A non-stationary climate-informed weather generator for assessing of future flood risks

Viet Dung Nguyen¹, Sergiy Vorogushyn², Katrin M. Nissen³, Lukas Brunner⁴, and Bruno Merz⁵

¹Helmholtz-Zentrum Potsdam - Deutsches Geoforschungszentrum
²Unknown
³Freie Universität Berlin
⁴University of Hamburg
⁵Deutsches GeoForschungsZentrum

June 03, 2024

Abstract

We present a novel non-stationary Regional Weather Generator (nsRWG) based on an auto-regressive process and marginal distributions conditioned on climate variables. We use large-scale circulation patterns as a latent variable and regional daily mean temperature as a covariate for marginal precipitation distributions to account for dynamic and thermodynamic changes in the atmosphere, respectively. Circulation patterns are classified using ERA5 reanalysis mean sea level pressure fields. We set up nsRWG for the Central European region using data from the E-OBS dataset, covering major river basins in Germany and riparian countries. nsRWG is meticulously evaluated, showing good results in reproducing at-site and spatial characteristics of precipitation and temperature. Using time series of circulation patterns and the regional daily mean temperature derived from General Circulation Models (GCMs), we inform nsRWG about the projected future climate. In this approach, we utilize GCM output variables, such as pressure and temperature, which are typically more accurately simulated by GCMs than precipitation. In an exemplary application, nsRWG statistically downscales precipitation from nine CMIP6 GCMs generating a long synthetic but spatially and temporally consistent weather series. The results suggest an increase in extreme precipitation over the German basins, aligning with previous regional analyses. nsRWG offers a key benefit for hydrological impact studies by providing long-term (thousands of years) consistent synthetic weather data indispensable for the robust estimation of probability changes of hydrologic extremes such as floods.

Hosted file

A non-stationary climate-informed weather generator for assessing of future flood risks

Viet Dung Nguyen¹, Sergiy Vorogushyn¹, Katrin Nissen², Lukas Brunner³ and Bruno Merz¹,⁴

¹ GFZ German Research Centre for Geosciences, Section Hydrology, 14473 Potsdam, Germany
² Institute of Meteorology, Free University, Berlin, Germany
³ Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria
now at: Research Unit Sustainability and Climate Risk, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, 20144 Hamburg, Germany
⁴ University of Potsdam, Institute of Environmental Science and Geography, Potsdam, Germany

Correspondence and requests for materials should be addressed to Viet Dung Nguyen (email: dung@gfz-potsdam.de)

Abstract

We present a novel non-stationary Regional Weather Generator (nsRWG) based on an auto-regressive process and marginal distributions conditioned on climate variables. We use large-scale circulation patterns as a latent variable and regional daily mean temperature as a covariate for marginal precipitation distributions to account for dynamic and thermodynamic changes in the atmosphere, respectively. Circulation patterns are classified using ERA5 reanalysis mean sea level pressure fields. We set up nsRWG for the Central European region using data from the E-OBS dataset, covering major river basins in Germany and riparian countries. nsRWG is meticulously evaluated, showing good results in reproducing at-site and spatial characteristics of precipitation and temperature. Using time series of circulation patterns and the regional daily mean temperature derived from General Circulation Models (GCMs), we inform nsRWG about the projected future climate. In this approach, we utilize GCM output variables, such as pressure and temperature, which are typically more accurately simulated by GCMs than precipitation. In an exemplary application, nsRWG statistically downscales precipitation from nine CMIP6 GCMs generating a long synthetic but spatially and temporally consistent weather series. The results suggest an increase in extreme precipitation over the German basins, aligning with previous regional analyses. nsRWG offers a key benefit for hydrological impact studies by providing long-term (thousands of years) consistent synthetic weather data indispensable for the robust estimation of probability changes of hydrologic extremes such as floods.

Keywords: non-stationary weather generator, circulation pattern, future projections, statistical downscaling
1. Introduction

Reliable climate and hydrological hazard and risk assessments require long time series of meteorological fields such as precipitation and temperature at regional scales. Despite recent advancements in observation technologies, the observed climatic records still remain relatively short for estimating the probability of extreme and rare events over large scales, e.g., at the scale of large river basins and entire countries. They represent a single realization of climate conditions within the range of possibilities due to natural variability. General Circulation Models (GCMs) can provide several realizations of continuous meteorological fields, but often have too coarse spatial resolution to be suitable for hydrological impact studies. Whereas Regional Climate Models (RCMs) provide sufficiently high resolution, their ensembles are limited compared to GCMs and require considerable production time. Given the computational constraints of climate models, there is a severe trade-off between the available number of climate model realizations and their spatial resolution. This is where stochastic weather generators become pivotal, as they allow to extend the time series while preserving the essential statistical properties of meteorological fields (Nguyen et al., 2021; Papalexiou et al., 2023). A weather generator (WG) is a stochastic model capable of generating long-term synthetic meteorological fields, e.g., precipitation, temperature, that have the temporal (e.g., autocorrelation) and spatial (e.g., spatial covariance) statistical properties of the fields on which the weather generator was conditioned. These fields can be provided by meteorological observations or physically-based climate models. In-depth overviews of the different types of WGs can be found in Haberlandt et al. (2011), Serinaldi and Kilsby (2014) with a subsequent update by Nguyen et al. (2021).

From the perspective of climate model downscaling, WGs represent a special type of statistical downscaling, which is not only able to bridge the scale gap and increase the spatial resolution of output fields, but also to extend the time series by generating synthetic fields of arbitrary length (Maraun et al., 2010). Particularly for flood design and risk assessment, long time series are essential for the robust estimation of high flood quantiles and associated risks. In the Derived Flood Frequency Analysis (DFFA), weather generators conditioned on past meteorological observations have been successfully applied in combination with rainfall-runoff models for estimating flood quantiles (Blázkova and Beven, 1997, Grimaldi et al., 2012, Haberland and Radtke, 2014, Winter et al., 2019). Further extended these model chains by including flood inundation and damage models have been used to estimate flood risks at catchment (Falter et al., 2015, 2016, Metin et al., 2018) and national scales (Sairam et al., 2021). The latter approach extends DFFA towards Derived Flood Risk Analysis (DFRA) (Falter et al., 2015). When conditioned on observed or simulated meteorological fields, WGs typically assume stationarity, i.e., entire time series of meteorological variables used to parameterize the marginal, at-site distribution functions and the spatial dependency are assumed to
be stationary. In this case, the long synthetic time series (e.g., 10,000 years of daily rainfall) represents a climate realization with statistical properties we would experience if we lived in this stationary climate for 10,000 years. Hence, the derived flood risk would represent the risk associated with this climate state. Under ongoing climate change, the assumption of stationary may no longer be valid for several atmospheric variables and is certainly no longer valid for temperature (IPCC, 2023). Hence, the assessment of future risks requires novel approaches that consider the non-stationarity of the climate and associated meteorological variables.

Assessments of future flood risks typically rely on climate model projections of precipitation, temperature and other relevant weather variables. Projections of precipitation, in particular of extremes, strongly depend on climate model resolution and are inherently inferior for coarser resolution models (Torma et al., 2015, Jong et al., 2023, Hohenegger et al., 2023). Projections from high-resolution RCMs are much less available and their production is delayed compared to the availability of GCMs. In addition, projections of extreme precipitation are strongly controlled by the GCM ensemble rather than by the selection of RCMs (Fowler et al., 2007). Hence, there is a need for weather generators and other statistical downscaling approaches that make timely use of large GCM ensembles to provide robust downscaling of weather variables for risk assessments of future periods. Moreover, impact and risk attribution studies require climate model runs without anthropogenic forcing, which are typically available only from coarse-resolution GCMs (Eyring et al., 2016). Simulation of precipitation by GCMs has been found to be poor compared to pressure and temperature fields (Johnson and Sharma, 2009). Thus, GCM precipitation output can hardly be directly used for conditioning WGs. On the other hand, indices of large-scale circulation dynamics are simulated more accurately than regional precipitation extremes (Farnham et al., 2018). The skill of CMIP6 GCMs to simulate large-scale circulation patterns (CPs) has improved compared to the previous CMIP5 ensemble (Cannon, 2020, Fernandez-Granja et al., 2021). Hence, there is a need for weather generator approaches that establish a link between robustly simulated large-scale atmospheric variables and locally variable precipitation.

Precipitation changes over time are controlled by changes in circulation dynamics, i.e., frequency and persistence of CPs, and by changes in thermodynamic properties of the atmosphere, including enhanced evapotranspiration and water holding capacity of the warmer air. Depending on the timescale (daily, monthly), season and location, a variable relative importance of dynamic and thermodynamic controls on past precipitation trends has been detected (Beck et al., 2007, Fleig et al., 2015, Cahynová and Huth, 2016). Changes in extreme rainfalls are particularly sensitive to thermodynamic changes compared to low-quantile precipitation (Haerter et al., 2010, Berg et al., 2013). On the other hand, Shepherd (2014) argued that most
of the uncertainty in climate model projections comes from the circulation dynamics which is sensitive to the
model forcing and chaotic variability. Pfahl et al. (2017) demonstrated this particularly for extreme
precipitation. They showed that the thermodynamic component of changes in extreme precipitation is
robust, i.e., consistent across models. However, this change is modulated by the dynamic component and
this effect is not consistent across models. This can even invert the sign of the precipitation change.
Therefore, both components of climatic change – dynamic and thermodynamic – need to be taken into
account when downscaling climate models in order to obtain the overall trend and assess the uncertainty of
hydrological changes.

To account for the non-stationarity introduced by climate change in WG-based downscaling, two main
approaches have been developed as discussed in the reviews by Wilks (2010, 2012). The first approach
adjusts the WG parameters for daily rainfall, temperature, etc. based on monthly change factors in mean
and variance inferred from climate model projections. The change factors can be static or vary gradually over
time (Wilks, 1992). Waswo and Sharma (2017), for example, correlated the parameters of a WG to changes in
mean monthly temperatures from climate model projections to simulate sub-daily rainfall for two stations in
Australia. Kiem et al. (2021) used annual maximum daily temperature to condition a WG for monthly to
annual rainfall simulation for Australia. This approach assumes a link between monthly mean changes and
changes in daily variability and extremes of specific meteorological variables. It also assumes that changes at
the GCM-grid scale (>100km) proportionally translate to the finer scale of a few kilometers. The approach
also relies on a robust simulation of, for example, monthly precipitation and the change signal by the GCMs.
Conditioning WGs solely on temperature changes tacitly considers only the thermodynamic climate change
signal.

The second approach conditions WGs on circulation patterns, either by adjusting the wet/dry day
probabilities and the mean of the simulated variable to large-scale circulation indices using regression
equations or by fitting different sets of WG parameters to groups of days characterized by specific circulation
patterns (Wilks, 2010). Conditioning weather generators on large-scale CPs has been explored by several
studies, e.g., Bárdossy and Plate (1992) and Fowler et al. (2005) to name a few. Ailliot et al. (2015) provide a
brief overview of weather-pattern-based WG approaches. Also, Haberlandt et al. (2015) approached
precipitation downscaling in a non-stationary climate with their WG conditioned on CPs derived by
simulated annealing. One of the fundamental assumptions is that climate change is only manifested in
changes in atmospheric dynamics. For a case study in Northern Germany, Haberlandt et al. (2015) showed
that the largest portion of precipitation change is actually driven by changes of the precipitation distribution
within individual patterns, i.e., by thermodynamic changes. Even the inclusion of predictors such as humidity
and temperature in the weather pattern classification in addition to pressure variables (with these additional predictors one no longer refers to circulations patterns, but to weather patterns) could not explain past trends in precipitation in the Rhine basin in Germany (Murawski et al., 2016, 2018). Recently, Steinschneider et al. (2019) suggested an integrated concept of conditioning a semi-parametric WG on both dynamic and thermodynamic changes in the atmosphere. They used a Markov chain-based simulation of circulation patterns combined with a block-bootstrapping to generate daily meteorological variables (Steinschneider and Brown, 2013). To extend the variability range of generated meteorological fields beyond the observed one, copula-based jitters were introduced while preserving the multi-site correlation and temporal persistence. The simulation of circulation patterns and associated daily meteorological fields can be perturbed based on ENSO dynamics as well as Clausius-Clapeyron precipitation scaling and elevation dependent warming, respectively. The authors explored the sensitivity of temperature and precipitation to these perturbations (Steinschneider et al., 2019), and Rahat et al. (2022) further assessed climate change scenarios for the Tuolumne River basin in the western US based on the proposed approach for the management of a downstream reservoir.

To overcome the limitations of stationary parameterizations described above and to leverage large and timely GCM ensembles, we further develop the idea of conditioning a stochastic weather generator on GCM-based indices of dynamic and thermodynamic changes. To this end, we simultaneously use large-scale circulation patterns as a latent discrete variable and the average regional air temperature as a covariate for marginal non-stationary precipitation distributions in a multi-site auto-regressive WG. The marginal distributions are parameterized using observed precipitation and temperature data, while the circulation patterns are derived from atmospheric reanalysis. When applied to GCM model projections, our approach takes advantage of the strength of GCMs to reproducing these properties more reliably than local precipitation. Considering the two fundamental controls of change in the climate system – dynamic and thermodynamic – may allow to disentangle future flood changes into those driven by dynamic changes, i.e., changes in frequency and persistence of circulation patterns, and by thermodynamic changes, i.e., changes due to increasing regional temperature.

This paper presents this novel methodological approach for a non-stationary weather generator conditioned on circulation patterns and regional daily mean temperature for the purpose of flood risk estimation and flood impact attribution. We demonstrate an implementation of the weather generator for a domain in Central Europe followed by a comprehensive evaluation of the presented approach. Finally, changes in downscaled precipitation from climate model projections are analyzed in preparation of subsequent analyses of changes in flood risk.
2. Study area and data

The weather generator is set up for a domain between 45.125°N and 55.125°N latitude and 5.125°E to 19.125°E longitude (Figure 1). This domain encompasses the five major river basins in Germany – Rhine, Danube, Elbe, Weser and Ems – which are targeted for flood risk assessment in future studies. This domain is further termed ‘German domain’ and covers more than 650,000 km². Given the size of the domain we further speak of regional weather generator. The regional WG is set up based on two types of meteorological data: (1) small-scale observational data to calibrate the weather generator and (2) synoptic-scale reanalysis data used to characterize circulation dynamics. Both datasets are used in daily resolution, spanning from January 1st, 1950 to December 31st, 2021.

We use the E-OBS dataset version 25.0e (Cornes et al., 2018), which contains gridded observed mean daily temperature and precipitation totals. For the German domain, 540 grid cells with a spatial resolution of 0.5° x 0.5° are selected for parameterizing the WG after remapping the E-OBS data (Figure 1). To derive the daily series of circulation patterns over Europe to be used as a latent variable for the WG, we adopt the ERA5 dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach et al., 2020). We use mean sea level pressure (MSLP) for the region from 25°N to 70°N latitude and from 15°W to 30°E longitude (same as Nied et al., 2014), encompassing a substantial portion of Europe and adjacent regions (‘European domain’) (Figure 1). The pressure data is aggregated to a spatial resolution of 1° x 1° prior to CP classification. Additionally, we extract the mean daily 2-meter air temperature grid (t2m) for the German domain and aggregate it to the mean daily regional temperature to be used as a covariate in the WG.
Figure 1. (a) European domain (35°N – 70°N and 15°W – 30°E) used for circulation pattern classification and (b) German domain (45.125°N – 55.125°N and 5.125°E – 19.125°E) covering the five major river basins in Germany (Danube, Elbe, Rhine, Weser and Ems) for which the nsRWG is set up.

In this study, we apply the proposed regional WG for developing synthetic weather series for future climate scenarios. We condition the WG on CPs and regional temperature derived from an array of GCMs included in the CMIP6 (Coupled Model Intercomparison Project Phase 6) (Eyring et al., 2016). We preselected 15 CMIP6 GCMs (Table 1) that have been evaluated by Cannon (2020) with regards to their ability to reproduce circulation dynamics prior to subsequent screening and weighting (section 3.5).

Table 1. CMIP6 General Circulation Models used for subsequent screening and weighting in the ClimWIP approach.

<table>
<thead>
<tr>
<th>Modelling center</th>
<th>GCM</th>
<th>Realization</th>
<th>Spatial resolution (in °)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing Climate Center</td>
<td>BCC-CSM2-MR</td>
<td>r1i1p1f1</td>
<td>1.13</td>
</tr>
<tr>
<td>Canadian Centre for Climate Modelling and Analysis</td>
<td>CanESM5</td>
<td>r1i1p1f1</td>
<td>2.81</td>
</tr>
<tr>
<td>National Center for Atmospheric Research</td>
<td>CESM2</td>
<td>r1i1p1f1</td>
<td>1.10</td>
</tr>
<tr>
<td>Centre National de Recherches Météorologiques and Centre Européen de Recherche et de Formation Avancée</td>
<td>CNRM-CM6-1</td>
<td>r1i1p1f2</td>
<td>1.41</td>
</tr>
</tbody>
</table>
The daily MSLP grids for the historical period (1985-2014) and two future periods 2031-2060 (near future) and 2071-2100 (far future) are extracted and remapped to a consistent resolution of 1°x1°. Further, mean daily t2m grids are extracted for the same periods and regional daily mean temperature is computed for the German domain. We consider two shared socio-economic pathways, SSP245 and SSP585, for the generation of synthetic weather series. The SSP245 pathway represents a middle-of-the-road emissions trajectory marked by moderate attempts to address greenhouse gas emissions. It strives for a harmonious balance between economic growth and environmental sustainability, envisioning a world where society aims for equitable socio-economic progress while acknowledging the paramount importance of climate action. In contrast, the SSP585 pathway portrays a scenario of exceedingly high emissions with limited mitigation efforts, where economic expansion and fossil fuel usage continue to surge, casting a shadow over environmental concerns.

### 3. Methods

#### 3.1. Circulation pattern classification

We employ a CP classification approach based on the SANDRA (Simulated ANnealing and Diversified RAndomization) objective classification algorithm (Philipp et al., 2007, Philipp et al., 2016). This method is based on k-means clustering and is designed to minimize the within-cluster variance of the Euclidean distance between the cluster elements and their respective centroids. To circumvent the limitations of conventional k-means, which often converges to local optima, the SANDRA algorithm introduces random...
reassignments of cluster elements, facilitating the search for the global optimum. Beck and Phillip (2010) and Philip et al. (2016) found a very good performance of SANDRA compared to other classification algorithms. For use with a weather generator, it is desirable to have a CP classification, in which different CP classes have as distinct local weather characteristics as possible, e.g., precipitation distributions for various CPs differ strongly. This is typically achieved with a higher number of classes (Murawski et al., 2016). However, we need to ensure that the WG has sufficient data to robustly parameterize the distributions of the weather variables for each class. With an increasing number of classes, less data is available within each class. Therefore, we test the SANDRA classifications with 4, 5, 6, 7 and 8 classes. To capture the seasonality in precipitation, we consider two distinct seasons for each CP: winter (November-April) and summer (May-October). Hence, in total we test classifications with 8, 10, 12, 14 and 16 classes.

We evaluate the stratification of observed precipitation for each classification using two metrics: explained variation (EV) and Pseudo-F statistic (PF). EV is defined as the ratio of the sum of squared deviations from the mean within classes to the total sum of squared deviations from the overall mean. The Pseudo-F statistic is the ratio of the sum of squared deviations between the class means to the mean within classes, weighted by the number of classes and cases (i.e., days). Values close to zero indicate poor stratification. EV = 1 indicates perfect stratification. Higher values for both metrics indicate better stratification. More details about these metrics can be found in Beck and Philipp (2010) and Murawski et al. (2016).

3.3. Multi-site non-stationary weather generator

In this study, we introduce a non-stationary version of the Regional Weather Generator (nsRWG) building upon the original stationary model developed by Hundecha et al. (2009) and further refined by Nguyen et al. (2021). Like its predecessors, nsRWG is a multi-variate auto-regressive (MAR-1) model designed to simulate daily weather variables including both precipitation and temperature. The non-precipitation variables are conditioned on the wet/dry state of the respective day according to the precipitation generated. Contrary to the stationary versions (Nguyen et al., 2021), the precipitation generation in nsRWG is now conditioned on CPs as a latent variable characterizing changes in atmospheric dynamics. Additionally, the mean regional daily mean temperature characterizing thermodynamic changes is used as a covariate for the marginal non-stationary probability distributions.

3.3.1 Spatio-temporal dependence model

The multi-variate auto-regressive model follows Bárdossy and Plate (1992): Let $W(t) = (W(t,u_1),...,W(t,u_n))$ be a multi-variate standard normal random vector of n locations $u = u_1,..., u_n$ at day $t$ with the zero mean. For the moment, we can think of $W(t)$ as standardized
precipitation at day $t$. In the next step, we introduce circulation pattern $CP_i$ as a latent variable, where $i$ is the circulation pattern index. We further extend the approach of Bárdossy and Plate (1992) for the estimation of the state variable by considering the transition between all pairs of circulation patterns. The MAR-1 model for the day $t$ which is characterized by $CP_i$ reads now as follows:

$$W(t) = \begin{cases} B_i W(t-1) + C_i \Psi(t), & \text{if day } t-1 \text{ has the same } CP_i \\ B_i W(t-1) + \overline{C}_i \Psi(t), & \text{if day } t-1 \text{ has a different } CP \\ D_i \Psi(t), & \text{if } t = 1 \text{ (starting day)} \end{cases}$$

(1)

where $\Psi(t) = (\psi(t,u_1), ..., \psi(t,u_n))$ is a random vector of the independent standard normal variable. The matrices $B_i, \overline{B}_i, C_i, \overline{C}_i$ and $D_i$ are related to the lag-0 correlation matrix ($M_{i0}$), the lag-1 correlation matrix ($M_{i1}$) within a single $CP_i$ and the lag-1 correlation matrix ($\overline{M}_{i1}$) for the transition between $CPs$ other than $CP_i$ to $CP_i$:

$$B_i = M_{i1} M_{i0}^{-1}$$

(2)

$$\overline{B}_i = \overline{M}_{i1} \overline{M}_{i0}^{-1}$$

(3)

$$C_i \overline{C}_i^T = M_{i0} - B_i \overline{M}_{i1}^T$$

(4)

$$\overline{C}_i \overline{C}_i^T = M_{i0} - \overline{B}_i \overline{M}_{i1}^T$$

(5)

$$D_i \overline{D}_i^T = M_{i0}$$

(6)

where the superscripts $-1$ and $T$ indicate the matrix inversion and matrix transpose operator, respectively.

The introduction of $\overline{M}_{i1}, \overline{B}_i$ and $\overline{C}_i$ represents an enhancement compared to the previous works by Bárdossy and Plate (1992), Hundecha et al. (2009) and Nguyen et al. (2021) when dealing with frequently alternating circulation patterns. Previous studies re-initialized the precipitation state every time when there is shift in a latent variable state, i.e., shift in CP (Bárdossy and Plate, 1992) or change between months (Hundecha et al., 2009, Nguyen et al., 2021).

We estimate the correlation matrices ($M_{i0}, M_{i1}$ and $\overline{M}_{i1}$) through Kendall correlation and then transform them into Pearson’s correlation (Serinaldi and Kilsby, 2014, Nguyen et al. 2021). To correct for poorly defined (not positive definite) matrices, we use the method of Higham (2002) to find the nearest positive definite correlation.
### 3.3.2. Marginal distributions

We use the 3-parameter Extended Generalized Pareto Distribution (Naveau et al. 2016, Nguyen et al., 2021) to model the non-zero daily precipitation distribution at every location \( u \) for each CP, i.e., marginal distribution. The cumulative probability distribution function for non-zero precipitation \( x \) is given by:

\[
F(x(t, u)) = \begin{cases} 
(1 - (1 + \sigma(t, u)^{-1}\xi x(t, u))^{-1})^k & \text{for } \xi > 0 \\
(1 - \exp(-\sigma(t, u)^{-1}x(t, u)))^k & \text{for } \xi = 0 
\end{cases}
\]  

(7)

where \( \kappa \) controls the shape of the lower tail, \( \sigma \) is a scale parameter and \( \xi > 0 \) controls the decay rate of the upper tail. The scale parameter \( \sigma \) is allowed to covary with the regional daily mean temperature (Equation 8). The exponential function is used to ensure that the scale parameter is positive:

\[
\sigma(t, u) = \exp(\sigma_0(u) + t\cdot m(t) \cdot \sigma_1(u)) 
\]

(8)

We employ the SCE-UA global optimization algorithm (Duan et al., 1992) to estimate the parameters \( \kappa, \xi, \sigma_0 \) and \( \sigma_1 \) by optimizing the log-likelihood function.

The complete precipitation process including wet and dry conditions with non-negative precipitation \( x \) for a certain CP at an individual location \( u \) is modelled using the cumulative distribution \( H(x) \):

\[
H(x(t, u)) = \begin{cases} 
(1 - p(u)) + p(u) \cdot F(x(t, u)) & x(t, u) > 0 \\
1 - p(u) & x(t, u) = 0 
\end{cases}
\]

(9)

where \( p \) represents the wet frequency and \( (1 - p) \) stands for the probability of zero rainfall. The link between the marginal distribution of precipitation (Equation 9) and the MAR-1 model (Equation 1) is given by:

\[
\Phi(W(t, u)) = H(x(t, u)) 
\]

(10)

where \( \Phi \) stands for the cumulative distribution function of a standard normal distribution.

For simulating daily average temperature fields consistent with non-stationary precipitation fields, we condition marginal temperature distributions for each month on the wet/dry state of precipitation. Similar to Nguyen et al. (2021), we apply a normal distribution to model daily temperature data. To accommodate the non-stationary change due to increasing regional temperature, we use it as a covariate for the location parameter of the non-stationary normal distribution:

\[
\mu(t, u) = \mu_0(u) + \mu_1(u) \cdot t \cdot m(t) 
\]

(11)

where \( \mu \) is location parameter of the normal distribution and parameters \( \mu_0 \) and \( \mu_1 \) are estimated by optimizing the log-likelihood function analogously to the parameters of the marginal precipitation.
distribution. Finally, temperature fields are also simulated using a multi-variate MAR-1 model analogously to precipitation.

3.4. Model setup and performance evaluation

The nsRWG is set up for the study area using the observed gridded dataset E-OBS v25.0e. The model is calibrated for 540 grid cells. For each combination of CP and season (winter/summer), 100 realizations are generated with a time series length of 72 years, the same length as the observed data. Synthetic and observed climate are compared using several statistical metrics introduced below. This evaluation procedure is commonly applied to assess stochastic weather models (Kleiber et al., 2012; Breinl et al., 2013; Serinaldi and Kilsby, 2014; Baxevani and Lennartsson, 2015; Nguyen et al., 2021).

In the evaluation process, special attention is given to both the local and spatial model performance. For the local evaluation, the following at-site metrics are computed and compared with observations:

- **Precipitation intermittence properties**: wet-day frequency and four transition probabilities (wet-to-wet, wet-to-dry, dry-to-wet and dry-to-dry). We consider days to be dry if the recorded daily precipitation is below 0.1 mm.
- **Daily precipitation for each CP**: mean and 99.5\textsuperscript{th}-percentile.
- **Seasonal precipitation sum for each CP**: mean and 98\textsuperscript{th}-percentile.

Both 99.5\textsuperscript{th} and 98\textsuperscript{th}-percentiles are estimated using semi-parametric quantile estimation proposed by Hutson (2002).

- **n-day maxima**: total precipitation for \( n = 5 \) and 10 days is compared to the observed statistics to analyze the plausibility of wet-spell precipitation amounts. We consider these durations to be important for the generation of flood events by single cyclones and for flood events resulting from subsequent storms, with the preceding storms contributing to the catchment wetness.

The ability of nsRWG to reproduce *daily average temperature* is assessed for each month by comparing the observed mean and 99.5\textsuperscript{th}-percentile values from the observation period to the simulated values.

To evaluate the spatial representation of precipitation fields, for each circulation pattern and season, we examine:

- **Correlation of precipitation as a function of distance between pairs of locations**: lag-0 \((M_0)\), lag-1 \((M_1)\) for the transition between days with the same CP, and lag-1 \((\bar{M}_1)\) for the transition between days characterized by different CPs.
• **Catchment areal precipitation**: 99.5\textsuperscript{th} percentile of catchment average precipitation for the five major river basins in Germany.

For a consistent and comparable assessment of model performance, we adopt the evaluation and performance framework (CASE) for weather generators proposed by Bennett et al. (2018). In the first step, the at-site performance is assessed for each at-site metric. The performance is categorized as "good" (G), "fair" (F), and "poor" (P) at each location i.e., grid cell. Model performance is considered "good" if the observed metric falls within the 90% range of the metric values computed for 100 model realizations. "Fair" performance is assigned when the observations are outside the 90% range but within the 99.7% limits, or if the absolute relative difference (RD) between the observed and simulated metric means is 5% or less. "Poor" performance is indicated when neither of these conditions is met. RD is defined as follows:

\[
RD = \left| \frac{M_{obs} - M_{sim}}{M_{obs}} \right| \times 100
\]  

(13)

where \(M_{obs}\) is metric value based on observations and \(M_{sim}\) is the mean metric value based on simulated data.

The overall performance is assessed by computing the share of sites exhibiting good, fair and poor performance. Overall performance is classified into 6 categories (Bennett et al., 2018):

• "Overall good" if more than half of the locations show good performance,
• "Overall fair" if more than half of the locations show fair performance,
• "Overall poor" if more than half of the locations show poor performance,
• "Overall fair-good" if the total percentage of fair and good performance locations exceeds the percentage of locations with poor performance,
• "Overall fair-poor" if the total percentage of fair and poor performance locations exceeds the percentage of locations with good performance,
• "Overall variable" if the total percentage of good and poor performance locations exceeds the percentage of locations with fair performance.

### 3.5. Downscaling future precipitation

To showcase the practical value of the newly developed weather generator, we employ it to downscale precipitation for future climate projections from nine global climate models. The GCMs are selected from the list of models in Table 1 using a mixture of qualitative and quantitative considerations about their performance in simulating European climate and taking into account their independence. The selection is
carried out to reduce the computational load for the nsRWG and subsequent future climate impact studies on flood risk change. The main guide for the selection are the model performance weights calculated using the ClimWIP (Climate model Weighting by Independence and Performance) method (Brunner et al. 2020) as implemented in the ESMValTool (Earth System Evaluation Tool) version v2.6.0 (Eyring et al. 2020; https://docs.esmvaltool.org/en/latest/recipes/recipe_climwip.html). As performance metrics we use the models’ distance to ERA5 in the European domain for the 1985-2014 climatology and annual variability of temperature and sea level pressure as well as the temperature trend. This follows the metric selection of Brunner et al. (2020) but targeted to Europe. We also ensure that the final selection of models follows the recommendations for model selection from the recent work by Merrifield et al. (2023) who considered model performance, independence, and spread as criteria. The resulting ClimWIP performance weights are summarized in Figure 2.

Figure 2. Performance weights for 15 GCMs (Table 1) resulting from the ClimWIP procedure based on the preselected evaluation criteria for the historical period 1985-2014.

Based on the resulting weights, we select nine models to be used in this study. The only exceptions are the two GFDL models from which we select only the (better performing) Earth System version even though they both received high weights. This is done to limit the inter-dependency in our model pool. Our selection includes UKESM1-0-LL, CanESM5, CESM2, CNRM-CM6-1, INM-CM5-0, MPI-ESM1-2-HR, MRI-ESM2-0, GFDL-
ESM4 and IPSL-CM6A-LR. For these models, also the independence weights are computed and can be used in future climate impact studies.

Overall, we investigate 36 distinct cases, which result from the combination of nine GCM models, two pathways, and two future periods. To this end, for each of these cases, we generate 100 realizations of daily time series with the nsRWG conditioned on respective circulation patterns and regional daily mean temperature corrected for bias with respect to ERA5 using quantile mapping (R-package qmap by Gudmundsson et al. (2012)).

4. Results and Discussion

4.1. Circulation pattern classification and mean regional temperature

The results of the classification of circulation patterns indicate relatively low values of EV below 0.1 and log(PF) between 2 and 2.3, in a similar range as in Murawski et al. (2016). The differences between classifications with different number of classes between 4 and 8 are relatively small, especially for cases with more than six classes. Therefore, we adopt the final classification with six CPs and consider this to be a good compromise between the degree of stratification (i.e., EV and PF) and the data available to estimate the marginal distributions. In total, the entire period is stratified into twelve classes (i.e., six CPs and two seasons), which is similar to a classical monthly-based parameterization used by Hundecha et al. (2012) and Nguyen et al. (2021), with the difference that days are unevenly distributed between classes. We also compare the EV and PF metrics between the CP-based classifications and the monthly-based classification (not shown). The CP-based classification clearly outperforms the monthly-based classification, although the overall values of EV and PF are low.

Figure 3 shows the mean sea level pressure (MSLP) of the six patterns of the selected classification. CP2 and CP4 are characterized by high-pressure systems covering a large region, particularly around Central Europe. The descending air associated with these high-pressure systems inhibits cloud formation and precipitation, leading to stable and dry weather conditions, often characterized by sunny days. CP3 includes the days with weak pressure gradients and is associated with moderate precipitation and weather conditions. CP1, CP5 and CP6 exhibit distinct low-pressure systems over large areas. As a result, these regions are prone to cloudy and wet weather conditions. CP5, in particular, shows an extremely low-pressure area in the North Atlantic region, creating a steep positive pressure gradient towards Central Europe. The steep pressure gradient can drive strong winds and lead to intense precipitation events in Central Europe. The winter, summer and annual frequencies of the CPs are summarized in Table 2. In winter, CP3 is the most frequent with 36.8%, while CP5 is the rarest with only 3.8%. In summer, CP2 takes the lead with 19.6%, and CP4 and CP6 are close
behind with 17.0% each. For the whole year, CP3 is the most frequent with 26.3%, followed by CP1 with 17.3%. CP5 is again the least frequent with 10.2%. These numbers give an idea of how different CPs play out over the seasons.

Table 2. Winter, summer and annual frequencies of the selected classification with six circulation patterns.

<table>
<thead>
<tr>
<th>CP / Season</th>
<th>CP1</th>
<th>CP2</th>
<th>CP3</th>
<th>CP4</th>
<th>CP5</th>
<th>CP6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>20.4%</td>
<td>12.2%</td>
<td>36.8%</td>
<td>11.8%</td>
<td>3.8%</td>
<td>15.0%</td>
</tr>
<tr>
<td>Summer</td>
<td>14.2%</td>
<td>19.6%</td>
<td>15.6%</td>
<td>17.0%</td>
<td>16.6%</td>
<td>17.0%</td>
</tr>
<tr>
<td>Annual</td>
<td>17.3%</td>
<td>15.9%</td>
<td>26.3%</td>
<td>14.3%</td>
<td>10.2%</td>
<td>16.0%</td>
</tr>
</tbody>
</table>

The regional average daily temperature over the period 1950-2021 shows a mean of 8.6°C and a standard deviation of 7.3°C. In this period, a significant positive trend of 0.27°C per decade (p-value: 4.3e-11) is detected based on the E-OBS data. The overall change in regional average annual temperature amounts to 1.9 °C since 1950. The pronounced increase in temperature underscores the potential relevance of thermodynamic changes in the atmosphere to be considered in the nsRWG parameterization.
Figure 3. (top) Selected circulation pattern classification with six patterns. The maps show the average mean sea level pressure of all the days falling into the same pattern. The red box in the maps shows the German domain. (middle) Mean daily precipitation observed for the six CPs during winter and summer seasons and (bottom) 99.5\textsuperscript{th}-percentile daily precipitation observed for six CPs during winter and summer seasons.

Figure 3 illustrates the stratification of daily precipitation intensity, examining the mean and 99.5\textsuperscript{th}-percentile of the classification based on the six CPs further divided into summer and winter. We observe a clear distinction between CPs in terms of mean and extreme precipitation. CP2 and CP4 stand out as rather dry patterns on average, whereas CP5 and CP1 are relatively wet in both seasons. CP3 and CP6 show average
wetness which slightly varies between seasons. CP1 and CP5 exhibit high extreme precipitation in both seasons. Additionally, CP3 and CP6 bring extreme rainfall in summer season throughout the German domain.

4.2. Marginal distribution fitting

We model the marginal precipitation distributions with the non-stationary extended Generalized Pareto (extGPD) distribution based on data stratified into twelve classes according to CPs and seasons. To assess the fitting performance of the non-stationary model, we compare it to the stationary version. The comparison is based on the Akaike Information Criterion (AIC) which considers model complexity and goodness of fit.

The non-stationary model provides a better fit for more than 50% of the cells for CP1, CP3, CP4 and CP6 in both seasons; for CP1, CP3 and CP4 even in over 75% of the cells (Figure 4). For CP2 characterized by rather dry conditions, the non-stationary model is better for about 70% of the cells in summer. In winter, the stationary model has a slight edge. For CP5 (wet conditions), the non-stationary model has a slight advantage in winter, but only 25% of the cells are better simulated in summer. Overall, the non-stationary model outperforms the stationary one, with on average better fitting for 70.5% of the model domain.

Figure 4. Difference in model performance of fitting marginal distributions of extGPD to precipitation data between the non-stationary and stationary models for six CPs and two seasons (left) and summarized for two seasons (right). A positive value indicates better performance for the non-stationary case.

4.3 Evaluation of nsRWG performance

The at-site and spatial performance of nsRWG is evaluated using the performance metrics described in Section 3.4. Table 3 summarizes the performance statistics which are discussed in detail in the following sections. Although a direct comparison with the performance of the stationary RWG model of Nguyen et al.
(2021) is not straightforward due to the different underlying datasets (gridded E-OBS vs. station-based), we discuss the nsRWG performance in the context of the stationary model evaluation.

### Table 3. Summary of the model performance statistics for nsRWG and categorization of the model performance according to Bennett et al. (2018). GFP [%] indicates the percentage of locations having “good”, “fair”, and “poor” performance considering all 100 realizations.

<table>
<thead>
<tr>
<th>Metric</th>
<th>GFP</th>
<th>Overall performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet frequency</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>Transitional probability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wet-wet (p11)</td>
<td>(81,13,6)</td>
<td>Good</td>
</tr>
<tr>
<td>wet-dry (p10)</td>
<td>(61,15,24)</td>
<td>Good</td>
</tr>
<tr>
<td>dry-wet (p01)</td>
<td>(51,18,31)</td>
<td>Good</td>
</tr>
<tr>
<td>dry-dry (p00)</td>
<td>(91,8,1)</td>
<td>Good</td>
</tr>
<tr>
<td>Transitional probability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily intensity for each CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>99.5th-percentile</td>
<td>(85,8,7)</td>
<td>Good</td>
</tr>
<tr>
<td>Seasonal sum for each CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>99th-percentile</td>
<td>(91,7,2)</td>
<td>Good</td>
</tr>
<tr>
<td>n-day maxima</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-day sum</td>
<td>(84,12,4)</td>
<td>Good</td>
</tr>
<tr>
<td>10-day sum</td>
<td>(84,12,4)</td>
<td>Good</td>
</tr>
<tr>
<td>Inter-site correlation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M_0 ) (lag-0)</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>( M_1 ) (within-type with lag-1)</td>
<td>(40,15,45)</td>
<td>Variable</td>
</tr>
<tr>
<td>( \overline{M}_1 ) (between-type with lag-1)</td>
<td>(41,15,44)</td>
<td>Variable</td>
</tr>
<tr>
<td>Areal precipitation (mean daily precipitation for the five major basins: Danube, Elbe, Ems, Rhine, Weser)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>99.5th-percentile</td>
<td>(51,34,15)</td>
<td>Good</td>
</tr>
<tr>
<td>5-day maxima</td>
<td>(83,17,0)</td>
<td>Good</td>
</tr>
<tr>
<td>10-day maxima</td>
<td>(83,17,0)</td>
<td>Good</td>
</tr>
<tr>
<td>Daily average temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>(100,0,0)</td>
<td>Good</td>
</tr>
<tr>
<td>99.5th-percentile</td>
<td>(70,20,10)</td>
<td>Good</td>
</tr>
</tbody>
</table>
4.3.1 At-site nsRWG performance

The nsRWG excels in resembling the observed wet-day frequency (Figure 5, top row) with a perfect GFP-score (100,0,0) across the model domain. This exceptional performance is evident in both seasons, as shown by the red dots closely aligned with the 1:1 line. The narrow grey bars reflect small uncertainty. CP1, CP5 and CP6 display a higher number of wet days compared to CP2, CP3 and CP4, consistent with our inference from the circulation patterns in Section 4.1 (Figures 2 and 3).

nsRWG reproduces the four transition probabilities reasonably well (Table 3). The model is particularly good at capturing the wet-to-wet (p11) and dry-to-dry (p00) transitions. Wet-to-dry (p10) and dry-to-wet (p01) transitions are more challenging, but the overall performance is still categorized as good. The performance across the CPs and seasons is nearly uniform with some small variations (Figure 5). For instance, the transition probabilities from dry-to-wet are slightly overestimated for CP1, CP3 and CP4 in both seasons. CP5 generally shows larger variability in performance across model realizations. The nsRWG performance with regards to wet-day frequency and transition probabilities is comparable to the stationary model performance (Nguyen et al., 2021).
Figure 5. Comparison of observed and simulated wet-day frequency (top) and transition probabilities (p11: wet-to-wet; p01: dry-to-wet) of daily precipitation (middle and bottom) at all grid cells. Red dots represent the median of the grey range corresponding to 100 model realizations.
nsRWG accurately reproduces mean daily precipitation sums and the 99.5th percentile of daily precipitation (Table 3). The performance for extreme precipitation is fairly uniform across CPs and seasons (Figure 6), i.e., most of the red dots are close to the 1:1 line.

Figure 6. Comparison of observed and simulated extreme (99.5th-percentile) daily precipitation (top) and 98th-percentile seasonal precipitation sum (bottom) for each CP at all grid cells. Red dots represent the median of the grey range corresponding to 100 model realizations.

The nsRWG performance with regard to the seasonal precipitation sum for each CP is good (Table 3). The mean of the seasonal sum is perfectly matched (Table 3), but also the 98th percentile is very well reproduced for all CPs (Table, Figure 6). Dry CPs (CP2 and CP4) show quite strong variability in 98th-percentile of seasonal precipitation sums in comparison to wetter CPs. CP5 in summer also exhibits strong variability between different model realizations (Figure 6). Though this pattern is associated with high mean and extreme daily precipitation (Figure 3), the total seasonal precipitation sum is relatively small (Figure 6). For all CPs and
seasons, the median of model realizations are close to the 1:1 line, what should be expected for a good model performance.

Figure 7 shows the good performance of nsRWG in reproducing 5-day and 10-day precipitation maxima. This statistic reflects the model’s ability to correctly generate maximum multi-day precipitation which is particularly relevant for flooding. This metric integrates the model performance with respect to auto-correlation, transition probabilities and marginal probabilities. The overall performance (Table 3) is good and has remarkably improved compared to the stationary model application (Nguyen et al., 2021).

Figure 7. Comparison of observed and simulated multi-day precipitation sums accumulated over 5-day and 10-day periods at all grid cells. Red dots represent the median of the grey range corresponding to 100 model realizations.

nsRWG excels in reproducing the mean of daily average temperature in each month at all locations (Table 3). The performance with regards to the 99.5\textsuperscript{th}-percentile of daily average temperature is slightly weaker but still overall good (Table 3). This is obviously more challenging for nsRWG to match the extreme percentiles compared to the mean. The performance with regards to the 99.5-percentile is fairly stable across all months (Figure 8) and most of the red dots corresponding to the mean of 100 realizations align closely to the 1:1 line. The spread in winter half-year (Nov-Mar) is however slightly stronger than in summer half-year (May-Oct). The performance of the non-stationary model version is comparable to or event slightly better than of the stationary RWG (Nguyen et al., 2021).
Figure 8. Comparison of observed and simulated 99.5\textsuperscript{th}-percentile average daily temperature for all months and grid cells. Red dots represent the median of the grey range corresponding to 100 model realizations.

4.3.2 Spatial nsRWG performance

Figure 9 provides an overview of the nsRWG ability to replicate the spatial dependence structure, as characterized by three different types of correlation: $M_0$ (lag-0), $M_1$ (within-type with lag-1), and $\overline{M}_1$ (between-type with lag-1). The simulated $M_0$ correlation closely aligns with the observed correlation structure demonstrating a close match even for inter-site distances of up to 1100 kilometers. However, the other two correlation types $M_1$ and $\overline{M}_1$ are partly significantly underestimated across all CPs and seasons. This is particularly evident in the case of $\overline{M}_1$, i.e., the model has some difficulties in representing spatial rainfall during transitions between days characterized by two different CPs.
Figure 9. Comparison of observed and simulated spatial correlation versus inter-site distance: $M_0$ lag-0 correlation (top), $M_1$ “within type” correlation (middle) and $\bar{M}_1$ “between type” correlation (bottom). Increasing density of points for the observed series is indicated in shaded colors from yellow to red. The density of points for the simulated series is indicated by the contour lines.
Despite some deficiencies in reproducing the spatial correlation, nsRWG is capable of reproducing various characteristics of the catchment average precipitation (mean, 99.5th percentile, 5-day and 10-day maxima) for the five major river basins in Germany (Table 3). The performance in terms of areal extreme (99.5th-percentile) daily precipitation is fairly good and consistent across all basins and CPs (Figure 10). The model slightly underestimates the extreme areal precipitation, particularly in summer, which is the consequence of the underestimation of spatial correlations (Figure 9).

**Figure 10.** Comparison of observed and simulated extreme (99.5th-percentile) daily precipitation averaged over the five major river basins in Germany (Danube, Elbe, Ems, Rhine and Weser) and over the whole domain (all). The color dots represent the median of the grey range corresponding to 100 realizations.

### 4.4. Future projected changes

In the following section we demonstrate an application of nsRWG for the generation of long synthetic precipitation series conditioned on the selected CMIP6 GCMs. Here the aim is not to provide a detailed analysis of projected changes in dynamics and thermodynamics, but rather to demonstrate how these explain changes in statistically downscaled extreme precipitation over Germany. A more comprehensive analysis of projected changes in the atmosphere and the associated flood hazard and risk is the focus of future research.

#### 4.4.1 Changes in circulation pattern frequency and persistence

Projected changes in the frequency of CPs are more pronounced in summer than in winter across all GCMs (Figure 11). The nine selected GCMs are mostly consistent in projecting frequency changes of the six CPs in summer, but show a more mixed pattern in winter. This behavior is similar in both near and far future periods and for both SSPs. In summer, wetter CPs such as CP1, CP5 and CP6 become less frequent in nearly
all GCMs. In particular, the reduction in frequency of CP6 is pronounced and consistent in all but one GCM. On the contrary, dryer patterns (CP2 and CP4) and the average wet pattern CP3 become more frequent in summer. The driest CP4 experiences the strongest positive change. The UKESM1-0-LL model projects that the frequencies of this pattern almost doubles in the near future. In winter, changes are less consistent across models and periods. On average, a slight increase is projected for the wetter CP5 characterized by westerly flows and responsible for extreme precipitation particularly in western and southwestern Germany (Figure 3). The dryer winter pattern CP3 and CP4 show a slight reduction of occurrence frequency (factors between 0.78 and 0.94 on average), which is mostly consistent across GCMs. Here, UKESM1-0-LL stands out and shows an opposite tendency (Figure 11).

### Near future

<table>
<thead>
<tr>
<th>GCMs</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKESM1-0-LL</td>
<td>0.78</td>
<td>0.61</td>
</tr>
<tr>
<td>CanESM5</td>
<td>0.88</td>
<td>1.11</td>
</tr>
<tr>
<td>CESM2</td>
<td>1.14</td>
<td>1.11</td>
</tr>
<tr>
<td>CNRM-CM6-1</td>
<td>1.15</td>
<td>1.14</td>
</tr>
<tr>
<td>INM-CM5-0</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>MPI-ESM1-2-HR</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>MRI-ESM2-o</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>GFDL-ESM4</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>IPSL-CM6A-LR</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>Average</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>

### Far future

<table>
<thead>
<tr>
<th>GCMs</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKESM1-0-LL</td>
<td>0.6</td>
<td>0.77</td>
</tr>
<tr>
<td>CanESM5</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>CESM2</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>CNRM-CM6-1</td>
<td>1.18</td>
<td>1.27</td>
</tr>
<tr>
<td>INM-CM5-0</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>MPI-ESM1-2-HR</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>MRI-ESM2-o</td>
<td>1.28</td>
<td>1.28</td>
</tr>
<tr>
<td>GFDL-ESM4</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>IPSL-CM6A-LR</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Average</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>

### Circulation pattern

<table>
<thead>
<tr>
<th>GCMs</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKESM1-0-LL</td>
<td>0.6</td>
<td>0.77</td>
</tr>
<tr>
<td>CanESM5</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>CESM2</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>CNRM-CM6-1</td>
<td>1.18</td>
<td>1.27</td>
</tr>
<tr>
<td>INM-CM5-0</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>MPI-ESM1-2-HR</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>MRI-ESM2-o</td>
<td>1.28</td>
<td>1.28</td>
</tr>
<tr>
<td>GFDL-ESM4</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>IPSL-CM6A-LR</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Average</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>
Figure 11. Changes in projected circulation pattern frequency for the near future (top) and the far future (bottom). The changes in frequency are represented by the ratio between CP frequency in the future and the control periods. ‘Average’ represents the average frequency changes across the nine selected GCMs.

4.4.2 Changes in regional temperature

Changes in mean regional temperature corrected for bias between the historical and control periods are analyzed across SSPs, GCMs, CPs and seasons (Figure 12). Since the marginal precipitation distributions are conditioned on the regional temperature for each CP, this is a valuable insight into how temperature changes for different CPs. The vast majority of GCMs indicate a positive regional temperature change for future periods. Only a few GCMs (e.g., INM-CM5-0 and GFDL-ESM4) show negative changes for some CPs. The far future and the more pessimistic SSP585 scenario show stronger positive changes as expected. The positive signals are stronger in summer than in winter. The temperature change is weakest for the wettest CPS in all models and scenarios. One of the driest patterns CP4 shows the strongest temperature increase in both summer and winter. Circulation patterns CP1 and CP3, which are medium wet patterns concerning both average and extreme precipitation (Figure 3), show a relatively strong positive temperature change. While the CP1 frequency is consistently decreasing (Figure 11), the projected frequency of CP3 in combination with positive temperature will increase the importance of CP3 for total precipitation input in the future. The average summer and winter temperature changes for the far future and SSP585 scenario are in the range of changes assessed by Coppola et al. (2021a) for a somewhat larger CMIP6 GCM ensemble in the central European region.
Figure 12. Changes in average regional temperature between historical and the near future (top) and the far future (bottom), respectively, across nine GCM projections, two SSPs, six CPs and two seasons. ‘Average’ represents the average change across the nine selected GCMs.

4.4.2. Changes in future extreme precipitation generated with nsRWG

Here, we present the changes in extreme precipitation conditioned on future climate projections. We examine the plausibility of seasonal and spatial patterns of changes in extreme precipitation in relation to the available literature. This discussion is impaired by the still small number of CMIP6 based analyses of extreme precipitation over the Central European domain. An in-depth analysis of downscaled precipitation data and a comprehensive flood impact assessment will be addressed in a separate, forthcoming study.
Figure 13. Changes in extreme (99.5\textsuperscript{th}-percentile) daily precipitation between historical and the near future (left) and historical and far future (right) periods for winter (DJF), spring (MAM), summer (JJA) and autumn (SON) and two SSP scenarios. The results are averaged over nine GCM models and 100 nsRWG realizations.

Extreme daily precipitation from nsRWG is projected to consistently increase over the target region, except over northern Italy in the summer months (JJA) (Figure 13). The overall increase in extreme precipitation is in line with the assessment of the CMIP6 GCM ensembles for western and central Europe for the 99\textsuperscript{th}-percentile daily precipitation (Coppola et al., 2021\textsuperscript{a,b}) and for the seasonal 20-year return period precipitation (Ritzhaupt & Maraun, 2023), although the model spread is considerable, particularly in summer. Decreases in summer over Italy for daily mean and hourly extreme precipitation are consistent with the MPI-ESL-LL downscaled by the WRF regional climate model driven by the CMIP5 RCP4.5 scenario (Knist et al., 2020). According to nsRWG, the precipitation increase for SSP245 is mostly in the range of up to 20% and up to 40% in the near and far future, respectively, for all seasons (Figure 13). The increase is stronger for SSP585 compared to SSP245 and is particularly pronounced in the summer months (JJA) in the far future. However, the autumn months (SON) also show a notable increase. Given the decreasing frequency of the wetter circulation patterns in the summer half-year, the increase in extreme precipitation is likely to be dominated by thermodynamic changes. However, the question of attributing precipitation and flood changes to changes in the dynamic and thermodynamic components is subject of future research.

5. Conclusions

In this paper, we develop a non-stationary version of the auto-regressive Regional Weather Generator (nsRWG) conditioned on circulation patterns (CPs) and regional daily mean temperature. nsRWG is designed to generate a synthetic long-term (thousands of years) daily weather series for use with hydrological impact models to assess future flood risks. By conditioning nsRWG on circulation patterns and regional...
temperature, we consider the effect of changes in the dynamic and thermodynamic properties of the atmosphere on changes in local precipitation. Non-stationary extended Generalized Pareto distribution is used to simulate the marginal precipitation distributions. CPs are used as a latent variable to parameterize the marginal non-stationary precipitation distributions, whose scale parameter is conditioned on the regional daily mean temperature. Power-transformed temperature data are modelled using a non-stationary normal distribution conditioned on the regional daily mean temperature.

The nsRWG is set up for the domain of more than 650,000 km² covering five major river basins in Germany – Danube, Elbe, Rhine, Weser and Ems – using E-OBS gridded observation data of daily precipitation and temperature. Circulation patterns and the regional temperature are derived based on the ERA5 reanalysis. The evaluation of nsRWG following the CASE framework by Bennet et al. (2018) shows overall good results with regards to the at-site precipitation intermittency properties, mean and extreme (99.5th-percentile) daily and multi-day precipitation sums. The comparison with a stationary precipitation model version, which does not include temperature as a covariate in the marginal precipitation distributions, shows a superior performance to the non-stationary model of more than 70% of grid cells in the study area. Matching spatial correlation structure of precipitation remains a challenge for nsRWG, in particular when transitioning between days with different circulation patterns. Nevertheless, the areal extreme precipitation for the major German river basin is very well reproduced by the model. nsRWG excels in simulating a minimum, mean and maximum daily temperature and their extreme percentiles.

The link between large-scale atmospheric characteristics such as CPs and regional temperature on one side and local precipitation and temperature on the other allows us to use pressure and temperature variables from the general circulation models (GCMs) to generate local weather and at the same time, to account for the climate change signal manifested in changes of frequency and persistence of circulation patterns and regional warming. This approach is charming as it relies on the mean sea level pressure for CP classification and on the regional temperature – two variables that are simulated by the global climate models more skillfully than precipitation. We demonstrate the application of nsRWG for downscaling the precipitation from nine CMIP6 GCMs weighted by the ClimWIP approach. The latter is used to assess the skill of GCMs to reproduce the mean, variability and trend of the covariates (mean sea level pressure and regional temperature) in the historical period (1985-2014). The results suggest a consistent increase in extreme precipitation over the German basins in the near (2031-2060) and far (2071-2100) future, in line with the previous regional analyses of various CMIP6 ensembles. By generating a long synthetic series for each of the historical and future periods, we can estimate precipitation change more robustly than if we were only using a 30-year series directly available from a climate model. Hence, nsRWG offers a key benefit for hydrological
impact studies by providing long-term (thousands of years) consistent synthetic weather data indispensable for the robust estimation of high flood flow quantiles and future flood risk changes.

Acknowledgments

This research has been funded by the Federal Ministry of Education and Research of Germany in the framework of the project FLOOD (project numbers 01LP1903E and 01LP2324E) as a part of the ClimXtreme Research Network on Climate Change and Extreme Events within framework program Research for Sustainable Development (FONA3). Funding by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for the research group FOR 2416 "Space-Time Dynamics of Extreme Floods (SPATE)" (project number 278017089) is gratefully acknowledged. The results are generated using Copernicus Climate Change Service information 2023. Lukas Brunner has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2037 “CLICCS – Climate, Climatic Change, and Society” – project no. 390683824, a contribution to the Center for Earth System Research and Sustainability (CEN) of the Hamburg University.

Author contributions

SV and BM developed the concept and acquired funding. DVN coded, parameterized and validated the weather generator and carried out the analyses. SV and DVN processed climate data. KN developed circulation pattern classification. LB performed ClimWIP analysis. SV and DVN wrote the manuscript with contributions from all authors.

Data and code availability

ERA5 and CMIP6 GCM output data were accessed through the XCES at the Deutsches Klimarechenzentrum (DKRZ). nsRWG is available at GFZ GitLab repository. The access can be granted upon request. Daily temperature and precipitation data generated by nsRWG for the ensembles of 72x100 years, nine GCMs, two future periods and two SSP scenarios will be made available at GFZ data repository (https://dataservices.gfz-potsdam.de/portal/index.html) upon the acceptance of the manuscript.

References


