



Long-term variance of heavy precipitation across central Europe using a large ensemble of regional climate model simulations

Florian Ehmele, Lisa-Ann Kautz, Hendrik Feldmann, and Joaquim G. Pinto

Institute of Meteorology and Climate Research, Department Troposphere Research (IMK-TRO), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

Correspondence: Florian Ehmele (florian.ehmele@kit.edu)

Received: 26 August 2019 – Discussion started: 2 September 2019

Revised: 9 March 2020 – Accepted: 26 April 2020 – Published: 26 May 2020

Abstract. Widespread flooding events are among the major natural hazards in central Europe. Such events are usually related to intensive, long-lasting precipitation over larger areas. Despite some prominent floods during the last three decades (e.g., 1997, 1999, 2002, and 2013), extreme floods are rare and associated with estimated long return periods of more than 100 years. To assess the associated risks of such extreme events, reliable statistics of precipitation and discharge are required. Comprehensive observations, however, are mainly available for the last 50–60 years or less. This shortcoming can be reduced using stochastic data sets. One possibility towards this aim is to consider climate model data or extended reanalyses. This study presents and discusses a validation of different century-long data sets, decadal hindcasts, and also predictions for the upcoming decade combined to a new large ensemble. Global reanalyses for the 20th century with a horizontal resolution of more than 100 km have been dynamically downscaled with a regional climate model (Consortium for Small-scale Modeling – CLimate Mode; COSMO-CLM) towards a higher resolution of 25 km. The new data sets are first filtered using a dry-day adjustment. Evaluation focuses on intensive widespread precipitation events and related temporal variabilities and trends. The presented ensemble data are within the range of observations for both statistical distributions and time series. The temporal evolution during the past 60 years is captured. The results reveal some long-term variability with phases of increased and decreased precipitation rates. The overall trend varies between the investigation areas but is mostly significant. The predictions for the upcoming decade show ongoing tendencies with increased areal precipitation. The presented regional climate model (RCM) ensemble not only allows for more robust statistics in general, it is also suitable for a better estimation of extreme values.

1 Introduction

Ongoing climate change affects not only the global scale but also impacts the regional climate. Regarding air temperature, there is a more or less clear trend in the recent past, which reveals a clear anthropogenic signal. However, various climate simulations show distinct differences for precipitation trends, especially for heavy precipitation (e.g., Moberg et al., 2006; Zolina et al., 2008; Toreti et al., 2010). A review of observed variability and trends in extreme climate events states that it is difficult to find significant relations between the greenhouse-gas-enhanced climate change and increases or

decreases in extreme precipitation events (Field et al., 2012). This is attributed to their rare occurrence, the general high spatial variability of precipitation, and a lack of long-term high-quality observations.

Magnitude and sign of heavy precipitation trends strongly depend on various factors such as the regarded area or the considered time period (e.g., Easterling et al., 2000). Global tendencies towards more intense precipitation throughout the 20th century between summer and winter seasons also account for precipitation trends. For example, Moberg and Jones (2005) found an increase in winter precipitation across central and western Europe between 1901 and 1999, while

Pal et al. (2004) found a decrease in summer precipitation for the period 1951–2000. Dittus et al. (2016) found an increasing trend between 1951 and 2005 in extreme total precipitation amounts for Europe in global climate model (GCM) simulations (Coupled Model Intercomparison Project phase 5; CMIP5). Similar trends were found in global reanalyses (e.g., ERA-20C; Poli et al., 2016) but not in observations. In contrast, Primo et al. (2019) found positive trends for two ground-based observational stations in Germany using extreme precipitation indices.

Model resolution is another crucial factor. The use of high-resolution regional climate models (RCMs) instead of global data sets revealed a more detailed and orographically related spatial structure of the precipitation fields and trends (e.g., Feldmann et al., 2013). An increase of both areal mean precipitation and extremes in central Europe on the order of 5%–10% was found in RCM simulations by Feldmann et al. (2013), which will continue with almost same magnitude for the next decade. Differences in precipitation trends also stem from varying definitions of extreme events such as certain thresholds, percentile-based indices, or return periods (e.g., Maraun et al., 2010). While most of these studies show trends in daily precipitation, just a few deal with subdaily trends. Barbero et al. (2017), for instance, compared trends in subdaily and daily extremes. Although significant increasing trends were found for both time ranges, trends in daily extremes are better detected than in subdaily extremes.

Spatially extended intensive rainfall events are frequently related to widespread flooding along the main river networks of central Europe causing major damage on the order of several billion euros (EUR) per event (e.g., Uhlemann et al., 2010; Kienzler et al., 2015; Schröter et al., 2015; MunichRe, 2017). A prominent example of such an extreme and devastating event is the flood in 2012 along the Elbe and Danube rivers (Ulbrich et al., 2003a, b). Such outstanding events are by definition extremely rare, which makes the risk estimation difficult or almost impossible due to the limited time period with available area-wide observations (e.g., Pauling and Paeth, 2007; Hirabayashi et al., 2013). However, trend analyses of such extreme events and the related risks during the past and for the future are of great importance for insurance purposes or flood protection (e.g., Merz et al., 2014; Schröter et al., 2015; Ehmele and Kunz, 2019).

A possible way of dealing with the unsatisfactory data availability is through century-long simulations using climate models (e.g., Stucki et al., 2016) or stochastic approaches (e.g., Peleg et al., 2017; Singer et al., 2018; Ehmele and Kunz, 2019). The currently used GCMs were found to be in good agreement with the available but limited observations (Fischer and Knutti, 2016). Brönnimann et al. (2013) and Brönnimann (2017) analyzed historical extreme events using century-long reanalysis data sets and concluded that the quality of the reanalyses strongly depends on the number and type of the assimilated observations. The investigated historical events were reproduced, but the magnitudes were

underestimated. A possible reason is the decreasing number and quality of observations early in the century and therefore a lack of assimilation data. The suitability of reanalysis data to investigate extreme precipitation for England and Wales was investigated by Rhodes et al. (2015). While time series of daily precipitation totals are well represented in both data sets, timing errors of heavy precipitation events were identified as one of the major problems. Stucki et al. (2012) investigated historical flooding events in Switzerland and indicate that the reanalyses underestimate precipitation in Switzerland which may result from the insufficient representation of the alpine topography. The timing and the exact location of heavy precipitation were also found to be inaccurate.

As shown by van der Wiel et al. (2019) or Martel et al. (2020), large ensembles can have an added value for flood risk estimation and for the calculation of return periods of heavy precipitation. van der Wiel et al. (2019) found a clear benefit in using an ensemble approach for the estimation of changes in hydrological extremes including compound events compared to traditional approaches. Martel et al. (2020) found similar results, namely a reduction in the projected return period of 100-year annual maximum precipitation with the different ensembles, albeit having different model structures and resolutions. Furthermore, it was emphasized that a higher resolution is advantageous to predict climate change signals over complex terrain. Other studies also highlighted the improvements of using high-resolution RCMs for the investigation of climate extremes (e.g., Feser et al., 2011; Feldmann et al., 2008, 2013; Schewe et al., 2019), especially over complex terrain (e.g., Torma et al., 2015).

The studies mentioned above document partly contrasting results and demonstrate the challenges arising when dealing with extreme precipitation and related phenomena. In this study, a set of different realizations with one RCM is used and combined to the new LAERTES-EU (LArge Ensemble of Regional climaTe modeL Simulations for EUrope), which can be used for more profound statistical analyses. The basis is the global 20th Century Reanalysis (20CR) data set (Compo et al., 2011), which was dynamically downscaled for Europe. LAERTES-EU consists of a handful of 20th century reanalysis data sets and a large ensemble of decadal hindcast simulations mainly for the second half of the century. Although all simulations were performed with the same RCM version and setup, LAERTES-EU is a combination of different external forcings, boundary conditions, and/or assimilation. Predictions for the upcoming decade will round up our analysis. The investigative focus lies on daily values of intensive areal precipitation which can be associated with major flood events in central Europe. As demonstrated, for example, by Schröter et al. (2015), severe flood events along the major river networks in central Europe are related to long-lasting and widespread precipitation events of mainly stratiform origin with embedded convective precipitation. Typically, intensities do not reach the most extreme rates of the

distribution but are characterized by high spatial mean values.

LAERTES-EU is validated in terms of coincidence with observations regarding temporal variability, statistical distributions, and possible long-term trends. The following research questions will be addressed.

1. How well is extreme areal precipitation represented in the RCM ensemble LAERTES-EU?
2. What are potential benefits of LAERTES-EU compared to other available data sets?
3. Which temporal evolution and variability of extreme areal precipitation over central Europe have manifested during the past?
4. Which tendency is expected for the upcoming decade?

A better interpretation of RCM data and a more profound understanding of extreme areal precipitation may have several applications such as risk assessments. Although they are relevant, we do not handle the potential mechanisms behind temporal variance and trends as well as spatial and seasonal differences as this goes beyond the scope of this study.

This paper is structured as follows. The data sets which were used in this study are introduced in Sect. 2. Section 3 sums up the methods used for the analysis and the validation. In Sect. 4, LAERTES-EU is validated with observations for a reference period. The investigation of temporal variabilities and trends is given in Sect. 5. Finally, Sect. 6 gives a summary and lists our conclusions.

2 Data sets

Two different types of data sets are applied in this study: gridded precipitation data based on observations and partly century-long climate model simulations (LAERTES-EU). The observational data sets are primarily available for the second half of the 20th century and serve as reference data for the validation of the ensemble. Furthermore, we compare LAERTES-EU with the forcing global model and also with the global 20CR data set (Compo et al., 2011), which were used as initial data for some of the simulations.

2.1 Observations

The European observational data set (E-OBS) version 17 including daily precipitation (Haylock et al., 2008; van den Besselaar et al., 2011) is a gridded data set with a horizontal resolution of 0.22° (≈ 25 km) covering the years 1950 to 2017. This version shows some improvements towards older versions, since updated algorithms and new stations have been included in some areas (e.g., for Poland). The E-OBS algorithm interpolates observations from weather stations to a regular grid using geostatistical methods (e.g., Journel and Huijbregts, 1978; Goovaerts, 2000). Note that E-OBS

is a land-only data set and ocean grid points are set to a missing value. Haylock et al. (2008) stated that rainfall totals in E-OBS are reduced by up to almost one-third compared to the raw station data at the corresponding grid cells. Regarding extremes, the deviation of E-OBS is even more pronounced (Hofstra et al., 2009). Nevertheless, both studies stated that the spatial mean precipitation in E-OBS is very close to other observations.

Although E-OBS has some limitations, we use it as the main reference for this study, as there is no other comparable high-resolution daily precipitation data set available that covers entire Europe for a long time period. Other products like satellite data with a very limited time frame are not helpful and also have limitations. There are single ground-based observations with very long time series, but as the focus of this study is on intensive areal precipitation, these data are of limited usefulness for validation.

In addition to E-OBS, we compare the RCM simulations with the central European high-resolution gridded daily data sets (Hydrological RASter data sets; HYRAS) provided by the German Weather Service (DWD; Rauthe et al., 2013). HYRAS is a gridded precipitation data set with a horizontal resolution of up to 1 km for the time period 1951–2006 and covers Germany and the surrounding river catchments. The HYRAS algorithm also uses ground-based measurements and interpolates the point observations to the regular grid. For this study, the HYRAS data were first aggregated to the E-OBS/RCM 25 km grid. HYRAS hereafter refers to this aggregated 25 km data set.

2.2 Regional climate model simulations

LAERTES-EU combines a large number of regional dynamical downscaling simulations for Europe performed with a single RCM. The used RCM is the non-hydrostatic model of the Consortium for Small-scale Modeling (COSMO) CLimate Mode (CLM) version 5 (CCLM5; Rockel et al., 2008), which has a spatial resolution of 0.22° (≈ 25 km). The model covers the European domain of the Coordinated Downscaling Experiment (EURO-CORDEX)¹ (Jacob et al., 2014). Overall, the simulations use the same domain, model version, and setup, which was adapted from EURO-CORDEX (Kotlarski et al., 2014). According to Feldmann et al. (2008), a dry-day correction is important as climate models tend to overestimate the number of wet days with low intensities below 0.1 mm, known as the drizzle effect (Berg et al., 2012). In order to reduce this typical bias, a dry-day adjustment was first applied to LAERTES-EU. The E-OBS data were used for this correction, as they have the same spatial extension and resolution as the CCLM simulations. All simulations are performed within the BMBF (Federal Ministry of Education and Research of Germany) MiKlip II² project (Marotzke

¹<http://www.euro-cordex.net> (last access: 20 May 2020).

²<https://www.fona-miklip.de/> (last access: 20 May 2020).

et al., 2016) to create and test a decadal prediction system including a regional downscaling component for Europe.

For all downscaling simulations, the boundary conditions were derived from the Max Planck Institute of Meteorology coupled Earth System Model (MPI-ESM). This global model consists of an atmospheric component (ECHAM6) (Stevens et al., 2013), an ocean component (MPI-OM) (Jungclauss et al., 2013), and a land-surface model (JSBACH) (Hagemann et al., 2013).

LAERTES-EU is divided into four different data blocks (Table 1) depending on the setup of the forcing MPI-ESM ensemble simulations. The differences between the four data blocks stem from the setup, external forcing, and initialization of the MPI-ESM simulations. Data blocks 1 and 2 of the RCM ensemble (compare Table 1) obtained the boundary values from the MPI-ESM-LR simulations using a T63 resolution and 47 vertical layers. Data blocks 3 and 4 used the MPI-ESM-HR version (Müller et al., 2018) as their driving model. In this version, the horizontal resolution is T127 and 95 vertical layers are applied. Three types of forcing ensembles can be distinguished:

- i. MPI-ESM assimilates reanalysis data for long-term simulations (data block 1);
- ii. long-term historical-type simulations, according to the CMIP5 specifications (data block 3; Taylor et al., 2012); and
- iii. initialized decadal (10-year) hind- and forecast simulations (data blocks 2 and 4).

In data block 1, the first type (I) is applied. Here, the 20CR data (Compo et al., 2011) are assimilated into the MPI-ESM-LR (Müller et al., 2014). The 20CR data set has a spatial resolution of approximately 2° (T62) and was generated using the Global Forecast System (GFS; Kanamitsu et al., 1991; Moorthi et al., 2001) of the National Centers for Environmental Prediction (NCEP)³. It used a 56-member ensemble Kalman filter approach to assimilate surface pressure, monthly sea surface temperature, and sea-ice observations. Three of the 20CR members are assimilated into MPI-ESM to provide long-term (110 years each) climate reconstruction simulations over the period 1900–2009 (Müller et al., 2014). Afterwards, a downscaling with CCLM uses these global simulations as boundary conditions (e.g., Primo et al., 2019).

Data block 3 consists of the second type (II), where five so-called historical simulations of MPI-ESM-HR with CMIP5 observed natural and anthropogenic external climate forcing (Taylor et al., 2012) are used as boundary conditions for CCLM. The ensemble was generated by starting the MPI-ESM from arbitrary dates in a pre-industrial control simulation (Müller et al., 2014). Three of the five CCLM members cover the period 1900–2005 (106 years each). The two ad-

ditional simulations cover the period 1960–2005 (46 years each).

Data blocks 2 and 4 consist of initialized decadal simulations (type III). The starting conditions are derived from an observed state (Müller et al., 2012; Marotzke et al., 2016). For each starting year, an ensemble of decadal simulations is generated and then, the initialization point is shifted by 1 year (e.g., 1961–1970, 1962–1971, and so on). Due to the overlap, a specific calendar year may be covered by several decadal hindcasts with different starting years. These decadal hind- and forecasts thus represent the current state of the major modes of climate variability compared to the so-called uninitialized historical simulations (data block 3). The downscaling procedure, the skill, and the added value are described in Mieruch et al. (2014), Feldmann et al. (2019), and Reyers et al. (2019).

In data block 2, the starting conditions of the three decadal hindcast members with MPI-ESM-LR are derived from the assimilation experiments in data block 1. The starting years of the CCLM downscaling range from 1910 to 2009. This means the last simulated year is 2019.

Data block 4 consists of two parts. Both of them use the MPI-ESM-HR version. The so-called preop-ensemble has five members. The external climate forcing is derived from CMIP5. The starting years range from 1960 to 2016 (last simulated year is 2026). The so-called dcppA-hindcast ensemble has 10 members and uses the external forcing for CMIP6 (Eyring et al., 2016). The global simulations are a contribution to the Decadal Climate Prediction Project of CMIP6 (DCPP; Boer et al., 2016). The starting years are 1960 to 2018 (last simulated year is 2028).

In total, LAERTES-EU consists of 1183 simulation runs (sample size) with approximately 12 500 simulated years. The number of ensemble members for a specific year varies from six at the beginning of the century to a maximum of 188 members between 1970 and 2000 (see Fig. S1 in the Supplement). The simulations in all four data blocks are affected by the observed external climate forcing, but they differ with respect to the representation of the observed climate variability; whereas data block 1 uses assimilated 20CR data, data blocks 2 and 4 contain initialized hindcasts, which to some degree follow the observed low-frequency variability, and data block 3 only uses the external forcing information. Nonetheless, the four groups of downscaling simulations can be grouped into a large ensemble, since the regional simulations were all performed with the same setup of the RCM. Despite the same initial conditions and model setup, the temporal evolution of the day-to-day weather is (statistically) independent between the members after a few weeks. This is an advantage, since the data set is homogeneous over time but also covers uncertainties in the observations including unknown and not-yet-observed events. The validity of this combination approach is tested within Sect. 4.

³<http://www.ncep.noaa.gov/> (last access: 20 May 2020).

Table 1. Overview of the RCM ensemble LAERTES-EU with the name of the simulation within the MiKlip project, the classification into data blocks, the underlying setup (experiment), the covered time period, and the number of simulation years; XX stands for the ensemble number and YYYY stands for the initialization year.

Name	Block	Experiment	Period	Years	Comment
as20ncepXX	1	20CR via MPI-ESM-LR	1900–2009	330	3 members of 110 years each
decXXoYYYY	2	MPI-ESM-LR DROUGHTCLIP	1911–2019	3000	3 members with 100 decades each
historical_rXi1p1-HR	3	MPI-ESM-HR HISTORICAL	1900–2005	410	run 1–3 each with 106 years, run 4–5 each with 46 years (1960–2005)
preop	4	MPI-ESM-HR CMIP5	1961–2026	2850	5 members with 57 decades each
dcppA-hindcast	4	MPI-ESM-HR CMIP6	1961–2028	5900	10 members with 59 decades each

3 Methods

The ability of LAERTES-EU to simulate realistic precipitation amounts and distribution is an important requirement. Moreover, temporal variability and possible trends should also be well represented for trustworthy data sets. The methods were applied to different investigation areas and time periods. Equations and additional information can be found in Appendix A–C. As the focus of this study is intensive areal precipitation, we concentrate on high percentiles of spatially aggregated daily rainfall totals, namely 99 %, and 99.9 %. The percentiles are based on wet days only. First, a spatial aggregation of daily precipitation values was applied. Afterwards, the percentiles of these areal precipitation were calculated for each year separately. In all data sets, ocean grid cells were set to a missing value and therefore neglected.

3.1 Validation methods

LAERTES-EU is analyzed and validated using various methods. The intensity spectrum gives the statistical probability of each precipitation amount by taking into account all grid points and all time steps within the investigation area and without any aggregation. Therefore, the range of occurred values is divided into evenly spaced histogram classes, which then are normalized with the total sample size. The resulting intensity–probability curve (IPC) is a good indicator of whether the model is capable to simulate realistic precipitation intensity distributions.

As an extension to the IPCs, the linear error in probability space L (cf. Eqs. A1–A3) is analyzed (e.g., Ward and Folland, 1991; Potts et al., 1996). Therefore, empirical cumulative density functions (ECDFs) are calculated for each simulation run and for the observations. The data basis is the same as for the IPCs. The value ΔC_r (Eq. A1) is defined as the difference between the ECDF of a model run r and that of the observation (difference of probabilities) up to a specific precipitation intensity. It is therefore a measure for the over- or underestimation of the model. Using ΔC_r , the linear error in probability space (L_r ; Eq. A2) is the mean of the absolute

values $|\Delta C_r|$ over the entire precipitation range as defined by Déqué (2012) or Wahl et al. (2017). The better both density functions coincide, the lower the value of L_r . According to Eq. (A2), L_r is always positive. The ensemble mean is given by \bar{L} (Eq. A3).

The internal variability of LAERTES-EU on different time intervals is compared to that of the observations. Given that the focus of this study is on intensive widespread precipitation, this analysis is performed using spatial mean precipitation amounts averaged over the investigation areas. First, the time series of daily spatial means are aggregated over different intervals, namely monthly, seasonal, and yearly precipitation sums, as well as 5-, 10-, or 30-year running means. In a second step, the standard deviation of a gamma distribution σ_r is calculated for each of these interval series (see Appendix A; Eq. A4), for every single member of LAERTES-EU, and for the observations. Finally, the ensemble mean of the four data blocks and of the complete ensemble is built. This method enables the analysis of how well the internal variability on different timescales is captured by LAERTES-EU.

The quantile–quantile (Q–Q) plot compares the simulated distribution with the observed one using different percentiles of daily spatial mean precipitation. The Q–Q distributions are used to calculate the coefficient of determination R^2 with R being the Pearson correlation coefficient (Eq. A5).

The added value of the ensemble size is analyzed by using the signal-to-noise ratio (S2N) (Eq. A6). Therefore, we determine a Gumbel distribution (cf. Appendix A) for different sample sizes and the corresponding 90 % confidence interval. The S2N is then the ratio of the return value of the Gumbel distribution divided by the 90 % confidence interval (Früh et al., 2010).

3.2 Decadal variability and trend analysis

For the analysis of the temporal evolution of heavy precipitation, we use time series of different percentiles of spatial mean precipitation and quantities introduced and recommended by the Expert Team on Climate Change Detection

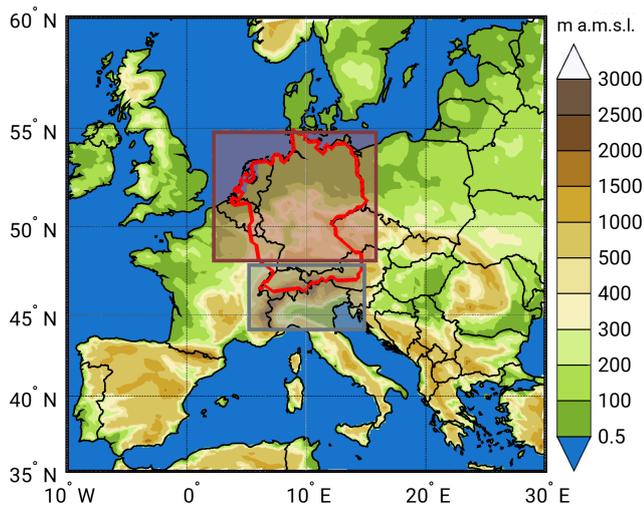


Figure 1. Topographic map of Europe at model resolution 0.22° (in meters above mean sea level; m a.m.s.l.) with the PRUDENCE regions Mid-Europe (ME; dark red box) and the Alps (AL; gray box), state borders (black contours), and the HYRAS area (light red contour). Ocean grid cells are set to a missing value.

and Indices (ETCCDI; Karl et al., 1999; Peterson, 2005). Currently, 27 indices for temperature and precipitation are defined by the ETCCDI. These indices can be used from local to global scales. Additionally, they combine extremes with a mean climatological state (Zwiers et al., 2013). In this study, we use the two indices (R95pTOT and R99pTOT; Eqs. B1 and B2), which indicate the amount of precipitation above the 95th or 99th percentile, respectively.

In terms of trend analysis, a Mann–Kendall test (Mann, 1945; Kendall, 1955) is performed with related significance investigations (Appendix C). Regarding possible oscillations, the complete time series is split into subseries with a minimum length of 10 years and up to 130 years (trend matrix). The Mann–Kendall test is applied to each of these subseries.

3.3 Investigation areas and time periods

The focus of this study is central Europe, implying the countries Germany, Switzerland, the Netherlands, Belgium, Luxembourg, and parts of France, Poland, Austria, the Czech Republic, and Italy. Following Christensen and Christensen (2007), these countries are mostly coincident with two of the areas defined in the PRUDENCE (prediction of regional scenarios and uncertainties for defining European climate change risks and effects) project, namely the PRUDENCE regions (PRs) Mid-Europe (ME) and the Alps (AL; Fig. 1). Although these boxes contain both land and ocean, the latter was set to a missing value and neglected. During validation, ME and AL were reduced to the HYRAS grid cells lying within the corresponding box, hereafter referred to as ME* and AL*.

The data sets are investigated on different time periods (TPs): TP1 covers the past from 1900 to 2017, which is divided into a subperiod (TP1b) only containing the period 1951 to 2006, with both observations (E-OBS and HYRAS) being available. TP2 is used for the predictions from 2018 to 2028. Note that the simulations were performed within the MiKlip project back in 2018 (using observations until 2017), which is the reason why the prediction period starts in 2018.

For climatological aspects, we use the time period of 1961–1990, hereafter referred to as climTP. A couple of studies (e.g., Cahill et al., 2015; Folland et al., 2018) showed that the climate change signal for global mean temperature significantly increased since the early 1980s. Therefore, using the time period of 1981–2010 as reference would possibly include a strong changing signal to the analysis. Using 1961–1990 reduces the influence of these effects, as this period shows more stable conditions to a certain degree. This also permits more room for the interpretation of the future predictions.

4 Validation of the RCM ensemble

In the following, the above-described methods are applied in order to validate LAERTES-EU concerning its representativeness with observations. With this aim, data for the investigation period TP1b are used, and the boxes ME and AL (cf. Fig. 1) are limited to the HYRAS area (ME* and AL*).

4.1 Statistical distributions and frequencies

The IPCs give the range of simulated (observed) precipitation intensities at any grid point within the investigation area and its corresponding probability (Fig. 2). For both investigation areas, the IPCs reveal a distinct added value of the RCM compared to the global model. Due to the coarse resolution, intensities greater than approximately 100 mm d^{-1} are not found in the GCMs, which underestimate by a large degree the probability of the high intensities. The same applies for the global reanalysis 20CR. On the other hand, the RCM tends to overestimate the probability for precipitation intensities above a threshold of approximately 50 mm d^{-1} but covers the entire range of values as the observations. The wider range of intensities at the upper tail of the distribution may include possibly not-yet-observed events.

For ME*, the IPCs of the RCM are close to HYRAS, but there is a systematic difference between HYRAS and E-OBS (Fig. 2a). As already mentioned by Haylock et al. (2008), E-OBS has a certain negative bias up to -30% when using grid-point-based quantities. The given deviation of HYRAS and E-OBS is within this range. Similar results can be found for AL* (Fig. 2b). The differences between the RCM simulations and the observations at a given probability are slightly less than those for ME*. For both areas, the range of simulated values is much higher (up to 400 mm d^{-1}). Naturally,

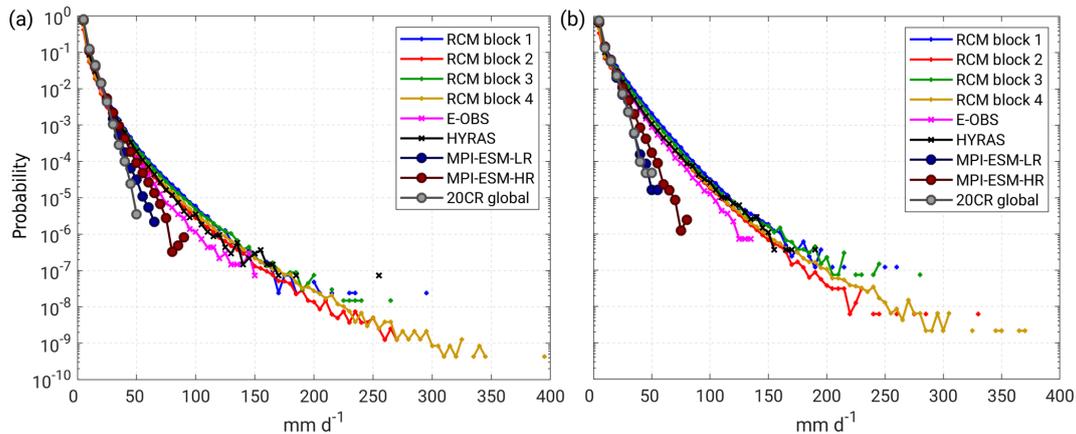


Figure 2. Intensity–probability curves (IPCs) of daily rainfall totals of the RCM simulations (dry-day adjusted), observations (E-OBS and HYRAS), GCM simulations (forcing MPI-ESM data at two resolutions – LR and HR), and global reanalysis data (20CR) for (a) Mid-Europe (ME*) and (b) the Alps (AL*), both limited to the HYRAS area during the investigation period TP1b (1951–2006). For the IPCs, every grid cell value at every time step was taken into account without any aggregation.

higher intensities are more likely in the mountainous AL* region.

In contrast to the grid-point-based IPCs, Fig. 3 shows the mean standard deviation of a gamma distribution (cf. Sect. 3.1 and Appendix A) for the time series of spatial mean precipitation amounts aggregated over different time intervals. For both areas, there is an expectable continuous decrease of internal variability towards longer periods for all data sets/data blocks. For ME*, LAERTES-EU is in good agreement with both observations at least up to a yearly perspective. For longer time periods, data block 1 shows a slightly different behavior compared to the other data blocks and observations. Nevertheless, data blocks 2–4 and the ensemble mean continue to match with the observations up to the 10-year running mean. Note that it is not possible to estimate the 30-year running mean for the decadal simulations of data blocks 2 and 4 given the data availability. For data block 3, only an external climate forcing was used, meaning these so-called historical are free runs in terms of daily weather evolution. Therefore, it is not expected that the multi-decadal variability is in phase to the observed circulation after a certain time, which can be a reason for slightly higher differences of data block 3 compared to the observations at the longest timescale. Furthermore, note that the results of Fig. 3 do not indicate a perfect match of LAERTES-EU in terms of absolute values, but rather that the internal variability (spread) of spatial mean precipitation totals is well captured. For the mountainous AL* region, the internal variability is higher and all data blocks have a higher standard deviation at all time intervals. This means that the spread of simulated precipitation amounts is increased compared to that of the observation. A possible reason for this difference can emerge from sparse measurements in that region considered for both E-OBS and HYRAS, especially for long-term observations. The more or less constant difference between

LAERTES-EU and the observations can be an indicator of a possibly systematic bias in this region.

The Q–Q plots of daily spatial mean precipitation fields for both investigation areas are shown in Fig. S2. Generally speaking, the distribution of the RCM is similar to those of the observations, at least to E-OBS, with little deviations from the optimum (diagonal line) for most of the spectrum and differences at around 10 % for the upper part of the distribution. In comparison to HYRAS, the maximum deviation is higher at around 20 %. For AL*, the differences between the RCM and HYRAS are larger than for ME* (Fig. S2). Even though HYRAS was aggregated to the E-OBS/RCM grid, the more pronounced differences especially for the extremes might be a result of the higher resolution of the HYRAS data, which, in particular, is of greater relevance in the mountainous region of AL*.

The findings of Fig. S2 are confirmed by the determination coefficients R^2 (Table 2). For both E-OBS and HYRAS, the coefficient is very high with $R^2 > 0.98$. There is a slightly higher R^2 for E-OBS than for HYRAS, which is an artificial effect of the data resolution. The region AL* shows a minimal higher skill compared to ME* in E-OBS and slightly lower values in HYRAS. Table 2 also reveals higher correlations of the CCLM simulations driven by the high-resolution MPI-ESM-HR data compared to those driven by the lower-resolved MPI-ESM-LR data. Even though this seems to be systematic, the differences are marginal.

Table 2 also contains the mean linear error in probability space \bar{L} for the different data blocks. Again, the differences between the data blocks are marginal with all cases being close to $\bar{L} = 0$, which indicates good agreement of LAERTES-EU with observations. In contrast to R^2 , \bar{L} has lower values for the simulations driven by MPI-ESM-LR. For all data blocks, \bar{L} is considerably higher for the mountainous AL* region. Note that both quantities being close to

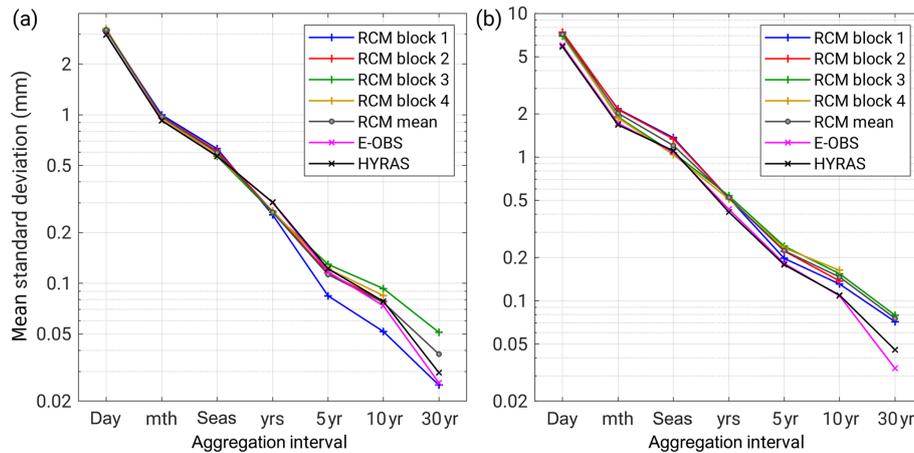


Figure 3. Mean standard deviation σ_T in mm (mean over data blocks) of spatially averaged precipitation aggregated over different time intervals: daily (day), monthly (month), seasonal (seas), yearly (years), and 5-/10-/30-year running mean (5 yr/10 yr/30 yr) for (a) ME* and (b) AL* (TP1b; 1951–2006). The four data blocks of LAERTES-EU are considered separately; RCM mean stands for the complete ensemble mean (gray). The results for E-OBS and HYRAS are given in black and magenta. Note that it is not possible to estimate the 30-year values for the decadal of data blocks 2 and 4.

Table 2. Coefficients of determination R^2 (top number) for the quantile–quantile contemplation of Fig. S2 and linear error in probability space \bar{L} (bottom number) between the RCM and both observations (E-OBS and HYRAS) for Mid-Europe (ME*) and the Alps (AL*), always using HYRAS grid cells only. Both quantities are based on daily spatial mean precipitation amounts.

RCM	E-OB		HYRAS	
	ME*	AL*	ME*	AL*
Data block 1	0.9914 0.0016	0.9924 0.0058	0.9876 0.0027	0.9835 0.0080
Data block 2	0.9914 0.0009	0.9925 0.0037	0.9878 0.0021	0.9848 0.0058
Data block 3	0.9963 0.0017	0.9976 0.0062	0.9936 0.0029	0.9930 0.0083
Data block 4	0.9966 0.0011	0.9981 0.0038	0.9943 0.0023	0.9938 0.0059

their optimum values does not indicate a perfect model. It rather means that the overall statistics regarding the entire range of intensities to a high degree coincide with the observations.

4.2 Time series

Besides overall statistics, other properties of LAERTES-EU like the temporal variability should cover the range of observations as well. Therefore, we analyze the time series of yearly values of different percentiles of the spatial mean precipitation for the investigation areas. In Fig. 4, the time series of the 99th percentile for ME* is shown. Both observa-

tional data sets have a high year-to-year variability with similar shape. The ensemble mean value of LAERTES-EU is higher, with a relative deviation of 1%–10% (TP1b average is 7%). The spread of both observational data sets is covered by the ensemble spread (minimum to maximum values) of LAERTES-EU except for few extreme peaks (e.g., 1985). In AL*, the E-OBS mean is about 5% higher than HYRAS but both time series have again a similar shape (Fig. S3). The ensemble mean again is higher with relative deviations of 12%–23% (16% on average) to E-OBS and 18%–29% (21% on average) to HYRAS. The ensemble spread also covers the observed variability.

Regarding more extreme values, namely the 99.9th percentile, similar results can be found (Figs. S4 and S5). Again, E-OBS and HYRAS show a similar behavior for both areas with mean value differences of less than 1%. The ensemble mean shows a mostly positive bias with deviations of less than 10% (6% on average during TP1b) compared to E-OBS for ME* and 6%–18% (average of 10%) for AL*. Furthermore, there is a distinctly higher spread and variability of the 99.9% for both the observations and LAERTES-EU. Except for a few peaks, LAERTES-EU covers the spread of the observations.

4.3 Added value of the sample size

In order to demonstrate the added value of the presented LAERTES-EU, we use the S2N (Eq. A6) for different sample sizes and return periods (cf. Appendix A). Sample size, in this case, means the number of simulation runs. Note that the simulations vary in length (number of years) with a minimum length of 10 years and a maximum of 110 years. In order to reduce the influence of the sample length on the results, the single simulation runs of LAERTES-EU were randomly con-

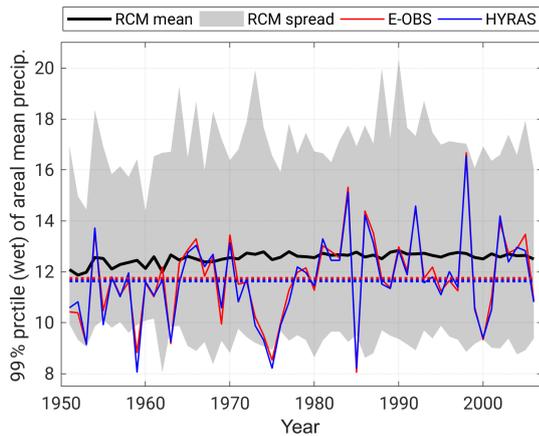


Figure 4. Time series of the yearly 99th percentile (wet days and HYRAS area only) of daily spatial mean precipitation values for Mid-Europe (ME*) during TP1b (1951–2006) of the LAERTES-EU ensemble mean (black), the ensemble spread (minimum to maximum; gray), E-OBS (red), and HYRAS (blue). The dotted lines symbolize the mean values of the observations throughout TP1b.

catenated using a 100-fold permutation. Observations have a sample size of 1. Again, S2N is calculated for daily spatial mean precipitation amounts during TP1b only using the HYRAS area.

For both ME* and AL*, S2N steadily increases with sample size for all calculated return values, indicating a more statistically robust estimate of the return values (Fig. 5). Furthermore, the S2N is lower for higher return periods which is a result of the increasing uncertainty of the best estimate due to fewer or even no data points for very high return periods. However, S2N also increases with sample size for the very high return periods. The robustness of a 2-year return value estimate of a sample of size 1 is about the same as the 1000-year estimate for a sample of size 20. This means that even for extremes, which have not been observed yet, some robust statistical analysis can be carried out.

5 Long-term variability and trends

The temporal evolution and variability of extreme precipitation throughout the past time period (TP1; 1900–2017) and also for the predictions (TP2; 2018–2028) are evaluated in this section. Besides time series of percentiles, we use climate change indices and statistical distributions. In this section, all land grid cells within the investigation areas ME and AL are used for calculating the daily areal mean precipitation amounts.

5.1 Precipitation distributions

Figure 6 shows the evolution of the distribution of areal mean precipitation throughout TP1 and TP2 by treating each decade independently. For the core of the distributions,

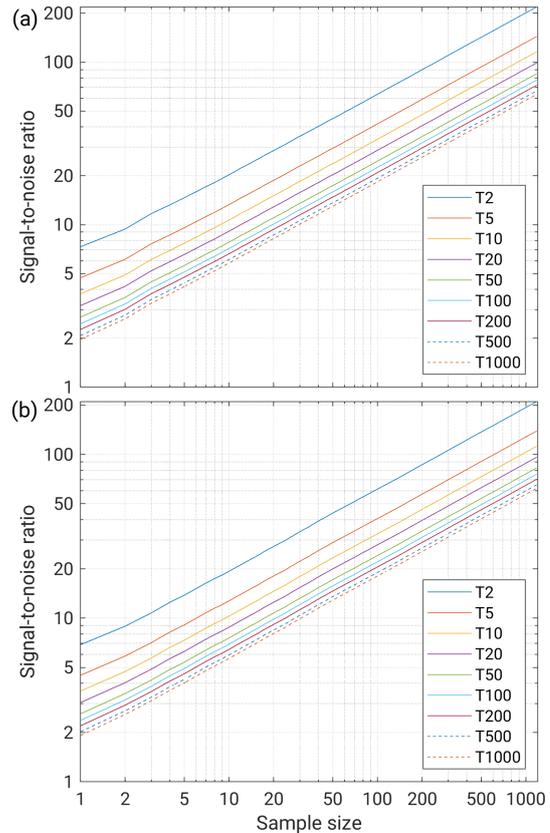


Figure 5. S2N ratio for different return periods T (colored lines) of daily spatial mean precipitation dependent on the sample size for (a) ME* and (b) AL*. The LAERTES-EU members were randomly stringed together permuting the order a hundred times. The shown S2N is the mean of this permutation.

namely medians, interquartile ranges, and upper whiskers, only small variance can be found between the different decades, which means that there is almost no change for the majority of the precipitation amounts. Nevertheless, a marked positive trend for the uppermost extremes of the distributions appears with maximum values around 18 mm d^{-1} at the beginning of the 20th century and about 24 mm d^{-1} in the 21st century. The distribution for the upcoming decade (2020–2028) shows only small differences to that of the present decade (since 2010), with an almost equal median and interquartile range but slightly higher maximum values (Fig. 6, green boxplot). Note that the decade of 2010–2019 contains the years 2018 and 2019 from the predictions, and that the last “decade” 2020–2028 is shorter with 9 years.

The boxplot for AL is shown in Fig. S6 and illustrates that not only the high percentiles reveal a decrease in the middle of the century, but the entire distribution is shifted towards lower values. Nevertheless, there is no clear tendency for the maximum values. For the upcoming decade, the distribution is similar to that of the present decade in the case of the median and the upper part of the distribution (Fig. S6,

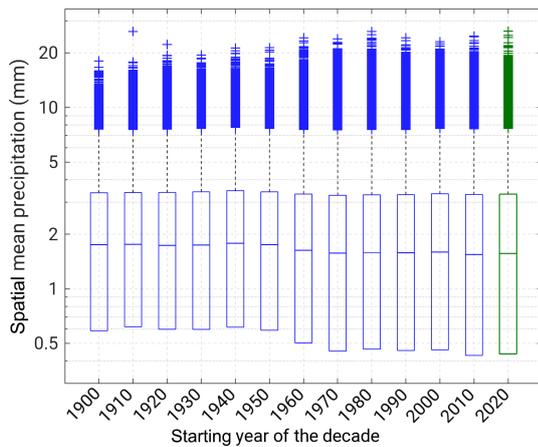


Figure 6. Boxplot of the distribution of daily spatial mean precipitation values (including dry days) for ME. Each decade was considered separately. The centerline of a box marks the median; the lower and upper ends of the box mark the 25th and 75th percentiles (interquartile range); the whiskers represent approximately the 99.9th percentile; the prediction part is marked in green.

green boxplot). The interquartile range is reduced due to an increased lower boundary of the boxplot.

5.2 Temporal evolution of yearly percentiles

5.2.1 Overview

The overall trend during TP1 and TP2 using a linear regression for both areas and percentiles is given in Table 3. While the ensemble mean shows a significant positive trend for ME for both percentiles, a small but significant negative trend can be found for the 99th percentile of AL, while there is almost no change in the 99.9th percentile of AL. In all cases, the ensemble spread increases due to both a decrease of the minimum values and an increase of the maximum values both being highly significant. The change of the maximums is stronger than the reduction of the minimums and more pronounced in AL than in ME.

Analogous to Table 3, we analyze the trend for TP1b only (Table S1 in the Supplement). The tendencies are the same for all cases but less pronounced, except for the mean 99.9% of AL where the negative trend during TP1b is slightly stronger than for the whole time series.

Figure 7 shows the temporal evolution of the 99th percentile during the 20th and the beginning of the 21st century for the whole LAERTES-EU. As given in Table 3, the lower boundary changes are small, while there is a visible positive trend of the ensemble mean and the upper boundary of the ensemble spread. Note that the larger spread from the 1960s onwards might be artificial due to the decisively larger number of members of data block 4. Nevertheless, there is a clear consistency in the time series for ME.

Table 3. Overall trend of daily spatial mean precipitation during TP1 and TP2 (1900–2028) using a linear regression of the yearly series of the 99th and 99.9th percentile (Pct; wet days only) for ME and AL. Given are absolute values and the relative changes (RCs) compared to the climatological mean (climTP; 1961–1990) for the ensemble minimum (min), the ensemble mean, and the ensemble maximum (max) percentile values within LAERTES-EU, and the related significance (p value; $\alpha = 0.05$).

Area	Pct	Variable	Trend (mm)	RC (%)	climTP (mm)	p_α
ME	99	min	−0.4	−4.6	7.8	0.9387
		mean	0.8	7.8	10.3	1.0000
		max	2.6	19.0	13.9	1.0000
ME	99.9	min	−1.0	−10.9	9.0	0.9974
		mean	1.1	8.4	13.5	1.0000
		max	6.7	31.0	21.6	1.0000
AL	99	min	−2.6	−17.1	15.4	1.0000
		mean	0.1	0.4	21.0	0.7208
		max	5.4	18.9	28.4	1.0000
AL	99.9	min	−4.3	−23.9	17.8	1.0000
		mean	−0.0	−0.0	27.3	0.0000
		max	9.0	20.0	44.7	1.0000

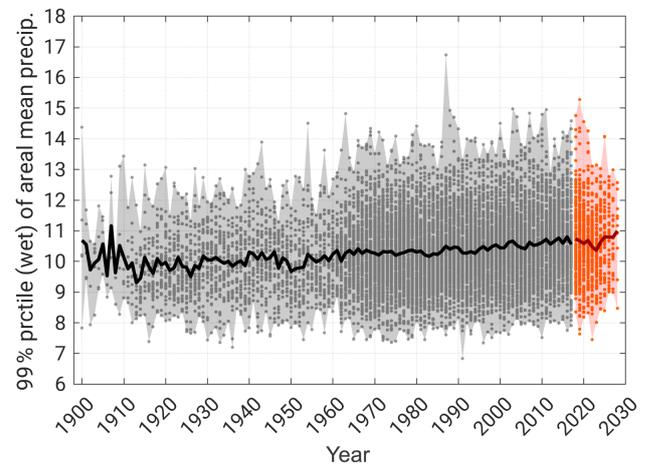


Figure 7. Time series of the yearly 99th percentile of daily spatial mean precipitation (wet days only) for Mid-Europe (ME; land only) of the LAERTES-EU ensemble mean (solid line), and the ensemble spread (minimum to maximum; dots and shaded area) during TP1 (1900–2017; black/gray) and TP2 (2018–2028; reddish).

Some differences emerge for AL (Fig. S7). At first, there is a distinct decrease of the ensemble mean between 1960 and 1970, which might be revealed from the rising number of members. As the ensemble matches well with the observations, we presume an overestimation of precipitation in the first half of the 20th century in that region, which could be a result of missing data for the applied dry-day correction. Due to the more complex terrain, the structure of the precip-

Table 4. Climatological mean (climTP; 1961–1990) of days per year exceeding the 99th and 99.9th percentiles (Pct; wet days only) for ME and AL, linear regression (LR) and relative change (RC) compared to climTP for different TPs, and related significance (p value; $\alpha = 0.05$).

Area	Pct	climTP	TP	LR	RC	p_α
ME	99	3.20	1 + 2	1.25	39 %	1.0000
			1b	0.76	24 %	1.0000
	99.9	0.60	1 + 2	0.36	60 %	1.0000
			1b	0.19	32 %	1.0000
AL	99	3.11	1 + 2	-0.17	-6 %	0.8262
			1b	-0.37	-12 %	0.9251
	99.9	0.62	1 + 2	-0.02	-2 %	0.7084
			1b	-0.04	-6 %	0.2973

itation fields is more complex and therefore more sensitive for different types of effects such as the dry-day correction.

The results for the 99.9th percentile are similar for both areas (Figs. S8 and S9). The positive trend for ME is even more pronounced, while the drop in the 1960s for AL is less visible, and therefore the time series is more constant.

For ME, the evolution of the number of days exceeding the climatological mean percentile reveals a strong positive and significant trend for both the 99th (Fig. 8, top) and 99.9th percentile (Fig. S10). The exact values of the climTP mean, the linear regression, the relative change, and the significance can be found in Table 4 (top numbers). For AL, the year-to-year variability is higher and the overall trend is slightly negative (Figs. 8, bottom, and S11) and at least significant for the 99th percentile. Again, we analyze the trend for TP1b separately (Table 4, bottom numbers). The tendencies for TP1b are the same but less pronounced except for the days exceeding the 99th percentile in AL, where there is a stronger trend signal in TP1b compared to the whole time series, which is also significant to a high degree.

5.2.2 Past trends and periodic oscillations

For a more detailed analysis of trends, the Mann–Kendall test described in Sect. 3.2 is applied to the time series of daily spatial mean precipitation percentiles. Figure 9a shows the relative number of LAERTES-EU members that show a positive or negative trend of the 99th percentile for ME. Only cases in which more than 60 % of the complete ensemble members reveal the same tendency are then considered for further investigation. For these cases, the ensemble mean trend is calculated (Fig. 9b) and the relative amount of significant members is displayed (Fig. 9c). All cases in which the ensemble reveals ambiguous tendencies are neglected (gray areas).

To a high degree, the single members show the same behavior, especially for the longer time series where positive trends are dominant. On a decadal timescale (diagonal line

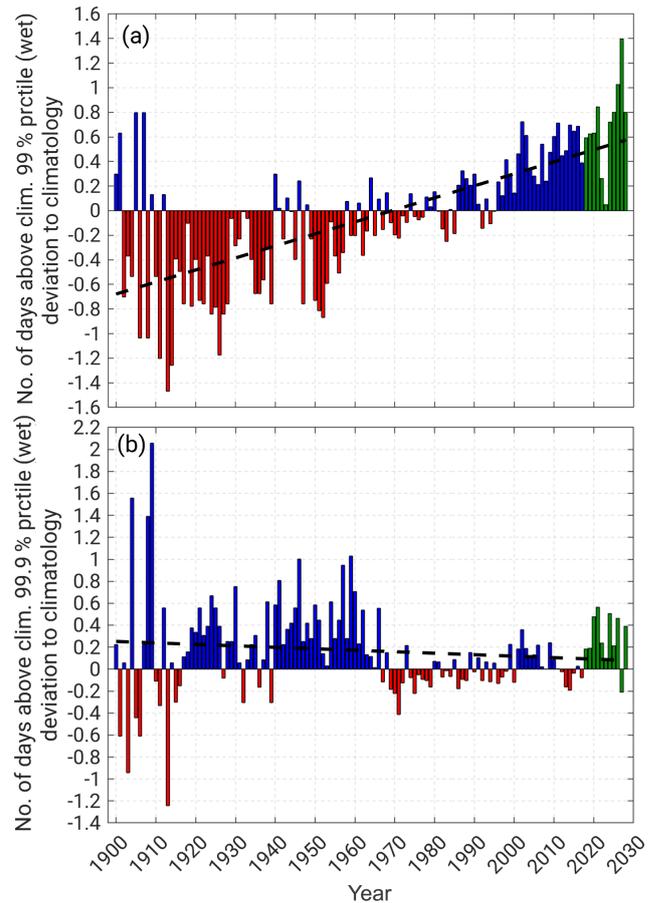


Figure 8. Deviation of the LAERTES-EU ensemble mean of the yearly number of days above the 99th percentile (wet days only) of daily spatial mean precipitation compared to the climatology (climTP; 1961–1990) for (a) Mid-Europe (ME) and (b) the Alps (AL). Red bars indicate negative anomalies (less days); blue bars indicate positive anomalies (more days). The predictions (TP2; 2018–2028) are given in green. The black line indicates a linear regression.

in Fig. 9), some oscillations appear with phases of increasing and decreasing precipitation. This signal might be smoothed, as it is not expected that the decadal simulations of data blocks 2 and 4 cover the natural variability at this timescale in detail. Furthermore, these simulations are not expected to be in phase with the long-lasting simulations of data blocks 1 and 3. The trends on this timescale reach rates of up to 0.1 mm a^{-1} or 1 mm per decade, respectively. The overall trend is weaker with a rate of $0\text{--}0.02 \text{ mm a}^{-1}$ or $0\text{--}2 \text{ mm}$ per century, respectively. Positive trends are more often significant than the negative, while only a small part of the ensemble shows significant trends. Similar results can be found for AL (Fig. S12). The trends on the decadal timescale reach higher rates but the oscillation is less pronounced than in ME. Again, most of the positive trends are significant, while just a few members with negative trends are significant.

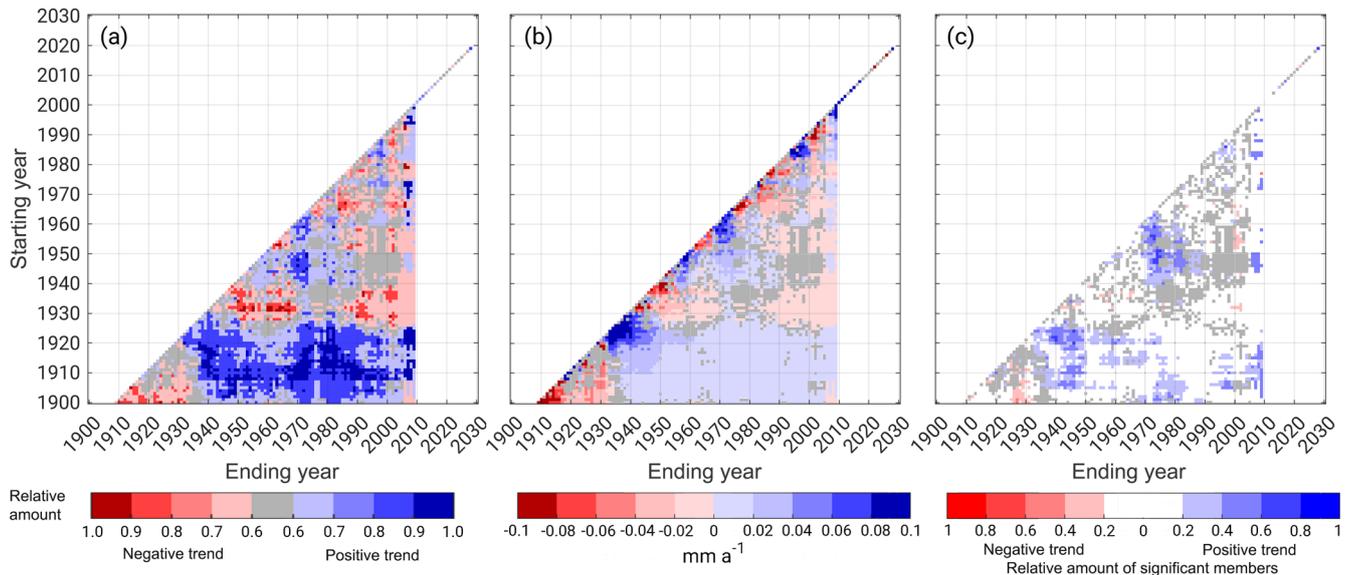


Figure 9. Trend analysis of the 99th percentile (wet days only) of daily spatial mean precipitation for ME with (a) the relative amount of members of LAERTES-EU with a positive (blue) or negative (red) trend; (b) the trend in millimeters per year averaged over the members from panel (a), and (c) relative amount of members from panel (a) with a significant trend; cases with no distinct number (less than 60 %) of members with same trend sign are marked in gray in panels (a)–(c).

For the 99.9th percentile of ME, large parts of LAERTES-EU show positive trends (Fig. S13). On the decadal timescale, a clear sequence of positive and negative trends is visible. Both the increases and decreases are more pronounced than for the 99th percentile but only a few members are significant. For AL, even more parts of the ensemble have the same tendency of heavy precipitation and a higher number of members have a significant trend (Fig. S14). These trends exceed rates of decisively more than $\pm 0.1 \text{ mm a}^{-1}$. In contrast to the results above, the 99.9th percentile for AL seems to have a multidecadal oscillation, while the overall trend of the complete time series is negative.

5.2.3 Future predictions

With respect to the upcoming decade (TP2; 2018–2028), LAERTES-EU predicts a continuation of the current trend with an increase especially for the 99.9th percentile (Figs. 7 and S6–S8; reddish area). In comparison to the last decade (2007–2017), the RCM mean of the 99th percentile increases of about 0.6 % for ME and about 2.1 % for AL. The 99.9th percentile increases about 2.0 % for ME and 3.0 % for AL.

Further to this absolute change, the number of days exceeding the climatological 99th percentile shows an increase of 4.9 % for ME and 8.4 % for AL, and 6.7 % (ME) and 22.4 % (AL) in the case of the 99.9 % compared to the mean of 2007–2017. This also manifests in the relative anomaly (Figs. 8 and S10–S11; green bars).

Nevertheless, a more detailed trend analysis illustrated in Fig. 9 and also Figs. S12–S14 reveals that LAERTES-EU shows no clear tendency for the 99 % during TP2. Just in a few cases, more than 60 % of the members have a similar mainly positive trend signal, which, however, is not significant. In the case of the 99.9th percentile, 60 %–70 % of the members show a strong positive trend of more than 0.1 mm a^{-1} with 20 %–40 % of them being significant. Although the tendency for TP2 is ambiguous and less significant, it shows continuity to the present decade.

5.3 Climate change indices

The results described in the previous sections also manifest in the considered ETCCDI climate change indices (Table 5). R95pTOT shows a positive trend for ME (Fig. 10a) with a relative change of about 18 % and a strong negative trend of approximately -15% for AL (Fig. S15). Remarkably, there is a high positive deviation in the first half of the 20th century compared to the climTP amount for AL which might be artificial due to the mentioned problems of the dry-day correction. R99pTOT shows a positive change for ME (Fig. 10b) and a slightly negative trend for AL (Fig. S16). The overestimation for AL in the early century is less pronounced for this index. Considering only the TP1b, the tendencies are the same in all cases. The positive trends for ME are less pronounced, while the negative trends for AL are stronger. The estimated trends are highly significant, except for the R99pTOT of AL for the whole time series.

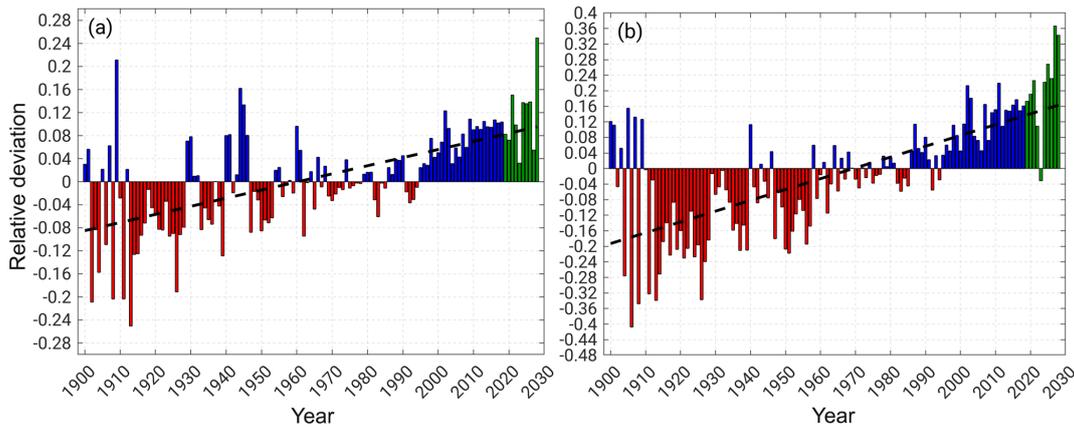


Figure 10. Relative deviation of (a) the R95pTOT index and (b) the R99pTOT index of the LAERTES-EU ensemble mean of daily spatial mean precipitation (wet days and land only) compared to the climatology (climTP; 1961–1990; Table 5) for Mid-Europe (ME). Red bars indicate negative (dry) anomalies; blue bars indicate positive (wet) anomalies. The predictions (TP2; 2018–2028) are given in green. The black line indicates a linear regression.

Table 5. Climatological mean (climTP; 1961–1990) of ETCCDI quantities for Mid-Europe (ME) and the Alps (AL), linear regression (LR) and relative change (RC) compared to climTP for different TPs, and related significance (p value; $\alpha = 0.05$). Both indices are based on wet days only of daily spatial mean precipitation (land only).

Area	ETCCDI	climTP (mm)	TP	LR (mm)	RC (%)	p_α
ME	R95pTOT	157.5	1 + 2 1b	28.4 20.1	18 13	1.0 1.0
	R99pTOT	43.8	1 + 2 1b	15.6 12.2	36 28	1.0 1.0
AL	R95pTOT	306.7	1 + 2 1b	-46.3 -54.3	-15 -18	1.0 1.0
	R99pTOT	88.5	1 + 2 1b	-4.5 -10.8	-5 -12	0.8953 0.9891

Compared to the present decade, the predictions show a continuation of the positive trend for ME with an increase of 2 % for R95pTOT and 5 % for R99pTOT. In contrast, both indices show a positive trend for AL with an increase of 7 % for R95pTOT and 8 % for R99pTOT, which is a complete reversal of the overall trend.

6 Summary and conclusions

We have presented the novel LAERTES-EU ensemble combining various regional climate model simulations done with COSMO-CLM to analyze long-term variability and trends of flood-related intensive areal precipitation across central Europe. The whole RCM ensemble was divided into four data blocks depending on forcing data, assimilation schemes, or the initialization of the driving global model MPI-ESM. The

setup of the COSMO model remained the same for all simulations. In total, the presented LAERTES-EU consists of over 1100 simulation runs with approximately 12 500 simulated years on a 25 km horizontal resolution.

The focus of investigation was laid on the PRUDENCE regions Mid-Europe (ME) and the Alps (AL). Regarding intensive areal precipitation, we concentrated on high percentiles, namely 99 % and 99.9 %, of spatially averaged daily precipitation amounts. Note that it was not expected that LAERTES-EU was able to reproduce historical precipitation events on a daily base in detail but have a more accurate performance regarding long-term variations and statistical distributions on a larger scale perspective. Furthermore, the given resolution restricts the consideration of convective processes, so we concentrated on larger-scale phenomena.

With respect to our initial research questions, the following main conclusions can be drawn and summed up out of the presented results, which will be discussed in more detail afterwards:

1. LAERTES-EU is capable of representing the range of extreme areal precipitation similar to the used observational data sets and also fits into the range of previous studies (e.g., Früh et al., 2010). The four data blocks are consistent and have similar precipitation distributions. The ensemble also covers the observed temporal evolution.
2. The benefits of the large ensemble size manifest in a strong increase of the signal-to-noise ratio beyond the typically used ensemble sizes and in high statistical significance of estimated trends for the ensemble mean. Furthermore, the distribution of precipitation totals is represented in a more concise way, taking the limitations of the considered observations into account.

3. Long-term trends reveal spatial differences in sign and strength. These tendencies are partly significant. Despite a quite large ensemble spread, the ensemble mean shows more explicit results. Distinct oscillations can also be found on shorter timescales (e.g., decades).
4. The predictions for the upcoming decade show a continuation of past tendencies in terms of both intensity and occurrence frequency for ME without any discontinuity to the previous time period. On the other hand, LAERTES-EU shows no clear signal for AL.

Regarding the validation (1), grid-point-based IPCs, areal mean precipitation distributions (internal variability σ_T and linear error in probability space \bar{L}), and Q–Q distributions have been analyzed. In all cases, the IPCs of the simulations show an overestimation of precipitation on the order of 10%–20% compared to E-OBS. Haylock et al. (2008) found that E-OBS can have a certain negative bias of up to 30% compared to single ground-based punctual observations. Taking this into account, the IPCs are almost coincident. Furthermore, the IPCs of LAERTES-EU show only small deviation compared to the HYRAS data set (aggregated to the model grid). The IPCs and also the Q–Q distributions of all four data blocks are coincident, which was a prerequisite for the combination to one large ensemble. The Q–Q distributions of spatially aggregated mean precipitation reveal fewer differences between LAERTES-EU and E-OBS, but an underestimation of simulated rainfall compared to HYRAS by about 10%. The linear error in probability space \bar{L} shows a good agreement of LAERTES-EU with observations in terms of the distribution of daily areal mean precipitation totals. For different aggregation intervals from daily values up to 10-year running means, the internal variability (standard deviation σ_T) of LAERTES-EU matches to a high degree with that of both observations. Note that both quantities \bar{L} and σ_T do not indicate whether the simulated absolute precipitation values coincide with the observations but rather show the agreement of statistical properties.

Regarding (2), LAERTES-EU reveals a clear added value due to the large sample size. Estimates of long return periods are more robust compared to smaller ensembles which is of importance, for instance, for risk and insurance applications. Furthermore, trends at least in the ensemble mean are highly significant. The IPCs also show a benefit of RCM data compared to the coarser global model (MPI-ESM) or the 20CR global reanalysis. Regarding extremes, LAERTES-EU includes a broader range of precipitation totals with even higher values, which are not covered by observations due to their limited temporal availability. Although the presented results reveal a broad range of realizations within LAERTES-EU, the statistics of the ensemble mean clearly benefit from the large ensemble size with a better signal-to-noise ratio.

Besides a proper representation of precipitation, long-term trends and temporal variations were of special interest. Regarding (3), the presented results show a reasonable agree-

ment of LAERTES-EU concerning the temporal evolution of the considered percentiles of spatially aggregated daily precipitation totals for the different investigation areas. The ensemble spread (minimum to maximum) covers the observed variability except a few peaks. The ensemble mean shows a small positive bias compared to both observational data sets. Throughout the complete time period of TP1 (1900–2017), positive and significant trends can be found for ME in both percentiles (99% and 99.9%) and also in the number of days exceeding the climatological mean (1961–1990). For AL, there is no clear trend signal in the ensemble mean but an increase in the maximum values. In contrast, the number of days exceeding the climatological mean percentiles is decreasing in this area. Comparing the trends of TP1 to the shorter TP1b (1951–2006), the tendencies are the same but less pronounced in TP1b. On a decadal timescale, some oscillations can be found with periods of increasing precipitation and such with decreasing values. Similar results as for time series of percentiles can be found using climate change indices (ETCCDI).

Regarding (4), the predictions for the next decade (2018–2028; TP2) reveal ongoing tendencies of heavy precipitation indices. A special case is AL where the slightly negative trends in the past (TP1) turn to positive ones. Both the continuity for ME and the reversal for AL appear in all time series, namely the number of days of threshold exceedance, ETCCDI variables, and investigated percentiles. While there is a clear signal and high significance for the ensemble mean, the trends were ambiguous and less significant when the ensemble members were considered separately. However, we conclude that this tendency is likely, as it is a continuation of the results of the present decade. Similar results for parts of LAERTES-EU were found by Meyers et al. (2019).

Precipitation remains a challenging task for both reanalyses and climate model simulations of the past and the future with partly contrasting results shown by several previous studies. Furthermore, long-term comprehensive observations are not available, which makes a validation difficult due to the high spatial variability of precipitation. This also affects analyses of trends or climate variability. What is known is a theoretical increase of the water vapor capacity according to the Clausius–Clapeyron (CC) equation of about 6%–7% per degree of temperature increase (e.g., Trenberth et al., 2003; Berg et al., 2009), which assumes a near-constant relative humidity. The CC rate is generally thought to be a proxy for future precipitation projections (Westra et al., 2013). A recent discussion about the validity of the CC rate as an estimate for future projections of heavy precipitation can be found in Zhang et al. (2017). They pointed out that besides the thermodynamic responses, changes in heavy precipitation may be also influenced by dynamical effects. Furthermore, Pfahl et al. (2017) and Kröner et al. (2017) showed that precipitation trends can be regionally influenced by contributions from both lapse-rate and circulation effects.

The ensemble mean of LAERTES-EU shows an increase of about 1.9 °C for ME and 2.3 °C for AL for the yearly mean 2 m temperature of spatial means during the 20th century (TP1; 1900–2017). Including the predictions (TP2), the increase is about 2.4 °C for ME and 2.8 °C for AL. For instance, Simmons et al. (2017) found an increase over European land masses of approximately 2 °C in the mean compared to pre-industrial conditions. Moberg et al. (2006) found an increase of about 1 °C for temperature extremes. Thus, LAERTES-EU is within the range of observed changes. The increase in temperature over the entire time period is equivalent to a CC scaling of about 15 %–20 %. The extracted changes of the high precipitation percentiles for ME make up to 50 % compared to the theoretical CC value. However, the negative tendencies for AL do not fit into this theoretical estimate.

The presented LAERTES-EU data set can be used for various applications fields. In particular, the simulations are used as input for hydrological modeling and further applications such as flood risk assessments. The presented ensemble in this case can be used as a stochastic weather generator treating the single simulations independently. This leads to the production of a quasi-stochastic hydrological discharge data set. Due to the large ensemble size, estimates of high return periods become more robust. However, it has to be mentioned that the composition of the four data blocks to one ensemble restricts the temporal homogeneity. Moreover, the validation showed a positive bias of the ensemble mean which, together with the overestimation of low intensities, requires a bias correction to avoid unrealistic discharges. This application as well as the bias correction of LAERTES-EU will be addressed in a consecutive study.

In this study, we have focused on all-year variance, oscillations, or trends. Future investigations can address a seasonal differentiated analysis of trends and oscillations as well as a more detailed investigation of the spatial distribution of these findings and potential mechanisms behind the observed variability. Previous studies indicated that there is a strong relation between precipitation in Europe and the North Atlantic Oscillation (NAO), especially during wintertime (e.g., Hurrell, 1995; Rîmbu et al., 2002; Haylock and Goodess, 2004; Nissen et al., 2010; Pinto and Raible, 2012). Moreover, Casanueva et al. (2014) found a connection between extreme precipitation and the Atlantic Multidecadal Oscillation (AMO) during the whole year.

Appendix A: Statistical quantities

The linear error in probability space L uses the difference of probabilities ΔC defined as

$$\Delta C_r(x) = \text{ecdf}_{\text{mod},r}(x) - \text{ecdf}_{\text{obs}}(x), \tag{A1}$$

where $\text{ecdf}_{\text{mod},r}$ is the empirical cumulative density function of the model run r , and ecdf_{obs} that of the observation up to precipitation intensity x . The linear error in probability space L_r for a model run r is then defined as (Déqué, 2012; Wahl et al., 2017)

$$L_r = \frac{1}{n} \cdot \sum_{x=1}^n |\Delta C_r(x)|. \tag{A2}$$

L_r describes the mean value of ΔC_r over the entire range of precipitation intensities x grouped into n classes. Using absolute values avoids a compensation of positive and negative values. The better both distributions coincide, the lower the value of L_r . The ensemble mean of L_r is given by

$$\bar{L} = \frac{1}{M} \sum_{r=1}^M L_r, \tag{A3}$$

with M being the total number of simulation runs.

The model performance on different frequency intervals is further validated using the standard deviation of a gamma distribution σ_Γ (Wilks, 2006), which is given by

$$\sigma_\Gamma^2 = \alpha \beta^2. \tag{A4}$$

In this formulation, α is the shape parameter of the gamma distribution, and β its scale parameter.

The quantile–quantile analysis uses the Pearson correlation coefficient (Wilks, 2006) given by

$$R = \frac{\sum_{k=1}^N \{[y_k - \bar{y}_k] \cdot [x_k - \bar{x}_k]\}}{\sqrt{\sum_{k=1}^N [x_k - \bar{x}_k]^2} \cdot \sqrt{\sum_{k=1}^N [y_k - \bar{y}_k]^2}}, \tag{A5}$$

with the data series x and y of length N . The range of R is $R \in [-1; +1]$ with a perfect anti-correlation at $R = -1$ and a perfect correlation at $R = +1$.

The S2N ratio in this study is defined as

$$\text{S2N} = \frac{\text{RV}_{T,\text{Gumbel}}}{\text{CI}_{90,T}}, \tag{A6}$$

with the return level RV of the Gumbel distribution at return period T divided by its 90 % confidence interval at T (Früh et al., 2010). Small values of S2N indicate a more uncertain estimate; high values indicate a more robust one. The Gumbel distribution (Wilks, 2006) is an extreme value type-I

distribution and often used for return period estimation. Its cumulative density function (cdf) is given by

$$F(x) = \exp\left(-\exp\left(-\frac{x-\beta}{\alpha}\right)\right), \tag{A7}$$

with the free parameters $\beta = \sigma\sqrt{6} \cdot \pi^{-1}$ and $\alpha = \bar{x} - \gamma\beta$, where σ is the standard deviation of the sample x assuming a normal distribution, and $\gamma = 0.57721$ Euler’s constant. For x , usually a series of yearly maximum values is used. The relationship between the cdf and the return period T is given by (Wilks, 2006)

$$T = \frac{1}{1 - F(x)}. \tag{A8}$$

Appendix B: ETCCDI quantities

Two out of the 27 indices introduced and recommended by the Expert Team on Climate Change Detection and Indices⁴ (ETCCDI; Karl et al., 1999; Peterson, 2005) are used in this study. R95pTOT describes the annual total precipitation sum of all values above the climatological 95th percentile of wet days ($\text{RR} > 1 \text{ mm}$) during the reference period of 1961–1990. The R95pTOT of the year k is defined as

$$\text{R95pTOT}_k = \sum_{w=1}^W \text{RR}_{w_k} \forall \text{RR}_{w_k} > \text{RR}_{p95}, \tag{B1}$$

where RR_{w_k} is the daily precipitation amount on a wet day during year k , RR_{p95} is the climatological 95th percentile, and W is the total number of wet days in year k . Analogously, the R99pTOT is defined by replacing the 95th with the 99th percentile:

$$\text{R99pTOT}_k = \sum_{w=1}^W \text{RR}_{w_k} \forall \text{RR}_{w_k} > \text{RR}_{p99}. \tag{B2}$$

Appendix C: Trends and significance

A Mann–Kendall test (Mann, 1945; Kendall, 1955) is performed for the detection of trends and its related significance. To account for possible oscillations within long time series, we first split the complete time series into subseries with a minimum length of 10 years and up to over 100 years (trend matrix). The Mann–Kendall test uses a standardized test statistic S_τ following a standard Gaussian distribution (SGD). S_τ is given by

$$S_\tau = \begin{cases} \frac{\tau-1}{\sqrt{\sigma_\tau^2}}, & \tau > 0, \\ 0, & \tau = 0, \\ \frac{\tau+1}{\sqrt{\sigma_\tau^2}}, & \tau < 0. \end{cases} \tag{C1}$$

⁴<http://etccdi.pacificclimate.org/> (last access: 20 May 2020).

Here, τ is known as the Kendall τ and σ_τ^2 is the variance of the SGD. A detected trend is significant if S_τ lies within the upper and lower quantiles z of the SGD at a given significance level α with $S_\tau \in \left[z_{\frac{\alpha}{2}} \sigma_\tau; z_{1-\frac{\alpha}{2}} \sigma_\tau \right]$, respectively (Yue et al., 2002).

Yue et al. (2002) pointed out some weaknesses of the Mann–Kendall test in the case of inherent autocorrelation. To avoid a distortion of the statistic by autocorrelation, Yue et al. (2002) presented the trend-free pre-whitening (TFPW) method. The first step is the estimation of a linear trend between two time steps ($t = i$ and $t = j$) using the Theil–Sen approach (TSA; Theil, 1950; Sen, 1968). The slope b of this linear regression is given by

$$b = \text{median} \left(\frac{x_j - x_i}{j - i} \right), \forall i < j. \quad (\text{C2})$$

In a second step, the original time series x is detrended by subtracting b at each time step t :

$$x'_t = x_t - b \cdot t. \quad (\text{C3})$$

Afterwards, the lag-1 autocorrelation coefficient r_1 is removed from the trend-free series x' :

$$x''_t = x'_t - r_1 \cdot x'_{t-1}, \quad (\text{C4})$$

where r_1 is given by

$$r_1 = \frac{\frac{1}{N-1} \cdot \sum_{i=1}^{N-1} (x'_i - \bar{x}') \cdot (x'_{i+1} - \bar{x}')}{\frac{1}{N} \cdot \sum_{i=1}^N (x'_i - \bar{x}')^2}. \quad (\text{C5})$$

The modified TFPW time series x^* results by re-adding the TSA slope b :

$$x^*_t = x''_t + b \cdot t. \quad (\text{C6})$$

This modified time series conserves the trend but is free of autocorrelation. The Mann–Kendall test is performed on the TFPW time series x^* . According to Yue et al. (2002), TFPW has to be considered in cases with non-zero TSA slope and significant lag-1 autocorrelation. The significance of a trend or autocorrelation is tested on the 90 % ($\alpha = 0.1$), 95 % ($\alpha = 0.05$), and 99 % ($\alpha = 0.01$) significance levels.

Data availability. The E-OBS data (Haylock et al., 2008) are available online after registration at <https://www.ecad.eu/download/ensembles/ensembles.php> (last access: 20 May 2020). The 20CR data (Compo et al., 2011) can be found on https://www.esrl.noaa.gov/psd/data/20thC_Rean/ (last access: 20 May 2020). HYRAS (Rauthe et al., 2013) can be requested at the German Weather Service (DWD). The RCM data (MiKlip data) will be made available via the CERA database (<http://cera-www.dkrz.de/>; last access: 20 May 2020) of the German Climate Computing Center (DKRZ).

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/esd-11-469-2020-supplement>.

Author contributions. FE, LAK, HF, and JGP designed the study. HF performed (parts of) the RCM simulations. LAK applied the dry-day correction. FE did the analysis and plots, and wrote the initial draft. All authors contributed with discussions and revisions.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Large Ensemble Climate Model Simulations: Exploring Natural Variability, Change Signals and Impacts”. It is not associated with a conference.

Acknowledgements. The authors thank the National Centers for Environmental Prediction (NCEP) for providing the 20CR data. We acknowledge the E-OBS data set from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>, last access: 20 May 2020) and the data providers in the ECA & D project (<http://www.ecad.eu>, last access: 20 May 2020). We also thank the German Weather Service (DWD) for providing HYRAS. In addition, we thank the Max Planck Institute for Meteorology (MPI-M) and DWD for the global model simulations and the German Climate Computing Center (DKRZ, Hamburg) for computing and storage resources. We also thank Martin Kadlec (AON) for discussions. We thank the reviewers for their valuable comments that helped to improve this study, and the handling editor for guidance throughout the entire process.

Financial support. We thank AON for funding the project “Hydrometeorological extreme events under recent climate conditions”. We also thank the BMBF MiKlip project II (FKZ: 01 LP 1518 A/D) and ClimXtreme Project (FKZ 01 LP 1901 A) for partial funding. Joaquim G. Pinto thanks the AXA Research Fund for support (<https://axa-research.org/en/project/joaquim-pinto>, last access: 20 May 2020). We thank open-access publishing fund of the Karlsruhe Institute of Technology (KIT).

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Review statement. This paper was edited by Ralf Ludwig and reviewed by Raul R. Wood and two anonymous referees.

References

- Barbero, R., Fowler, H. J., Lenderink, G., and Blenkinsop, S.: Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions?, *Geophys. Res. Lett.*, 44, 974–983, <https://doi.org/10.1002/2016GL071917>, 2017.
- Berg, P., Haerter, J. O., Thejll, P., Piani, C., Hagemann, S., and Christensen, J. H.: Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature, *J. Geophys. Res.-Atmos.*, 114, D18102, <https://doi.org/10.1029/2009JD012008>, 2009.
- Berg, P., Feldmann, H., and Panitz, H.-J.: Bias correction of high resolution regional climate model data, *J. Hydrol.*, 448–449, 80–92, <https://doi.org/10.1016/j.jhydrol.2012.04.026>, 2012.
- Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Müller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., and Eade, R.: The decadal climate prediction project (DCPP) contribution to CMIP6, *Geosci. Model Dev.*, 9, 3751–3777, <https://doi.org/10.5194/gmd-9-3751-2016>, 2016.
- Brönnimann, S.: Weather extremes in an ensemble of historical reanalyses, in: *Historical Weather Extremes in Reanalyses*, *Geographica Bernensia*, g92, 7–22, <https://doi.org/10.4480/GB2017.G92.01>, 2017.
- Brönnimann, S., Romppainen-Martius, O., Franke, J., Stickler, A. N., and Auchmann, R.: Historical weather extremes in the “Twentieth Century Reanalysis”, in: *Weather extremes during the past 140 years*, *Geographica Bernensia*, g89, 7–17, <https://doi.org/10.4480/GB2013.G89.01>, 2013.
- Cahill, N., Rahmstorf, S., and Parnell, A. C.: Change points of global temperature, *Environ. Res. Lett.*, 10, 084002, <https://doi.org/10.1088/1748-9326/10/8/084002>, 2015.
- Casanueva, A., Rodríguez-Puebla, C., Frías, M. D., and González-Reviriego, N.: Variability of extreme precipitation over Europe and its relationships with teleconnection patterns, *Hydrol. Earth Syst. Sci.*, 18, 709–725, <https://doi.org/10.5194/hess-18-709-2014>, 2014.
- Christensen, J. H. and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, *Climatic Change*, 81, 7–30, <https://doi.org/10.1007/s10584-006-9210-7>, 2007.
- Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, *Q. J. Roy. Meteorol. Soc.*, 137, 1–28, <https://doi.org/10.1002/qj.776>, 2011.

- Déqué, M.: Deterministic Forecasts of Continuous Variables, in: chap. 5, Forecast Verification, edited by: Jolliffe, I. T. and Stephenson, D. B., John Wiley & Sons Ltd, 77–94, <https://doi.org/10.1002/9781119960003.ch5>, 2012.
- Dittus, A. J., Karoly, D. J., Lewis, S. C., Alexander, L. V., and Donat, M. G.: A Multiregion Model Evaluation and Attribution Study of Historical Changes in the Area Affected by Temperature and Precipitation Extremes, *J. Climate*, 29, 8285–8299, <https://doi.org/10.1175/JCLI-D-16-0164.1>, 2016.
- Easterling, D. R., Evans, J. L., Groisman, P. Y., Karl, T. R., Kunkel, K. E., and Ambenje, P.: Observed Variability and Trends in Extreme Climate Events: A Brief Review, *B. Am. Meteorol. Soc.*, 81, 417–426, [https://doi.org/10.1175/1520-0477\(2000\)081<0417:OVATIE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<0417:OVATIE>2.3.CO;2), 2000.
- Ehmele, F. and Kunz, M.: Flood-related extreme precipitation in southwestern Germany: development of a two-dimensional stochastic precipitation model, *Hydrol. Earth Syst. Sci.*, 23, 1083–1102, <https://doi.org/10.5194/hess-23-1083-2019>, 2019.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Feldmann, H., Früh, B., Schädler, G., Panitz, H.-J., Keuler, K., Jacob, D., and Lorenz, P.: Evaluation of the precipitation for South-western Germany from high resolution simulations with regional climate models, *Meteorol. Z.*, 17, 455–465, <https://doi.org/10.1127/0941-2948/2008/0295>, 2008.
- Feldmann, H., Schädler, G., Panitz, H.-J., and Kottmeier, C.: Near future changes of extreme precipitation over complex terrain in Central Europe derived from high resolution RCM ensemble simulations, *Int. J. Climatol.*, 33, 1964–1977, <https://doi.org/10.1002/joc.3564>, 2013.
- Feldmann, H., Pinto, J. G., Laube, N., Uhlig, M., Moemken, J., Pasternack, A., Früh, B., Pohlmann, H., and Kottmeier, C.: Skill and Added Value of the MiKlip Regional Decadal Prediction System for Temperature over Europe, *Tellus A*, 71, 1–19, <https://doi.org/10.1080/16000870.2019.1618678>, 2019.
- Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., and Zahn, M.: Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples, *B. Am. Meteorol. Soc.*, 92, 1181–1192, <https://doi.org/10.1175/2011BAMS3061.1>, 2011.
- Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M. E.: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, in: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp., <https://doi.org/10.1017/CBO9781139177245>, 2012.
- Fischer, E. M. and Knutti, R.: Observed heavy precipitation increase confirms theory and early models, *Nat. Clim. Change*, 6, 986–991, <https://doi.org/10.1038/NCLIMATE3110>, 2016.
- Folland, C. K., Boucher, O., Colman, A., and Parker, D. E.: Causes of irregularities in trends of global mean surface temperature since the late 19th century, *Sci. Adv.*, 4, eaao5297, <https://doi.org/10.1126/sciadv.aao5297>, 2018.
- Früh, B., Feldmann, H., Panitz, H.-J., Schädler, G., Jacob, D., Lorenz, P., and Keuler, K.: Determination of precipitation return values in complex terrain and their evaluation, *J. Climate*, 23, 2257–2274, 2010.
- Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *J. Hydrol.*, 228, 113–129, [https://doi.org/10.1016/S0022-1694\(00\)00144-X](https://doi.org/10.1016/S0022-1694(00)00144-X), 2000.
- Hagemann, S., Loew, A., and Andersson, A.: Combined evaluation of MPI-ESM land surface water and energy fluxes, *J. Adv. Model. Earth Syst.*, 5, 259–286, <https://doi.org/10.1029/2012MS000173>, 2013.
- Haylock, M. R. and Goodess, C. M.: Interannual variability of European extreme winter rainfall and links with mean large-scale circulation, *Int. J. Climatol.*, 24, 759–776, <https://doi.org/10.1002/joc.1033>, 2004.
- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.-Atmos.*, 113, D20119, <https://doi.org/10.1029/2008JD010201>, 2008.
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood risk under climate change, *Nat. Clim. Change*, 3, 816–821, <https://doi.org/10.1038/nclimate1911>, 2013.
- Hofstra, N., Haylock, M., New, M., and Jones, P. D.: Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature, *J. Geophys. Res.-Atmos.*, 114, D21101, <https://doi.org/10.1029/2009JD011799>, 2009.
- Hurrell, J. W.: Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation, *Science*, 269, 676–679, <https://doi.org/10.1126/science.269.5224.676>, 1995.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsman, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, *Reg. Environ. Change*, 14, 563–578, <https://doi.org/10.1007/s10113-013-0499-2>, 2014.
- Journel, A. G. and Huijbregts, C. J.: Mining Geostatistics, Academic Press, London, 600 pp., 1978.
- Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.-S.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-earth system model, *J. Adv. Model. Earth Syst.*, 5, 146–172, <https://doi.org/10.1002/jame.20023>, 2013.
- Kanamitsu, M., Alpert, J. C., Campana, K. A., Caplan, P. M., Deaven, D. G., Iredell, M., Katz, B., Pan, H.-L., Sela, J., and White, G. H.: Recent changes implemented into the global forecast system at NMC, *Weather Forecast.*, 6, 425–435, [https://doi.org/10.1175/1520-0434\(1991\)006<0425:RCITG>2.0.CO;2](https://doi.org/10.1175/1520-0434(1991)006<0425:RCITG>2.0.CO;2), 1991.
- Karl, T. R., Nicholls, N., and Ghazi, A. (Eds.): CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary, in: Weather and Climate Extremes, Springer, Dordrecht, 3–7, https://doi.org/10.1007/978-94-015-9265-9_2, 1999.

- Kendall, M. G.: Rank correlation methods, Charles Griffin, London, UK, 196 pp., 1955.
- Kienzler, S., Pech, I., Kreibich, H., Müller, M., and Thielen, A. H.: After the extreme flood in 2002: changes in preparedness, response and recovery of flood-affected residents in Germany between 2005 and 2011, *Nat. Hazards Earth Syst. Sci.*, 15, 505–526, <https://doi.org/10.5194/nhess-15-505-2015>, 2015.
- Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, *Geosci. Model Dev.*, 7, 1297–1333, <https://doi.org/10.5194/gmd-7-1297-2014>, 2014.
- Kröner, N., Kotlarski, S., Fischer, E., Lüthi, D., Zubler, E., and Schär, C.: Separating climate change signals into thermodynamic, lapse-rate and circulation effects: theory and application to the European summer climate, *Clim. Dynam.*, 48, 3425–3440, <https://doi.org/10.1007/s00382-016-3276-3>, 2017.
- Mann, H. B.: Nonparametric Tests Against Trend, *Econometrica*, 13, 245–259, <https://doi.org/10.2307/1907187>, 1945.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Thiemel, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, 1–34, <https://doi.org/10.1029/2009RG000314>, 2010.
- Marotzke, J., Müller, W. A., Vamborg, F. S. E., Becker, P., Cubasch, U., Feldmann, H., Kaspar, F., Kottmeier, C., Marini, C., Polkova, I., Prömmel, K., Rust, H. W., Stammer, D., Ulbrich, U., Kadow, C., Köhl, A., Kröger, J., Kruschke, T., Pinto, J. G., Pohlmann, H., Reyers, M., Schröder, M., Sienz, F., Timmreck, C., and Ziese, M.: MiKlip: a national research project on decadal climate prediction, *B. Am. Meteorol. Soc.*, 97, 2379–2394, <https://doi.org/10.1175/BAMS-D-15-00184.1>, 2016.
- Martel, J.-L., Mailhot, A., and Brissette, F.: Global and Regional Projected Changes in 100-yr Subdaily, Daily, and Multiday Precipitation Extremes Estimated from Three Large Ensembles of Climate Simulations, *J. Climate*, 33, 1089–1103, <https://doi.org/10.1175/JCLI-D-18-0764.1>, 2020.
- Merz, B., Elmer, F., Kunz, M., Mühr, B., Schröter, K., and Uhlemann-Elmer, S.: The extreme flood in June 2013 in Germany, *La Houille Blanche*, 1, 5–10, <https://doi.org/10.1051/hb/2014001>, 2014.
- Mieruch, S., Feldmann, H., Schädler, G., Lenz, C.-J., Kothe, S., and Kottmeier, C.: The regional MiKlip decadal forecast ensemble for Europe: the added value of downscaling, *Geosci. Model Dev.*, 7, 2983–2999, <https://doi.org/10.5194/gmd-7-2983-2014>, 2014.
- Moberg, A. and Jones, P. D.: Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901–99, *Int. J. Climatol.*, 25, 1149–1171, <https://doi.org/10.1002/joc.1163>, 2005.
- Moberg, A., Jones, P. D., Lister, D., Walther, A., Brunet, M., Jacobeit, J., Alexander, L. V., Della-Marta, P. M., Luterbacher, J., Yiou, P., Chen, D., Klein Tank, A. M. G., Saladié, O., Sigró, J., Aguilar, E., Alexandersson, H., Almarza, C., Auer, I., Barriendos, M., Begert, M., Bergström, H., Böhm, R., Butler, C. J., Caesar, J., Drebs, A., Founda, D., Gerstengarbe, F.-W., Micela, G., Maugeri, M., Österle, H., Pandzic, K., Petrakis, M., Srnec, L., Tolasz, R., Tuomenvirta, H., Werner, P. C., Linderholm, H., Philipp, A., Wanner, H., and Xoplaki, E.: Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000, *J. Geophys. Res.-Atmos.*, 111, D22106, <https://doi.org/10.1029/2006JD007103>, 2006.
- Moorthi, S., Pan, H.-L., and Caplan, P.: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system, in: Technical Procedures Bulletin 484, US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Office of Meteorology, Program and Plans Division, Silver Spring, MD, 2001.
- Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, *Geophys. Res. Lett.*, 39, L22707, <https://doi.org/10.1029/2012GL053326>, 2012.
- Müller, W. A., Pohlmann, H., Sienz, F., and Smith, D.: Decadal climate predictions for the period 1901–2010 with a coupled climate model, *Geophys. Res. Lett.*, 41, 2100–2107, <https://doi.org/10.1002/2014GL059259>, 2014.
- Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornbluh, L., Li, H., Modali, K., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), *J. Adv. Model. Earth Syst.*, 10, 1383–1413, <https://doi.org/10.1029/2017MS001217>, 2018.
- MunichRe: NatCatSERVICE, available at: <https://natcatservice.munichre.com/> (last access: 23 August 2018), 2017.
- Nissen, K. M., Leckebusch, G. C., Pinto, J. G., Renggli, D., Ulbrich, S., and Ulbrich, U.: Cyclones causing wind storms in the Mediterranean: characteristics, trends and links to large-scale patterns, *Nat. Hazards Earth Syst. Sci.*, 10, 1379–1391, <https://doi.org/10.5194/nhess-10-1379-2010>, 2010.
- Pal, J. S., Giorgi, F., and Bi, X.: Consistency of recent European summer precipitation trends and extremes with future regional climate projections, *Geophys. Res. Lett.*, 31, L13202, <https://doi.org/10.1029/2004GL019836>, 2004.
- Pauling, A. and Paeth, H.: On the variability of return periods of European winter precipitation extremes over the last three centuries, *Clim. Past.*, 3, 65–76, <https://doi.org/10.5194/cp-3-65-2007>, 2007.
- Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P.: An advanced stochastic weather generator for simulating 2-D high-resolution climate variables, *J. Adv. Model. Earth Sy.*, 9, 1595–1627, <https://doi.org/10.1002/2016MS000854>, 2017.
- Peterson, T. C.: Climate change indices, *WMO Bull.*, 54, 83–86, 2005.
- Pfahl, S., O’Gorman, P. A., and Fischer, E. M.: Understanding the regional pattern of projected future changes in extreme precipitation, *Nat. Clim. Change*, 7, 423–427, <https://doi.org/10.1038/nclimate3287>, 2017.
- Pinto, J. G. and Raible, C. C.: Past and recent changes in the North Atlantic oscillation, *WIREs Clim. Change*, 3, 79–90, <https://doi.org/10.1002/wcc.150>, 2012.

- Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Tart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the twentieth century, *J. Climate*, 29, 4083–4097, <https://doi.org/10.1175/JCLI-D-15-0556.1>, 2016.
- Potts, J. M., Folland, C. K., Jolliffe, I. T., and Sexton, D.: Revised LEPS scores for assessing climate model simulations and long-range forecasts, *J. Climate*, 9, 34–53, [https://doi.org/10.1175/1520-0442\(1996\)009<0034:RSFACM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<0034:RSFACM>2.0.CO;2), 1996.
- Primo, C., Kelemen, F. D., Feldmann, H., Akhtar, N., and Ahrens, B.: A regional atmosphere–ocean climate system model (CCLMv5.0clm7-NEMOv3.3-NEMOv3.6) over Europe including three marginal seas: on its stability and performance, *Geosci. Model Dev.*, 12, 5077–5095, <https://doi.org/10.5194/gmd-12-5077-2019>, 2019.
- Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., and Gratzki, A.: A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), *Meteorol. Z.*, 22, 235–256, <https://doi.org/10.1127/0941-2948/2013/0436>, 2013.
- Reyers, M., Feldmann, H., Mieruch, S., Pinto, J. G., Uhlig, M., Ahrens, B., Früh, B., Modali, K., Laube, N., Moemken, J., Müller, W., Schädler, G., and Kottmeier, C.: Development and prospects of the regional MiKlip decadal prediction system over Europe: predictive skill, added value of regionalization, and ensemble size dependency, *Earth Syst. Dynam.*, 10, 171–187, <https://doi.org/10.5194/esd-10-171-2019>, 2019.
- Rhodes, R. I., Shaffrey, L. C., and Gray, S. L.: Can reanalyses represent extreme precipitation over England and Wales?, *Q. J. Roy. Meteorol. Soc.*, 141, 1114–1120, <https://doi.org/10.1002/qj.2418>, 2015.
- Rimbu, N., Boroneanț, C., Buță, C., and Dima, M.: Decadal variability of the Danube river flow in the lower basin and its relation with the North Atlantic Oscillation, *Int. J. Climatol.*, 22, 1169–1179, <https://doi.org/10.1002/joc.788>, 2002.
- Rockel, B., Will, A., and Hense, A.: The regional climate model COSMO-CLM (CCLM), *Meteorol. Z.*, 17, 347–348, <https://doi.org/10.1127/0941-2948/2008/0309>, 2008.
- Schewe, J., Gosling, S. ., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H. K., Seneviratne, S. I., van Vliet, M. T. H., Vautard, R., Wada, Y., Breuer, L., Büchner, M., Carozza, D. A., Chang, J., Coll, M., Deryng, D., de Wit, A., Eddy, T. D., Folberth, C., Frieler, K., Friend, A. D., Gerten, D., Gudmundsson, L., Hanasaki, N., Ito, A., Khabarov, N., Kim, H., Lawrence, P., Morfopoulos, C., Müller, C., Müller Schmied, H., Orth, R., Ostberg, S., Pokhrel, Y., Pugh, T. A. M., Sakurai, G., Satoh, Y., Schmid, E., Stacke, T., Steenbeek, J., Steinkamp, J., Tang, Q., Tian, H., Tittensor, D. P., Volkholz, J., Wang, X., and Warszawski, L.: State-of-the-art global models underestimate impacts from climate extremes, *Nat. Commun.*, 10, 1005, <https://doi.org/10.1038/s41467-019-08745-6>, 2019.
- Schröter, K., Kunz, M., Elmer, F., Mühr, B., and Merz, B.: What made the June 2013 flood in Germany an exceptional event? A hydro-meteorological evaluation, *Hydrol. Earth Syst. Sci.*, 19, 309–327, <https://doi.org/10.5194/hess-19-309-2015>, 2015.
- Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, *J. Am. Stat. Assoc.*, 63, 1379–1389, <https://doi.org/10.1080/01621459.1968.10480934>, 1968.
- Simmons, A. J., Berrisford, P., Dee, D. P., Hersbach, H., Hirahara, S., and Thepaut, J.-N.: A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets, *Q. J. Roy. Meteorol. Soc.*, 143, 101–119, <https://doi.org/10.1002/qj.2949>, 2017.
- Singer, M. B., Michaelides, K., and Hobley, D. E. J.: STORM 1.0: a simple, flexible, and parsimonious stochastic rainfall generator for simulating climate and climate change, *Geosci. Model Dev.*, 11, 3713–3726, <https://doi.org/10.5194/gmd-11-3713-2018>, 2018.
- Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblüeh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M earth system model: ECHAM6, *J. Adv. Model Earth Syst.*, 5, 422–446, <https://doi.org/10.1002/jame.20015>, 2013.
- Stucki, P., Rickli, R., Brönnimann, S., Martius, O., Wanner, H., Grebner, D., and Luterbacher, J.: Weather patterns and hydro-climatological precursors of extreme floods in Switzerland since 1868, *Meteorol. Z.*, 21, 531–550, <https://doi.org/10.1127/0941-2948/2012/368>, 2012.
- Stucki, P., Dierer, S., Welker, C., Gómez-Navarro, J. J., Raible, C. C., Martius, O., and Brönnimann, S.: Evaluation of down-scaled wind speeds and parameterised gusts for recent and historical windstorms in Switzerland, *Tellus A*, 68, 31820, <https://doi.org/10.3402/tellusa.v68.31820>, 2016.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.
- Theil, H.: A rank-invariant method of linear and polynomial regression analysis, in: *Proceedings of the Royal Netherlands Academy of Sciences*, 53, Part I: 386–392, Part II: 521–525, Part III: 1397–1412, https://doi.org/10.1007/978-94-011-2546-8_20, 1950.
- Toreti, A., Koplaki, E., Maraun, D., Kuglitsch, F.-G., Wanner, H., and Luterbacher, J.: Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns, *Nat. Hazards Earth Syst. Sci.*, 10, 1037–1050, <https://doi.org/10.5194/nhess-10-1037-2010>, 2010.
- Torma, C., Giorgi, F., and Coppola, E.: Added value of regional climate modeling over areas characterized by complex terrain – Precipitation over the Alps, *J. Geophys. Res.-Atmos.*, 120, 3957–3972, <https://doi.org/10.1002/2014JD022781>, 2015.
- Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, *B. Am. Meteorol. Soc.*, 84, 1205–1218, 2003.
- Uhlemann, S., Thielen, A. H., and Merz, B.: A consistent set of trans-basin floods in Germany between 1952–2002, *Hydrol. Earth Syst. Sci.*, 14, 1277–1295, <https://doi.org/10.5194/hess-14-1277-2010>, 2010.
- Ulbrich, U., Brücher, T., Fink, A. H., Leckebusch, G. C., Krüger, A., and Pinto, J. G.: The central European floods of August 2002: Part 1 – Rainfall periods and flood development, *Weather*, 58, 371–377, <https://doi.org/10.1256/wea.61.03A>, 2003a.
- Ulbrich, U., Brücher, T., Fink, A. H., Leckebusch, G. C., Krüger, A., and Pinto, J. G.: The central European floods

- of August 2002: Part 2 – Synoptic causes and considerations with respect to climatic change, *Weather*, 58, 434–442, <https://doi.org/10.1256/wea.61.03B>, 2003b.
- van den Besselaar, E. J. M., Haylock, M. R., van der Schrier, G., and Klein Tank, A. M. G.: A European daily high-resolution observational gridded data set of sea level pressure, *J. Geophys. Res.-Atmos.*, 116, D11110, <https://doi.org/10.1029/2010JD015468>, 2011.
- van der Wiel, K., Wanders, N., Selten, F. M., and Bierkens, M. F. P.: Added Value of Large Ensemble Simulations for Assessing Extreme River Discharge in a 2 °C Warmer World, *Geophys. Res. Lett.*, 46, 2093–2102, <https://doi.org/10.1029/2019GL081967>, 2019.
- Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A., Keller, J. D., and Ohlwein, C.: A novel convective-scale regional reanalyses COSMO-REA2: Improving the representation of precipitation, *Meteorol. Z.*, 26, 345–361, <https://doi.org/10.1127/metz/2017/0824>, 2017.
- Ward, M. N. and Folland, C. K.: Prediction of seasonal rainfall in the north nordeste of Brazil using eigenvectors of sea-surface temperature, *Int. J. Climatol.*, 11, 711–743, <https://doi.org/10.1002/joc.3370110703>, 1991.
- Westra, S., Alexander, L. V., and Zwiers, F. W.: Global increasing trends in annual maximum daily precipitation, *J. Climate*, 26, 3904–3918, 2013.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, in: vol. 91 of International Geophysics Series, 2nd Edn., Academic Press, San Diego, California, USA, 2006.
- Yue, S., Pilon, P., Phinney, B., and Cavadias, G.: The influence of autocorrelation on the ability to detect trend in hydrological series, *Hydrol. Process.*, 16, 1807–1829, <https://doi.org/10.1002/hyp.1095>, 2002.
- Zhang, X., Zwiers, F., Li, G., Wan, H., and Cannon, A.: Complexity in estimating past and future extreme short-duration rainfall, *Nat. Geosci.*, 10, 255–259, <https://doi.org/10.1038/NGEO2911>, 2017.
- Zolina, O., Simmer, C., Kapala, A., Bachner, S., Gulev, S., and Maechel, H.: Seasonally dependent changes of precipitation extremes over Germany since 1950 from a very dense observational network, *J. Geophys. Res.-Atmos.*, 113, D06110, <https://doi.org/10.1029/2007JD008393>, 2008.
- Zwiers, F. W., Alexander, L. V., Hegerl, G. C., Knutson, T. R., Kossin, J. P., Naveau, P., Nicholls, N., Schär, C., Seneviratne, S. I., and Zhang, X.: Climate extremes: challenges in estimating and understanding recent changes in the frequency and intensity of extreme climate and weather events, in: *Climate Science for Serving Society*, edited by: Asrar, G. and Hurrell, J., Springer, Dordrecht, 339–389, https://doi.org/10.1007/978-94-007-6692-1_13, 2013.