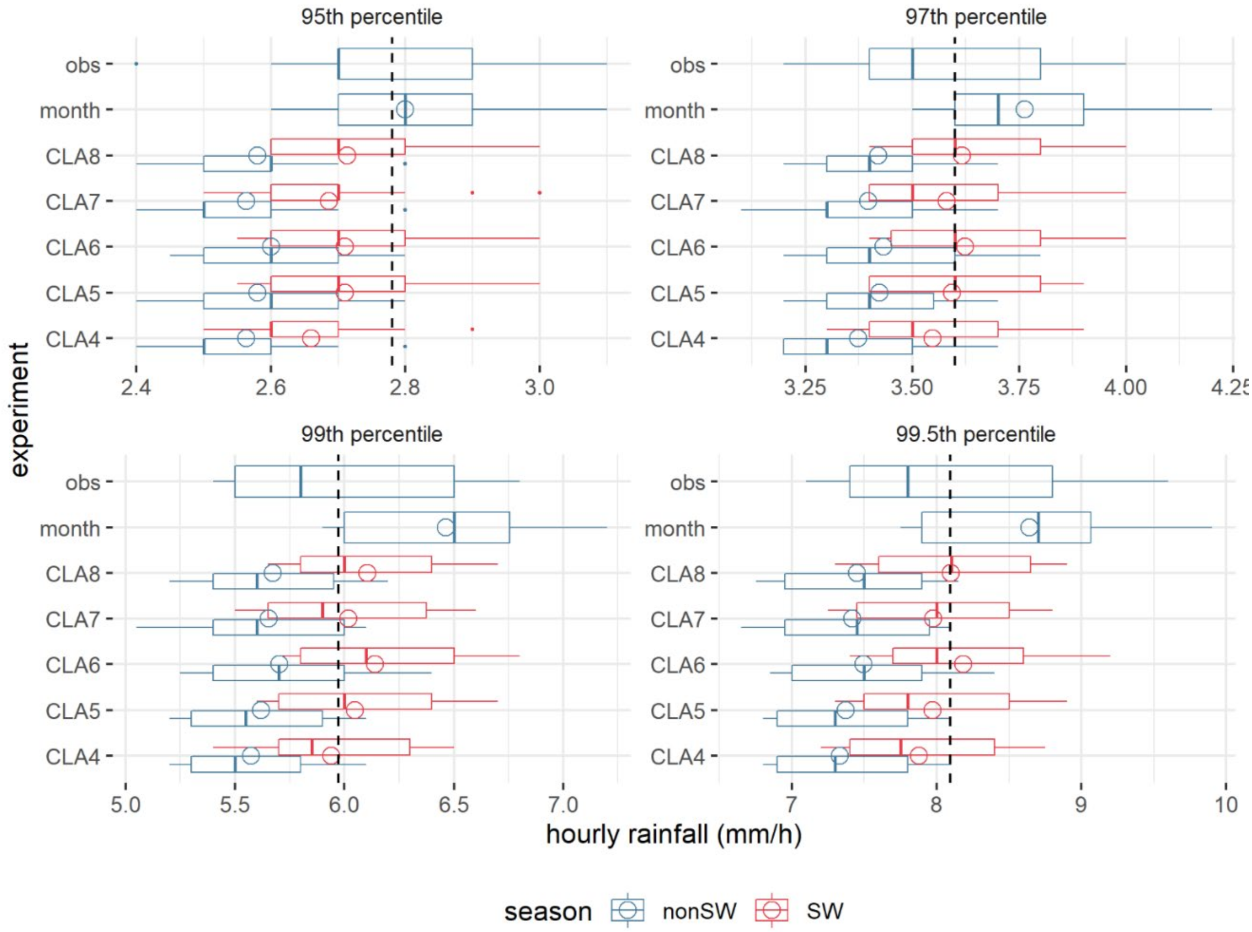

Figure 6: Comparison of extreme hourly rainfall percentiles for different disaggregation procedures and observations.

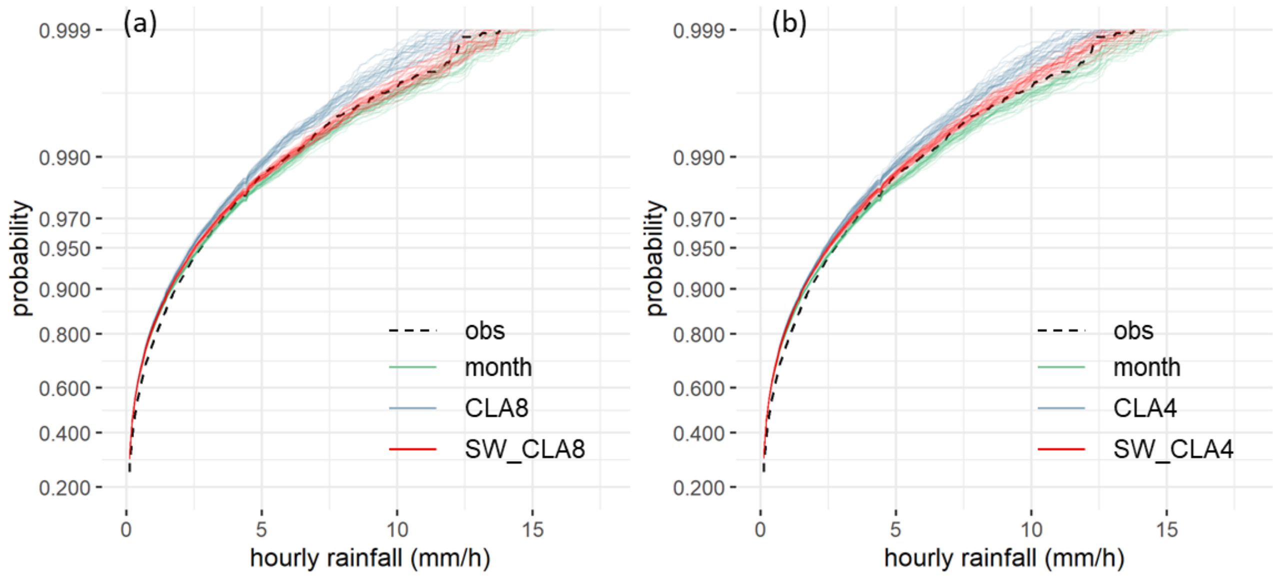

Figure 7 Comparisons of the cumulative distribution functions (shown as logit-scaled) of hourly rainfall for the example station Nürburg-Barweiler (Figure 1) for five disaggregation procedures (month, CLA4, SW_CLA4, CLA8, and SW_CLA8) and observations.

5. Discussion

In this study, we demonstrate that the temporal rainfall disaggregation based on MOF is improved in reproducing hourly rainfall extremes when conditioned on circulation patterns along with seasonal (winter/summer) stratification. We assume that the additional information injected by CPs and seasonality better describes precipitation formation mechanisms and hence the link between daily sums and hourly precipitation distributions. Additionally, the number of CP classes also plays a role. Models with more CP classes and seasonality considered perform better in rainfall temporal disaggregation and the SW_CLA8 disaggregation procedure shows the best performance in extreme hourly rainfall simulation (Figure 6 and Figure 7). It is expected that increasing the number of classes in CP stratification improves the performance in explaining the rainfall variability, as shown in Beck and Philipp(2010) and Huth et al. (2008), and further improves the performance of the disaggregation model. However, increasing the number of CP classes is computationally more demanding and also each class becomes less populated since the length of time series is limited, which will inevitably introduce more uncertainty in candidate days sampling. Another concern is related to the underestimation of hourly rainfall mean values in all disaggregation experiments, which results from the overestimation of fraction of wet hours in the fragments sampling procedure. Although monthly-based standard MOF shows lower underestimation (Figure 2), the difference between monthly-based model and CP-based approach with seasonal stratification and high number of CP classes (like SW_CLA8) is rather small. Apparently, all models underestimate precipitation in the range between 50^{th} and 95^{th} percentiles (Figure 7). CP-based models perform similarly to the monthly-based approach. However, for higher percentiles above 95^{th} the monthly-based approach strongly overestimates precipitation compared to the CP-based one (Figure 6). Hence, the monthly based approach compensates the underestimation for lower percentiles with overestimation for higher percentiles resulting in a better mean. We thus consider the CP-based approach to be better particularly with regards to disaggregation of extreme precipitation, which plays an important role in many practical applications in water management, such as flood design estimation and risk analysis.

The major limitation of the CP-based disaggregation is the stationarity assumption of the link between daily sums and subdaily distributions in each CP class. This assumption is challenged by the thermodynamic changes in the atmosphere, in particular increasing water vapor with increasing air temperature (Lenderink et al., 2017). The CPs in this study are classified by using only daily mean sea level pressure and seasonality (summer/winter), which does not account for this thermodynamic effect. Westra et al. (2013) examined the daily-to-subdaily disaggregation performance of MOF conditioned on a range of atmospheric covariates, such as air temperature and relative humidity, where the atmospheric covariates with the greatest influence on the sub-daily rainfall temporal pattern were identified by fitting a generalized additive model (GAM). They found that the temporal distribution of subdaily rainfall is sensitive to changes in atmospheric temperature. The maximum intensity of short-duration rainfall increased by 4.1 – 13.4% per degree change in air temperature for the maximum 6 min burst, and by 3.1 – 6.8% for the maximum 1 h burst. Rafatnejad et al. (2021) evaluated climate change impacts on extreme subdaily rainfall amounts by using MOF with inclusion of air temperature and other weather variables as influential factors, where the distances between target and candidate days were weighted by the correlation coefficients between conditional variables and rainfall series. Their results indicated an increase in the extremes, for instance in the mean and standard deviation of the 95th percentile. Therefore, one approach to further improve the CP-based rainfall temporal disaggregation model in reproducing subdaily rainfall extremes is to include additional conditional variables, in particular air temperature, into the CP classification, which is expected to represent the thermodynamic effect of climate warming on the link between daily rainfall totals and subdaily distributions. On the contrary, Breinl and Di Baldassarre (2019) used both precipitation and air temperature in a single matrix to compute the distances between the target day and candidate days and this approach delivered poorer performance with regards to subdaily rainfall extremes and inter-site correlation. So, the question remains, how to optimally incorporate additional weather variables to circulation patterns into the MOF disaggregation procedure and can be addressed in future research.

6. Conclusions

In this research, a multisite method of fragments-based rainfall temporal disaggregation model conditioned on circulation pattern (CP) classification is developed and applied for the German part of the Rhine river basin. Its performance in simulating standard rainfall statistics, spatial correlation, wet and dry spells features and extremes is examined and compared with the standard disaggregation procedure (monthly-based). The CP-based disaggregation shows good performance in representing standard rainfall statistics, including standard deviation, lag-1 autocorrelation, and fraction of wet hours, although the monthly-based method of fragments disaggregation performs slightly better with regard to mean. Both disaggregation procedures underestimate the mean duration of dry spells, while the CP-based models outperform monthly-based one in wet spell length estimation. The spatial correlation structure in terms of the inter-site correlation coefficients is well maintained by both procedures.

The CP-based rainfall disaggregation procedure significantly improves the simulation of rainfall extremes, especially for high percentiles. The performance gain may be explained by the improvement of CP classification in stratifying extreme rainfall features. It could be shown that model performance increases with the number of CP classes; this comes, however, at the costs of higher computational demands and higher uncertainty. The superior performance for rainfall extremes is a valuable improvement for many practical applications in water management, such as flood design estimation and risk analysis. In addition, the CP-based approach opens up the possibility of including climate change effects in generating subdaily rainfall series.

CRediT authorship contribution statement

Conceptualization: SV, XG, DVN, BM; Methodology: XG, KN, BW; Data curation and Coding: XG; Writing – original draft: XG; Writing – review & editing: SV, BM, DVN, KN and BW; Supervision: SV and BM

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Xiaoxiang Guan is sponsored by China Scholarship Council for his PhD research (Grant #: 202106710029). This study is carried out in the KAHR project (Grant #: 01LR2102F) and the contributions from the ClimXtreme projects FLOOD (Grant #: 01LP1903E) and LASLI (Grant #: 01LP1903A) all funded by the German Federal Ministry for Research and Education (BMBF) are acknowledged.

References

Anis, M.R. and Rode, M. 2015. A new magnitude category disaggregation approach for temporal high-resolution rainfall intensities. Hydrological Processes, 29(6), 1119-1128. doi: https://doi.org/10.1002/hyp.10227.

Beck, C. and Philipp, A. 2010. Evaluation and comparison of circulation type classifications for the European domain. Phys Chem Earth, 35(9-12), 374-387. doi: https://doi.org/10.1016/j.pce.2010.01.001.

Breinl, K. and Di Baldassarre, G. 2019. Space-time disaggregation of precipitation and temperature across different climates and spatial scales. Journal of Hydrology-Regional Studies, 21, 126-146. doi: https://doi.org/10.1016/j.ejrh.2018.12.002.

Breinl, K., Strasser, U., Bates, P. and Kienberger, S. 2017. A joint modelling framework for daily extremes of river discharge and precipitation in urban areas. Journal of Flood Risk Management, 10(1), 97-114. doi: https://doi.org/10.1111/jfr3.12150.

Carreau, J., Ben Mhenni, N., Huard, F. and Neppel, L. 2019. Exploiting the spatial pattern of daily precipitation in the analog method for regional temporal disaggregation. Journal of Hydrology, 568, 780-791. doi: https://doi.org/10.1016/j.jhydrol.2018.11.023.

Gutierrez-Magness, A.L. and McCuen, R.H. 2004. Accuracy evaluation of rainfall disaggregation methods. Journal of Hydrologic Engineering, 9(2), 71-78. doi: https://doi.org/10.1061/(asce)1084-0699(2004)9:2(71).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. 2020. The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999-2049. doi: https://doi.org/10.1002/qj.3803.

Huth, R., Beck, C. and Kucerova, M. 2016. Synoptic-climatological evaluation of the classifications of atmospheric circulation patterns over Europe. International Journal of Climatology, 36(7), 2710-2726. doi: https://doi.org/10.1002/joc.4546.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynova, M., Kysely, J. and Tveito, O.E. (2008) Trends and Directions in Climate Research. Gimeno, L., GarciaHerrera, R. and Trigo, R.M. (eds), pp. 105-152.

Khaliq, M.N. and Cunnane, C. 1996. Modelling point rainfall occurrences with the modified Bartlett–Lewis rectangular pulses model. Journal of Hydrology, 180, 109-138. doi: https://doi.org/10.1016/0022-1694(95)02894-3.

Koutsoyiannis, D. and Onof, C. 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. Journal of Hydrology, 246(1), 109-122. doi: https://doi.org/10.1016/S0022-1694(01)00363-8.

Kronenberg, R., Franke, J. and Bernhofer, C. 2012. Classification of daily precipitation patterns on the basis of radar-derived precipitation rates for Saxony, Germany. Meteorologische Zeitschrift, 21(5), 475 - 486. doi: https://doi.org/10.1127/0941-2948/2012/0343.

Lenderink, G., Barbero, R., Loriaux, J.M. and Fowler, H.J. 2017. Super-Clausius–Clapeyron Scaling of Extreme Hourly Convective Precipitation and Its Relation to Large-Scale Atmospheric Conditions. Journal of Climate, 30(15), 6037-6052. doi: https://doi.org/10.1175/jcli-d-16-0808.1.

Li, X., Meshgi, A., Wang, X., Zhang, J.J., Tay, S.H.X., Pijcke, G., Manocha, N., Ong, M., Nguyen, M.T. and Babovic, V. 2018. Three resampling approaches based on method of fragments for daily-to-subdaily precipitation disaggregation. International Journal of Climatology, 38, E1119-E1138. doi: https://doi.org/10.1002/joc.5438.

Lisniak, D., Franke, J. and Bernhofer, C. 2013. Circulation pattern based parameterization of a multiplicative random cascade for disaggregation of observed and projected daily rainfall time series. Hydrology and Earth System Sciences, 17(7), 2487-2500. doi: https://doi.org/10.5194/hess-17-2487-2013.

Lu, Y. and Qin, X.S. 2014. Multisite rainfall downscaling and disaggregation in a tropical urban area. Journal of Hydrology, 509, 55-65. doi: https://doi.org/10.1016/j.jhydrol.2013.11.027.

Lu, Y., Qin, X.S. and Mandapaka, P.V. 2015. A combined weather generator and K-nearest-neighbour approach for assessing climate change impact on regional rainfall extremes. International Journal of Climatology, 35(15), 4493-4508. doi: https://doi.org/10.1002/joc.4301.

Mezghani, A. and Hingray, B. 2009. A combined downscaling-disaggregation weather generator for stochastic generation of multisite hourly weather variables over complex terrain: Development and multi-scale validation for the Upper Rhone River basin. Journal of Hydrology, 377(3-4), 245-260. doi: https://doi.org/10.1016/j.jhydrol.2009.08.033.

Mohr, S., Ehret, U., Kunz, M., Ludwig, P., Caldas-Alvarez, A., Daniell, J.E., Ehmele, F., Feldmann, H., Franca, M.J., Gattke, C., Hundhausen, M., Knippertz, P., Küpfer, K., Mühr, B., Pinto, J.G., Quinting, J., Schäfer, A.M., Scheibel, M., Seidel, F. and Wisotzky, C. 2022. A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis. Nat. Hazards Earth Syst. Sci. Discuss., 2022, 1-44. doi: https://doi.org/10.5194/nhess-2022-137.

Müller, H. and Haberlandt, U. 2018. Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. Journal of Hydrology, 556, 847-864. doi: https://doi.org/10.1016/j.jhydrol.2016.01.031.

Nguyen, V.D., Merz, B., Hundecha, Y., Haberlandt, U. and Vorogushyn, S. 2021. Comprehensive evaluation of an improved large-scale multi-site weather generator for Germany. International Journal of Climatology, 41(10), 4933-4956. doi: https://doi.org/10.1002/joc.7107.

Philipp, A., Della-Marta, P.M., Jacobeit, J., Fereday, D.R., Jones, P.D., Moberg, A. and Wanner, H. 2007. Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering. Journal of Climate, 20(16), 4065-4095. doi: https://doi.org/10.1175/jcli4175.1.

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B. and Jeremiah, E. 2012. A comparison of alternatives for daily to sub-daily rainfall disaggregation. Journal of Hydrology, 470-471, 138-157. doi: https://doi.org/10.1016/j.jhydrol.2012.08.041.

Rafatnejad, A., Tavakolifar, H. and Nazif, S. 2021. Evaluation of the climate change impact on the extreme rainfall amounts using modified method of fragments for sub-daily rainfall disaggregation. International Journal of Climatology, 42(2), 908-927. doi: https://doi.org/10.1002/joc.7280.

Sharma, A. and Srikanthan, S. 2006 Continuous Rainfall Simulation: A Nonparametric Alternative. In: 30th Hydrology and Water Resources Symposium, Launceston, Tasmania, 4–7 December, 2006.

Ullrich, S.L., Hegnauer, M., Nguyen, D.V., Merz, B., Kwadijk, J. and Vorogushyn, S. 2021. Comparative evaluation of two types of stochastic weather generators for synthetic precipitation in the Rhine basin. Journal of Hydrology, 601, 126544. doi: https://doi.org/10.1016/j.jhydrol.2021.126544.

Vallorani, R., Bartolini, G., Betti, G., Crisci, A., Gozzini, B., Grifoni, D., Iannuccilli, M., Messeri, A., Messeri, G., Morabito, M. and Maracchi, G. 2018. Circulation type classifications for temperature and precipitation stratification in Italy. International Journal of Climatology, 38(2), 915-931. doi: https://doi.org/10.1002/joc.5219.

Westra, S., Evans, J.P., Mehrotra, R. and Sharma, A. 2013. A conditional disaggregation algorithm for generating fine time-scale rainfall data in a warmer climate. Journal of Hydrology, 479, 86-99. doi: https://doi.org/10.1016/j.jhydrol.2012.11.033.

Westra, S., Mehrotra, R., Sharma, A. and Srikanthan, R. 2012. Continuous rainfall simulation: 1. A regionalized subdaily disaggregation approach. Water Resources Research, 48(1). doi: https://doi.org/10.1029/2011WR010489.

Winter, B., Schneeberger, K., Dung, N.V., Huttenlau, M., Achleitner, S., Stötter, J., Merz, B. and Vorogushyn, S. 2019. A continuous modelling approach for design flood estimation on sub-daily time scale. Hydrological Sciences Journal, 64(5), 539-554. doi: https://doi.org/10.1080/02626667.2019.1593419.