

Do convection-permitting regional climate models have added value for hydroclimatic simulations? A test case over small and medium-sized catchments in Germany

Oakley Wagner^{1,3}, Verena Maleska², Laurens M. Bouwer^{1,3}

5 ¹Climate Service Center Germany (GERICS), Helmholtz-Zentrum hereon, 20095 Hamburg, Germany

²Chair of Environmental Development and Risk Management, Technical University of Dresden, 01062 Dresden, Germany

³Institute of Geography, University of Hamburg, 20146 Hamburg, Germany

Correspondence to: Oakley Wagner (oakley.wagner@hereon.de)

Abstract. Through fine grid structure and explicit representation of deep convective processes, convection-permitting regional climate models (CPRCMs) bear great potential for improved assessment of climate and hydrology under current and future climatic conditions. For a robust assessment of the added value of CPRCMs as climate service for hydrological impact modelling, the current scope of research needs to be expanded by studies on further model structures and study areas. The paper presented here considers the non-hydrostatic model ICON-CLM 2.6.4 at 3 km resolution (ICON3km) and its driving model ICON-CLM 2.6.4 with parametrised convection at 11 km (ICON11km) for a study area of 13,210 km² in East Central Germany, enclosing the small (107 km²) to medium-sized (529 km²) catchments of the upper and central part of the predominantly rural Weiße Elster river basin. The reanalysis-driven historical hourly air temperature, global radiation, relative humidity, wind speed and precipitation simulations are evaluated. ICON3km is further analysed for added value for discharge, soil moisture and evapotranspiration simulations using the distributed hydrological model WaSiM. Our results suggest primarily an improvement by ICON3km in the estimation of air temperature and summer global radiation, as well as a reduction of the overestimation of the left tail of the frequency distribution of wind speed. The most noticeable deficiencies of ICON3km are the strong overestimation of high precipitation intensity and too frequent heavy rainfall events. These shortcomings translate into a pronounced overestimation of discharge when uncorrected, dominating the hydrological impact estimations. As such, no added value in the use of ICON3km for hydrological impact modelling in the Weiße Elster basin was identified.

25 1 Introduction

Extreme rainfall events around the world are expected to become more intense and more frequent under global warming (IPCC, 2021; Myhre et al. 2019). In the UK for example, the frequency of rainfall events exceeding 20 mm/h, and thereby capable of producing flash floods, is estimated to be 4-fold by 2070 under a high emission scenario (Kendon et al., 2023). With urbanisation on the rise worldwide (Jiang & O'Neill, 2017) and a growth in population in flood-prone areas (Kam et

30 al., 2021), an increase in the frequency and intensity of local extreme rainfall events leads to large flood risks. Convection-permitting regional climate models (CPRCMs) bear great potential for improved projections of changes in extreme rainfall. These models operate on spatial resolutions smaller than 4 km (Prein et al., 2021) and are built with a non-hydrostatic dynamical core, since the hydrostatic approximation is no longer valid at this scale (Giorgi, 2019; Steppeler et al., 2003). As such, they resolve deep convection and no longer rely on its parametrisation, thereby circumventing a major source of
35 uncertainty (Prein et al., 2015). However, finer scale processes, such as cloud and precipitation microphysics, shallow convection, radiative transfer, and turbulence still need to be parametrised (Prein et al., 2015; Kendon et al., 2021). A full depiction of all energetic processes in clouds would require a downscaling to the smallest scale, i.e. the Kolmogorov scale, where viscous dissipation converts turbulent kinetic energy into heat (Prein et al., 2015; Friedlander and Topper, 1961 cited in Schulz & Sanderson, 2004). Besides a more process-based representation of atmospheric processes, CPRCMs also
40 introduce finer grid spacing, allowing to better capture complex topography and land surface heterogeneities (Gutowski et al., 2020), constituting an added value for impact modelling over small catchments of mountainous or highly urban character (e.g. Schaller et al., 2020; Tamm et al., 2023).

To justify their high computational costs and data storage requirements (Schär et al., 2020), CPRCMs must prove to indeed offer substantial improvement compared to coarser regional climate models with parametrised convection (RCMs).
45 However, in the field of hydrological impact assessment, to this day only a limited number of studies comparing the use of CPRCMs to RCMs have been conducted. An overview is given in Table 1, reflecting how from a state of unsatisfactory performance for impact modelling in the early stages of development (e.g. Kay et al., 2015; Reszler et al., 2018), CPRCMs have made remarkable progress, with a number of recent studies showing substantial added value of CPRCMs for flood simulation. However, with most studies only focusing on a single basin, using data from a single CPRCM, over a short
50 simulation time, there remains an urgent need for further studies (Lucas-Picher et al., 2021). Our paper introduces an additional geographical region and CPRCM to the discourse. It looks at the strengths and limitations of ICON-CLM (2.6.4) in convection-permitting setup at 3 km resolution in depicting hourly near-surface air temperature and relative humidity, wind speed at 10 m, global radiation and precipitation over a chosen catchment in East Central Germany, compared to its driving coarser RCM, ICON-CLM (2.6.4) at 11 km resolution. The climate model data is further used as input to the
55 distributed hydrological model WaSiM to identify potential added value of the studied CPRCM data for discharge simulations in the catchment, while underlining the results by analyses of soil moisture and evapotranspiration estimations.

Table 1: Previous studies investigating the added value of CPRCMs compared to RCMs, with non-bias-corrected (*) and bias-corrected (+) precipitation data, for hydrological impact modelling (exclusively model chains consisting of a climate or numerical weather prediction model and a hydrological model are considered)

Study	Study Region	Driving GCM/ RCM	CPRCM	Hydrological Model	Key Findings
Kay et al. (2015)	southern Britain	limited-area version of HadGEM3 (12 km)	modified version of the UK Met Office weather forecast model (resolution: 1.5 km) *	CLASSIC-GB (semi-distributed)	- Worse performance for hydrological simulations using CPRCM data due to its strong positive bias in precipitation intensity
Mendoza et al. (2016)	Colorado River basin, south-western USA	WRF (12 km, 36 km)	WRF (4 km) *	PRMS (distributed), VIC (semi-distributed), Noah LSM (semi-distributed), Noah-MP LSM (semi-distributed)	- Improvement in the simulation of historical runoff ratios by using CPRCM data
Reszler et al. (2018)	south-eastern Austria	COSMO-CLM, WRF (12.5 km, 50 km)	COSMO-CLM, WRF (3 km) *, + (scaled distribution mapping)	KAMPUS (distributed)	- CPRCM data was not found to offer added value for hydrological modelling
Schaller et al. (2020)	West Coast of Norway	EC-Earth v2.3 (25 km)	AROME-MetCoOp (2.5 km) *	Hydrologiska Byråns Vattenbalans (semi-distributed and distributed)	- spatial refinement and higher temporal resolution advantageous in the modelling of fast flood-generating processes in complex terrain
Rudd et al. (2020)	southern Britain	Met Office Hadley Centre model (12 km)	Met Office Hadley Centre model (1.5 km) *	Grid-to-Grid (distributed)	- finer grid spacing advantageous in exceedance count projections, especially for minimal and minor property severity impact level
Kay (2022)	Britain	HadREM3-GA705 (12 km)	HadREM3-RA11M (2.2 km) *, + (monthly factors)	G2G (distributed)	- added value by CPRCM in respect to percentage bias in low flow volume, median flow and high flow volume for baseline simulations - divergences between use of CPRCM and RCM data mostly abolished with bias-correction
Davis et al. (2022)	Bangalore catchment, South India	Global Forecasting System (25 km);	WRF (3km, 1 km; 15-min intervals) *	PCSWMM (semi-distributed)	- improvement in the simulation of urban flood level peaks using CPRCM data

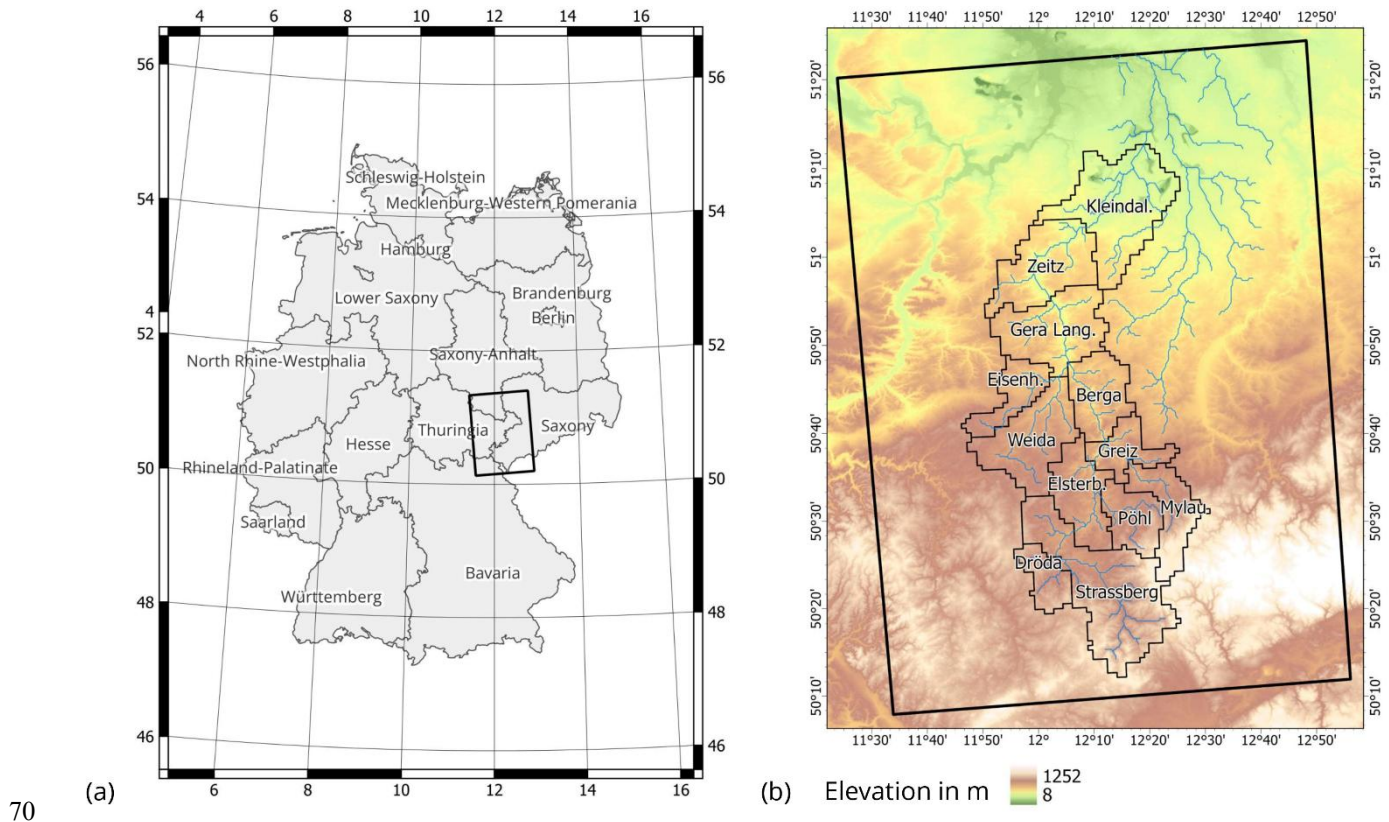
		3-hour intervals)			
Tamm et al. (2023)	Southern Finland	HARMONIE (12 km)	HARMONIE (3 km) + (empirical quantile mapping)	SWMM (distributed)	- more realistic annual maximum flow values using CPRCM data
Ascott et al. (2023)	Lake Victoria basin, Eastern Africa	Met Office Unified Model (25 km)	Met Office Unified Model (4 km) *	GR4J (lumped)	- higher flow estimates using CPRCM data
Poncet et al. (2024)	Gardon d'Anduze catchment, southern France	CNRM-ALADIN (12 km)	CNRM-AROME (2.5 km) *, + (CDF-t method)	GR5H (lumped), CREST (distributed)	- improved simulation of flood intensity and frequency using the uncorrected CPRCM data - most intense floods best modelled with bias-corrected CPRCM data compared to RCM data, but RCM data served for better depiction of the distribution of flow peaks over threshold
Archer et al. (2024)	Bristol and Bath, UK	HadREM3-GA705 (12 km)	HadREM3-RA11M (2.2 km) *	LISFLOOD-FP (distributed)	- Intensity-duration-frequencies from a CPRCM ensemble produce larger relative flooded areas compared to the intensity-duration-frequencies from interpolated rainfall observations
Xie et al. (2025)	Western Norway	HCLIM38-ALADIN (12 km)	HCLIM38-AROME (3 km) *	HBV (lumped), WRF-Hydro (distributed)	- CPRCM captures spatial variability of annual maximum 1-day and 1-hour precipitation better - no universal added value for flood simulation by CPRCMs, but improvement in the simulation of severe flood peaks

In this paper, section 2 outlines the study area considered in this paper, the observational and climate model data used, as well as the applied methodology. The results from the meteorological and hydrological data analysis are presented in section 3, while section 4 provides a discussion. A conclusion is given in section 5.

2 Methods and Data

65 2.1 Study Area

The study area was chosen to encompass the focus regions of the KlimaKonform project (Zorn et al., 2024) in East Central Germany (Fig. 1a). It spans 13,210 km² and hosts the upper and central part of the Weiße Elster river basin (Fig. 1b). The twelve enclosed catchments stretch over an area of 2,960 km², with the smallest having an area of 107 km² (Dröda) and the largest of 529 km² (Strassberg).



70 (a) Location of the study area in Germany and (b) the studied catchments within the area over a digital elevation model of 30 m resolution (SRTM 1 Arc-Second Global; USGS, 2018)

The study area is embedded between the Leipzig lowlands to the North and the Elster Mountains and Ore Mountains to the South, resulting in a topographically varied landscape encompassing both flat plains and steeply sloped terrain (Rode et al., 75 2008; Miera et al., 2022). Altitudes range from 8 m to 1,252 m on a 30 m grid, from 85 m to 1,006 m on a 3 km grid, and from 101 m to 852 m on an 11 km grid.

The northern part of the study area falls under the Köppen climate classification of Cfb (temperate, no dry season, warm summer), and the southern part is classified as Dfb (cold, no dry season, warm summer). In the study region, floods occur primarily in summer as a result of high-intensity convective storms, as well as in winter due to prolonged rainfall, sometimes amplified by rapid snowmelt.

The land use in the Weiße Elster river basin is predominantly cropland and pasture (Rode et al., 2008, see Fig. S1 in the appendix). Chernozem/phaeozem and luvisol soils on loess are prevalent in the northern part of the study area, while cambisols and stagnic gleysols dominate the southern part (Miera et al., 2022).

2.2 Observational Data

Hourly observational data from 20 stations for mean air temperature, two stations for global radiation, 20 stations for relative humidity at 2 m, and 11 stations for wind speed at 10 m has been obtained from the German Weather Service (DWD) for the period from 2005 to 2014 (Fig. S2 in the appendix). No spatial interpolation has been performed; the station measurements are compared point-based to the corresponding climate model grid cells. For the assessment of precipitation anomalies, 60-year time series of monthly point-based precipitation data (1963 - 2022) have been obtained from DWD for five stations in the study area.

For other evaluations of precipitation simulations and hydrological modelling, adjusted radar-based quantitative precipitation estimations are used as reference for the study period from 2005 to 2014 in hourly time steps at a resolution of 1 km. The data set constitutes a gap-filled version of DWD's RADOLAN product (Weigl & Winterrath, 2010) and is provided through the ReKIS portal (Körner, 2022) for the federal states of Saxony, Saxony-Anhalt and Thuringia. RADOLAN is a radar-based precipitation dataset at 1 km resolution, which has been adjusted with measurements from DWD's ground station network (Winterrath et al., 2012). The choice of adjustment procedure is determined based on a performance test with a few control stations (Winterrath et al., 2012). It is unknown whether any, respectively which of these control stations correspond precisely to the stations used in this study. However, the station values can generally be assumed to be preserved.

The data has been upscaled and regridded to the climate model grids using the area-weighted average method (becoming RADOLAN3km and RADOLAN11km). Kreklow et al. (2020) evaluated the performance of the original RADOLAN dataset against rain gauge data and found RADOLAN to overestimate the number of days with daily precipitation sums > 1 mm, as well as heavy rainfall days (i.e. daily precipitation sums > 20 mm), in the summer half-year. Especially up to the year 2010, extreme high outliers in the time series were not uncommon, with markable improvements through the introduction of additional gauges in 2007, as well as further quality checks and new processing routines in 2010 (Kreklow et al., 2020). However, the average daily precipitation sum for the summer half-year was underestimated, with a similar trend reported for the winter half-year, yet far less pronounced (Kreklow et al., 2020). The mean median of the absolute daily deviations, when

comparing RADOLAN data to daily point-based measurements of precipitation for the period from 2013 to 2016, was quantified as 0.761 mm/day (DWD, 2021).

110 The hydrological model was calibrated and validated using Thiessen-interpolated point-based meteorological observations, including hourly precipitation from 20 precipitation gauges in the catchment (see Fig. S2 in the appendix for a map of the observational network). An evaluation of how the RADOLAN product compares to the Thiessen-interpolated point-based precipitation measurements over the study area is outlined in Fig. S3 in the appendix by means of a QQ-plot.

Discharge data along the Weiße Elster has been obtained through the ReKIS project (Kronenberg et al., 2021), based on measurements by the Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG), Thüringer Landesamt für 115 Umwelt, Bergbau und Naturschutz (TLUBN) and Landesamt für Umweltschutz Sachsen-Anhalt (LAU).

Topographic information was obtained from SRTM 1 Arc-Second Global (USGS, 2018), land use is based on Corine Land Cover 2018 (Copernicus Land Monitoring Service, 2020) and used soil data refers to BÜK200 (BGR & SGD, 2018).

2.3 Climate Model Data

The regional climate model ICON-CLM (version 2.6.4, van Pham et al., 2021) has been run over the EURO-CORDEX 120 domain at 11 km resolution for hourly output (ICON11km). It is driven three-hourly with data from the global atmospheric reanalysis ECMWF-ERA5 and provides the lateral boundary conditions for the CPRCM. As CPRCM, ICON-CLM (version 2.6.4) is employed in its convection-permitting setup at 3 km resolution over the Central European (CEU) domain providing hourly weather data (ICON3km). The domains are illustrated in Fig. S4 in the appendix. Information on the model dynamics is provided by Zängl et al. (2015).

125 In ICON11km, deep, mid-level, and shallow convection have been parametrised using the Tietke/Bechtold scheme (Tiedtke, 1989; Bechtold et al., 2008). ICON3km, in turn, operates on a resolution fine enough to explicitly resolve deep and mid-level convection, with only the shallow convection parametrisation needing to remain switched on. Vertical turbulent transfer is parametrised according to Raschendorfer (2008), and cloud microphysics are calculated with the single-moment scheme of Seifert (2008).

130 Tracer transport and fast-physics parametrisations are computed at time steps of 25 s, while the dynamical core runs at 5 s.

2.4 Evaluation of Simulated Meteorological Data

Air temperature, global radiation, relative humidity, and wind speed station measurements within the catchment bounding box (see Fig. S2 in the appendix) are compared to their co-located climate model grid cells. Air temperature estimates by the climate models are projected to the elevation of the measurement stations using the environmental lapse rate, which is of

135 0.65 °C/ 100 m. Precipitation simulations for the area are evaluated raster-based against radar precipitation estimations
upscaled and regridded to the climate models' grids. Comparing precipitation simulations to observations on the same grid
circumvents the need to consider differences in the degree of smoothing of local heavy rainfall fields with resolution (Iles et
al., 2020). While the radar precipitation estimations are upscaled to the resolution of the climate models for comparison
between the climate model simulations and the observations, however the climate models themselves are not compared
140 among each other by scaling the simulations to a common grid spacing. In fact, upscaling the finer climate simulation results
to the grid of the coarser climate model would entail partial information loss (Iles et al., 2020), while downscaling the
coarser climate model simulations would bring them into resolutions where processes dominate which are on a scale too fine
to have been computed by the coarser model (Prein et al., 2016). It is however worth noting that first studies suggest that
improvements at higher resolution may still be apparent after upscaling (e.g. Prein et al., 2016; Fantini et al., 2018 for
145 upscaling from ~ 12 km to ~ 50 km resolution).

When temporally aggregating observational data, time periods are not summarised if more than 20% of their data is missing
(Haylock et al., 2008). Smaller gaps are filled by dividing the totals through the proportion of non-missing observations
(Haylock et al., 2008). Frequency analyses have been conducted based on frequency polygons with a number of bins
according to Sturges' rule (Sturges, 1926). It should be noted that for large datasets Sturges' rule is prone to oversmoothing
150 (Scott, 2009). For frequency analyses on the 99.5th-percentile of hourly precipitation, the quantile thresholds are deduced
from RADOLAN data for both resolutions and transferred to the respective ICON data sets.

The meteorological time series are of only 10 years and thereby only a third of the length recommended for climatological
analyses by the World Meteorological Organization (WMO, 2017). The short time series can only hint to strengths and
limitations of the climate models, but longer time series are needed to draw firm conclusions on the hydroclimatic model
155 performance.

Depth-duration-frequency curves have been constructed by sliding a moving window of a fixed duration (2 h, 4 h to 24 h by
2-hour-increments for 11-km-resolution data, and by 4-hour-increments for 3-km-resolution data) over the time series and
fitting the identified annual-maximum precipitation sums with a Gumbel distribution (Koutsoyiannis et al., 1998). The short
duration of the time series of only 10 years poses a strong limitation to this approach.

160 2.5 Hydrological Modelling

The distributed, physically based deterministic hydrological model WaSiM (Message Passing Interface, model version:
Richards 10.00.03; see Schulla, 2021) has been run for hourly time steps for the Weiße Elster river for the study period from
2005 to 2016. The first simulation year has been discarded to ensure adequate spin-up. Meteorological input variables consist
of hourly air temperature, global radiation, relative humidity, wind speed and precipitation. The modules in WaSiM for high-
165 resolution hydrological simulation in the temperate zone are for processes of evapotranspiration, snow accumulation and

snow melt, interception of snow and precipitation, as well as an unsaturated-zone model and a groundwater model. Potential evapotranspiration is calculated after Penman-Monteith, thereby showing dependency on the meteorological input variables of air temperature, global radiation, relative humidity and wind speed. Interception processes are described by an energy balance approach, considering the complete suite of meteorological input variables, just as is the case for the simulation of snow accumulation and snow melt. The unsaturated zone model is built on the Richards-equation and draws on the aforementioned modules, as well as on the groundwater model. The hydrological model was calibrated on (1) the single reservoir recession constant for surface runoff, (2) the single reservoir recession constant for interflow, (3) the drainage density for interflow and (4) the fraction of surface runoff on snow melt.

Topography, land use/ land cover (LULC) and soil characteristics have been embedded on a 1 km² grid for all simulation runs. The vertical water movement through the unsaturated zone was described by the Richards equation (Richards, 1931) and potential evapotranspiration was calculated according to the approach after Penman-Monteith (Monteith, 1975; Brutsaert, 1982). As interpolation technique for the meteorological input data, Thiessen polygons were chosen in order to keep the grid structure of the climate model data unchanged. Radiation and temperature were adjusted considering the impact of topography, according to the implemented scheme by Oke (1987).

Alongside discharge, WaSiM outputs a range of other hydroclimatic variables. Among them are catchment-average total evapotranspiration and catchment-average relative soil moisture simulated for the upper soil section reaching from the soil surface to the root depth. Evapotranspiration and soil moistures estimates are studied in this paper. While snowmelt simulation data is outputted as well, it is not analysed further. CPRCMs are not expected to offer their main added value in winter, when large synoptic weather systems dominate, as these are already well represented by the driving RCMs (Strandberg & Lind, 2020). The focus of the paper presented here is therefore on summer storms, where CPRCMs and RCMs are likely to show their greatest differences (Strandberg & Lind, 2020) as the CPRCMs no longer rely on the parametrisation of deep convection (Ban et al., 2021).

Through the complex interactions of the hydroclimatic variables, the flood generation is likely to show a non-linear behaviour. In a scatter plot of rainfall to runoff volume, the presence of a breakpoint in the linear regression reflects such non-linear behaviour (Buttle et al., 2018). Ross et al. (2021) found the existence of a breakpoint to be most common when considering the relation of runoff volume (Q_{tot}) to the sum of rainfall volume (T_{tot}) and antecedent rainfall (AT). This relation is therefore used in the study presented here to analyse the effect the use of a CPRCM has on the non-linearity of flood generation compared to the use of a coarser regional climate model with parametrised convection. The temporal delineation of runoff events over a specified threshold of 0.03 mm/h at the most downstream catchment (Kleindalzig) and the association of the triggering rainfall events was performed using a modified version of the HydRun tool (Tang & Carey, 2017; with modifications to the runoff peak search algorithm). The months of November to May were discarded to avoid the effect of snowmelt.

Boxplots of computed hourly discharge for a set of catchments are presented in the paper. Boxplots are built according to McGill et al. (1978), i.e. they represent the median with the 25th and 75th percentile, while whiskers extend to the largest/smallest value but no more than 1.5-times the interquartile range.

3 Results

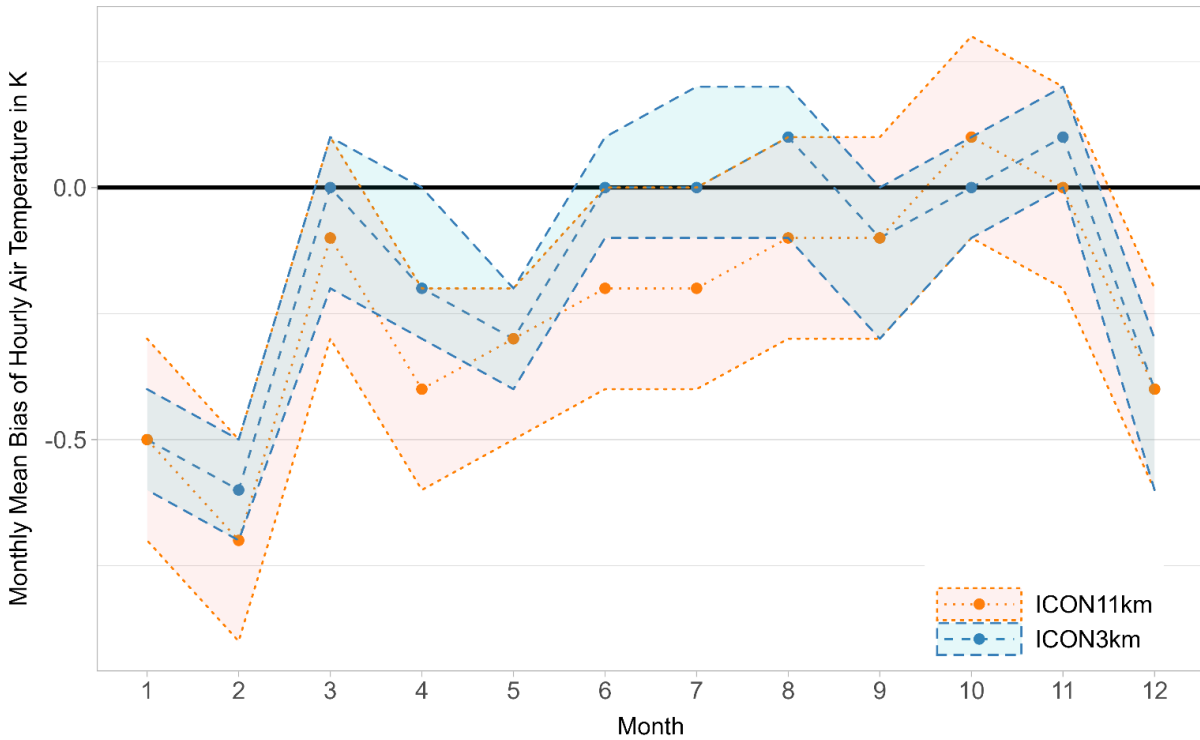
3.1 Evaluation of Simulated Meteorological Data

3.1.1 Near-Surface Air Temperature

According to a Welch two sample t-test, both the RCM and the CPRCM compute significantly lower hourly air temperatures than have been observed on average over the study period ($\bar{t}_{Obs} = 8.7$ °C, $\bar{t}_{ICON3km} = 8.5$ °C, $\bar{t}_{ICON11km} = 8.4$ °C) at 95% confidence level. However, these cold mean biases fall within the measurement uncertainty bandwidth of the temperature stations of ± 0.08 K to ± 0.76 K (Brinckmann & Dirksen, 2020). As such, no statement of added value by the CPRCM can be made for mean hourly air temperature simulation. It is however noticeable that the temperature estimates by the RCM are statistically significantly different from those by the CPRCM.

In the frequency distribution of hourly air temperature (see Fig. S5 in the appendix) and in the representation of its diurnal cycle for winter, spring and summer (see Fig. S6 in the appendix), ICON3km is not found to offer noticeable improvement compared to ICON11km. A better representation of the diurnal cycle of air temperature by ICON3km is however apparent for autumn (see Fig. S6 in the appendix).

Monthly mean biases of hourly air temperature of ICON11km and ICON3km fall within the uncertainty bandwidth of the observations (Fig. 2). Nevertheless, the results hint towards two points of improvement by the CPRCM. Firstly, while ICON11km only achieved zero monthly mean bias of hourly air temperature for the month of November, ICON3km did so for March, June, July and October. Consistently throughout the summer (JJA), and in April, ICON3km offered the greatest improvements compared to ICON11km, with a bias reduction by 0.2 K. Secondly, throughout the year the spatial bias variability from ICON3km is on average 38% lower than from ICON11km, as reflected by the 95% confidence interval from the station means. Monthly mean biases of hourly air temperature were negative or zero throughout most of the year, and greater in winter (DJF) than in summer (JJA) for both ICON3km and ICON11km.



225 **Fig. 2: Monthly mean bias of hourly air temperature of ICON3km and ICON11km for the period from 2005 to 2014, as well as the 95%-confidence intervals from the station means**

3.1.2 Global Radiation

Both ICON3km and ICON11km compute daily mean global radiation significantly higher than observed according to a Welch two sample t-test ($\bar{G}_{ICON3km} = 45.9 \text{ J/cm}^2$, $\bar{G}_{ICON11km} = 45.9 \text{ J/cm}^2$, $\bar{G}_{obs} = 44.5 \text{ J/cm}^2$). However, if it is assumed that the observations are systematically too low by 3%, which is also the pyranometers' measurement uncertainty (DWD, n.d.), the apparent overestimation by the climate models is not significant any more. Indeed, pyranometers face the challenge of temperature differences between their inner dome and detector, caused by infrared radiation from the instrument optics in conjunction with different thermal capacities and connectivities (Philipona, 2002), and further intensified by differential cooling by the colder sky, especially on cloud-free days (Sanchez et al., 2015). These effects lead to thermal negative offsets if not controlled sufficiently well, and hence to an underestimation of global radiation by the pyranometers (Philipona, 2002; Sanchez et al., 2015). Daily global radiation estimates by ICON3km were also not found to be significantly different from those computed by ICON11km.

230

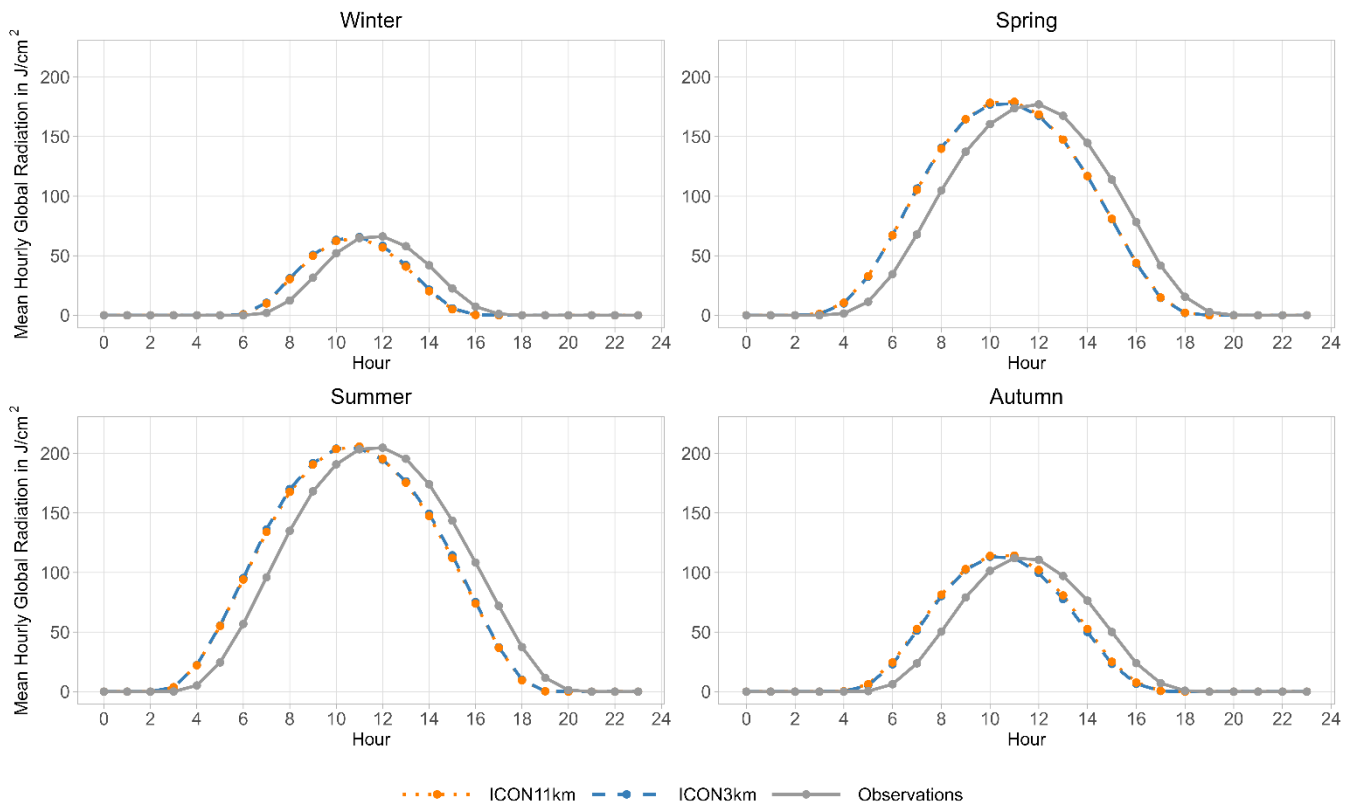
235

Both climate models were found to overestimate the frequency of daily mean global radiation for the range of bins between 45.7 and 106.7 J/cm^2 but underestimate the frequency of higher global radiation (see Fig. S7 in the appendix). A similar

240 picture can be drawn for the intensity estimates of global radiation. The climate models show a positive bias in the intensity of moderate daily mean global radiation. Extremes however were underestimated.

For most months of the year, the studied CPRCM does not offer noticeable improvement in the estimation of daily mean global radiation, apart during the peak of summer (July), where ICON3km was found to reduce the negative monthly mean bias of daily mean global radiation by 2.5 J/cm^2 (see Fig. S8 in the appendix). For reference, average daily mean global radiation in July was of 77.4 J/cm^2 .

245 When looking at the diurnal cycle of global radiation, estimates from the climate models peak one hour earlier than the observations, however they are able to adequately represent the amplitude of the diurnal cycle (Fig. 3). No apparent improvement is offered by ICON3km compared to ICON11km.



250 **Fig. 3: Seasonal diurnal cycles of mean hourly global radiation computed by ICON11km and ICON3km, as well as the observations for the study area for the period from 2005 to 2014**

3.1.3 Relative Humidity

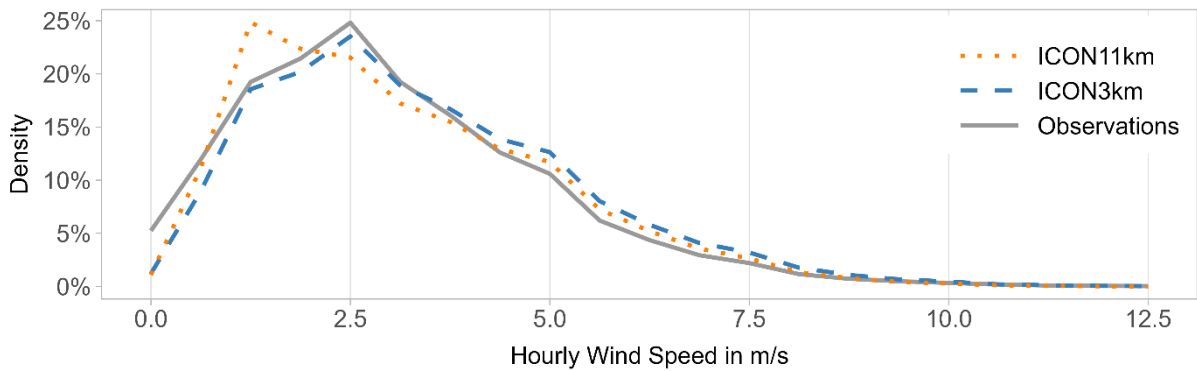
While hourly relative humidity computed by ICON3km is significantly higher than observed ($\overline{RH}_{ICON3km} = 78.7\%$ vs $\overline{RH}_{Obs} = 78.3\%$), this is not the case for the estimates by ICON11km according to a Welch two sample t-test (with $\overline{RH}_{ICON11km} = 78.3\%$). Under consideration of the measurement uncertainty of $\pm 2\%$ of the employed sensor HMP45D (Kyrouac & Theisen, 2017), however neither of the climate models show significant differences to the observations. The estimates by the climate models were found to be statistically significantly different from each other.

ICON11km tends to underestimate the frequency bins of relative humidity of below 46.0% and of above 89.8%, while overestimating those between 59.1% and 89.8% (see Fig. S9 in the appendix). ICON3km does not seem to offer noticeable improvement in the frequency distribution.

Overall, for most months of the year, monthly mean bias of hourly relative humidity is lower with ICON3km than with ICON11km (see Fig. S10 in the appendix). Only in the months of spring (MAM) was ICON11km found to outperform ICON3km. However, it should be noted that the largest absolute difference in the monthly means of hourly errors between the two models occurs in September and October and is of only 1.3% relative humidity. These deviations fall within the measurement uncertainty bandwidth. Both climate models tend to show negative relative humidity biases during winter (DJF) and autumn (SON), and positive biases during spring (MAM) and summer (JJA).

3.1.4 Wind Speed

According to a Welch two sample t-test, the climate models show significantly higher mean hourly wind speeds than have been observed ($\overline{v}_{ICON3km} = 3.4$ m/s, $\overline{v}_{ICON11km} = 3.2$ m/s, $\overline{v}_{Obs} = 3.1$ m/s). If the observations are assumed to be systematically too low in the order of the measurement uncertainty, the estimated mean from ICON11km is not significantly different from that of the observations. The measurement uncertainty of the employed anemometers (Ultrasonic Anemometer 2D) is of ± 0.1 m/s for wind speed below 5 m/s, and of $\pm 1.5\%$ for wind speed higher than that (METEK, n.d.). Computed wind speeds by ICON3km remain significantly too high in the mean even under consideration of the measurement uncertainty. While this might suggest a performance decline by the CPRCM, it was however found to be the result of an improved representation of the frequency distribution (Fig. 4). Both ICON3km and ICON11km tend to overestimate the frequency of moderate to high wind speeds (2.8 to 11.9 m/s, resp. 4.0 to 8.5 m/s), but ICON11km also does so for very low wind speeds (0.6 to 1.7 m/s), thereby attenuating its positive bias of the mean. While the climate models overestimate the frequency of high wind speeds, they underestimate their intensity, with slight improvements by ICON3km (see Fig. S11 in the appendix).



280 **Fig. 4: Frequency polygons for hourly wind speed as observed, and calculated by ICON3km and ICON11km over the study area for the period from 2005 to 2014**

ICON11km was found to compute a too pronounced amplitude of the diurnal cycle of hourly wind speed compared to the observations (Fig. 5). ICON3km shows a better performance yet tends to overestimate absolute values. While the observations plateau during the early afternoon, mean wind speed computed by ICON11km drops sharply after having
 285 reached its peak around noon, together with solar radiation. Mean wind speed simulations by ICON3km do not show a peak as sharp, but level out shortly, and are thereby in better alignment with the observations.

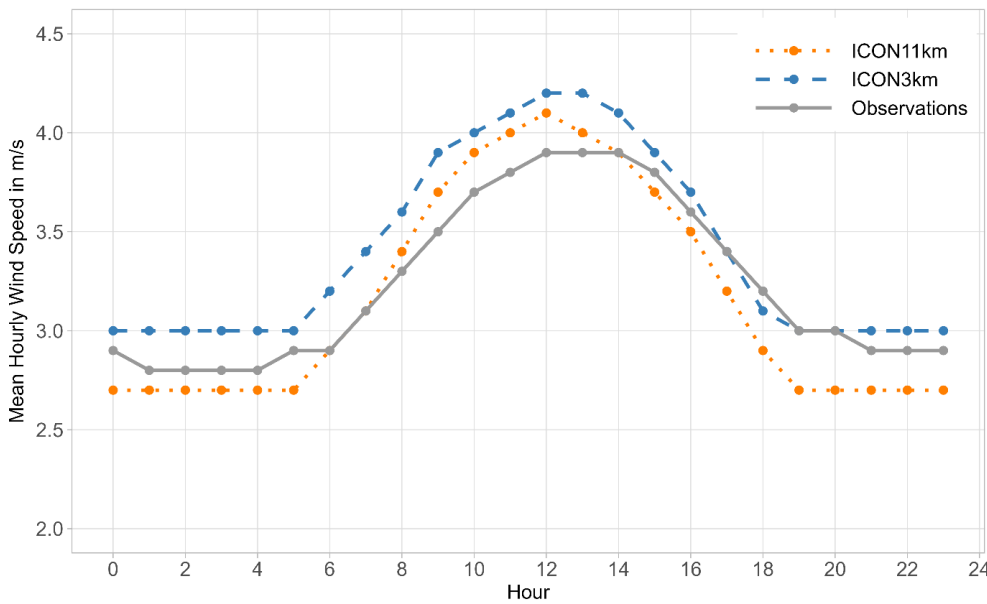
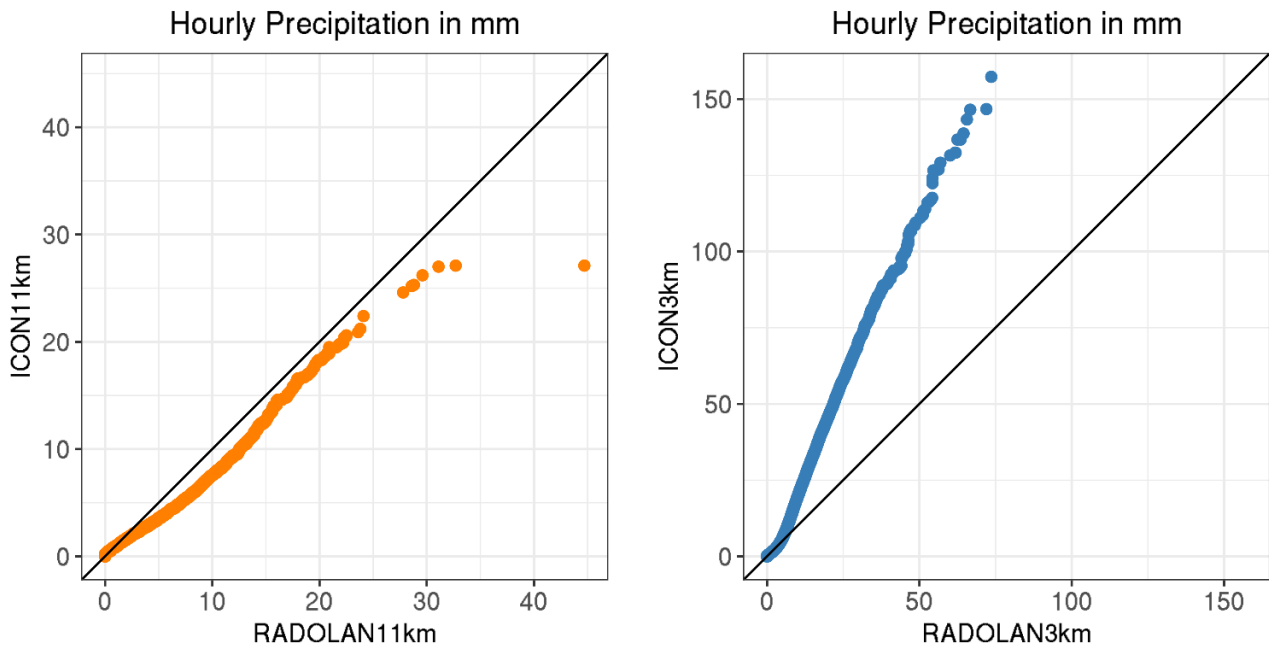


Fig. 5: Diurnal cycle of mean hourly wind speed computed by ICON11km and ICON3km, as well as the observations for the study area for the period from 2005 to 2014

290 **3.1.5 Precipitation**

With median hourly precipitation of 2.8 mm over the 99.5th percentile (left-closed interval), ICON11km shows an underestimation compared to adjusted radar data upscaled to the same grid (RADOLAN11km) with a respective estimate of

3.8 mm. In turn, ICON3km overestimates the median of the 99.5th percentile of hourly precipitation compared to upscaled adjusted radar estimates (RADOLAN3km), with 4.2 mm to 4.0 mm respectively. These trends are further visible in the QQ-plots (Fig. 6). ICON3km computes unrealistically high hourly rainfall intensities of up to 157.3 mm/h over the study region, which is around twice as high as the maximum seen in RADOLAN3km (73.6 mm/h). Looking exemplarily at all ICON3km hourly precipitation estimates surpassing a chosen threshold of 110% of the maximum hourly precipitation registered in RADOLAN3km (a total of 146 estimates), these extreme values were found to cover up to 8 cells per time step (see histogram of the number of covered cells per time step in Fig. S12 in the appendix). These extreme estimates all occurred between late April and Mid-September, with a concentration over the summer months. No geographic region in the study area was the main recipient of these storms. An example of such an unrealistic heavy precipitation event can be seen in the simulation time series in July of 2009, when ICON3km computed an area of extreme precipitation with local hourly rainfall intensities of up to 143.3 mm/h (peak on 2009-07-17 17:00 with 6 adjacent cells showing rainfall intensities above 100 mm/h). In Fig. S13 in the appendix, maps of the precipitation sums for the time window ± 12 h around the discussed precipitation peak time in ICON3km are shown for ICON3km, ICON11km, RADOLAN3km and RADOLAN11km. The areal means of the precipitation sums over the study area are of 47.9 mm in ICON3km, 18.6 mm in RADOLAN3km, 18.8 mm in ICON11km and 18.4 mm in RADOLAN11km. This specific event has been chosen for closer hydrological investigation, as the combination of extent and intensity characterises it as one of the most severe downpours simulated. In 3.2.5 *Event-based Simulation*, the hydrological impact of this simulated heavy precipitation event is discussed.



310

Fig. 6: QQ-plots for hourly precipitation for the period from 2005 to 2014 for ICON11km to RADOLAN11km (left) and ICON3km to RADOLAN3km (right)

The frequency polygons for the 99.5th-percentiles of hourly precipitation, with the thresholds drawn from the RADOLAN data for both resolutions (2.8 mm/h from RADOLAN3km and 2.7 mm/h from RADOLAN11km), are shown in Fig. 7.

315 ICON11km computes too frequent low intensity rainfall, while ICON3km gives an underestimation. The frequency of moderate and heavy rainfall in turn is underestimated by ICON11km but overestimated by ICON3km.

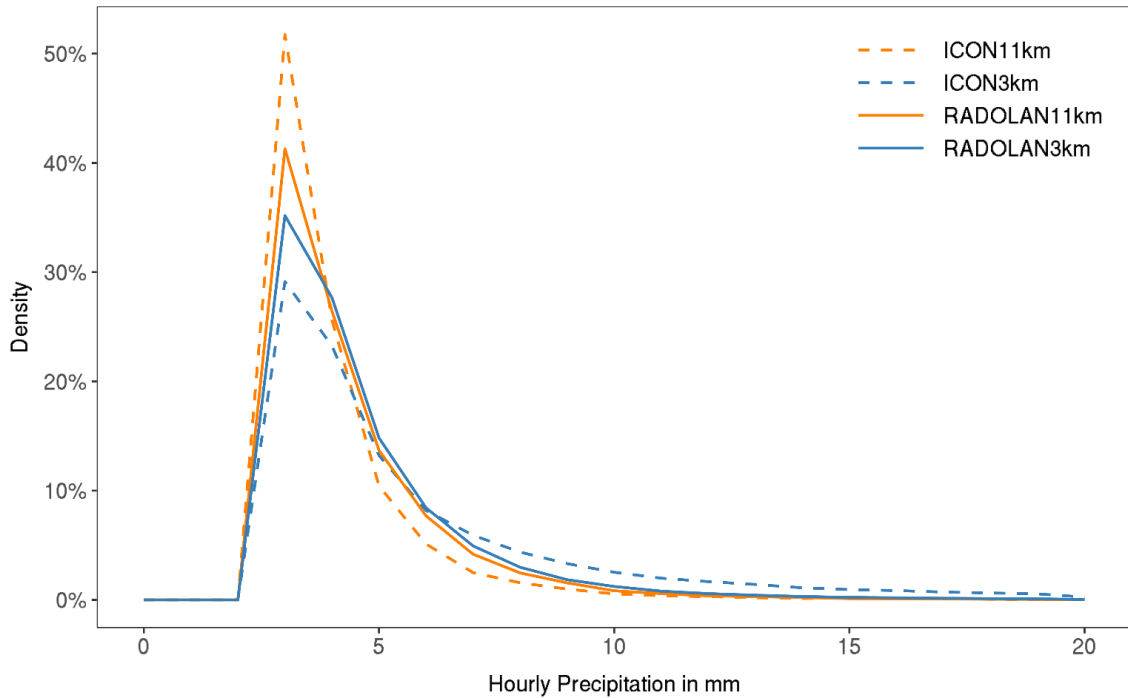


Fig. 7: Frequency polygons for ICON11km, ICON3km, RADOLAN11km and RADOLAN3km for the 99.5th-percentile of hourly precipitation for the period from 2005 to 2014

320 The overestimation of high precipitation intensities and their frequencies by ICON3km further results in too high depth-
duration-frequency curves, as shown for selected return periods in Fig. 8. In keeping is an overestimation of the mean yearly
precipitation totals by ICON3km compared to the radar data upscaled to equal resolution (790 mm compared to 670 mm).
An overestimation of the frequency of light precipitation by ICON11km in turn also results in too high mean yearly
precipitation totals by ICON11km compared to adjusted radar data estimates of respective resolution (729 mm compared to
325 665 mm).

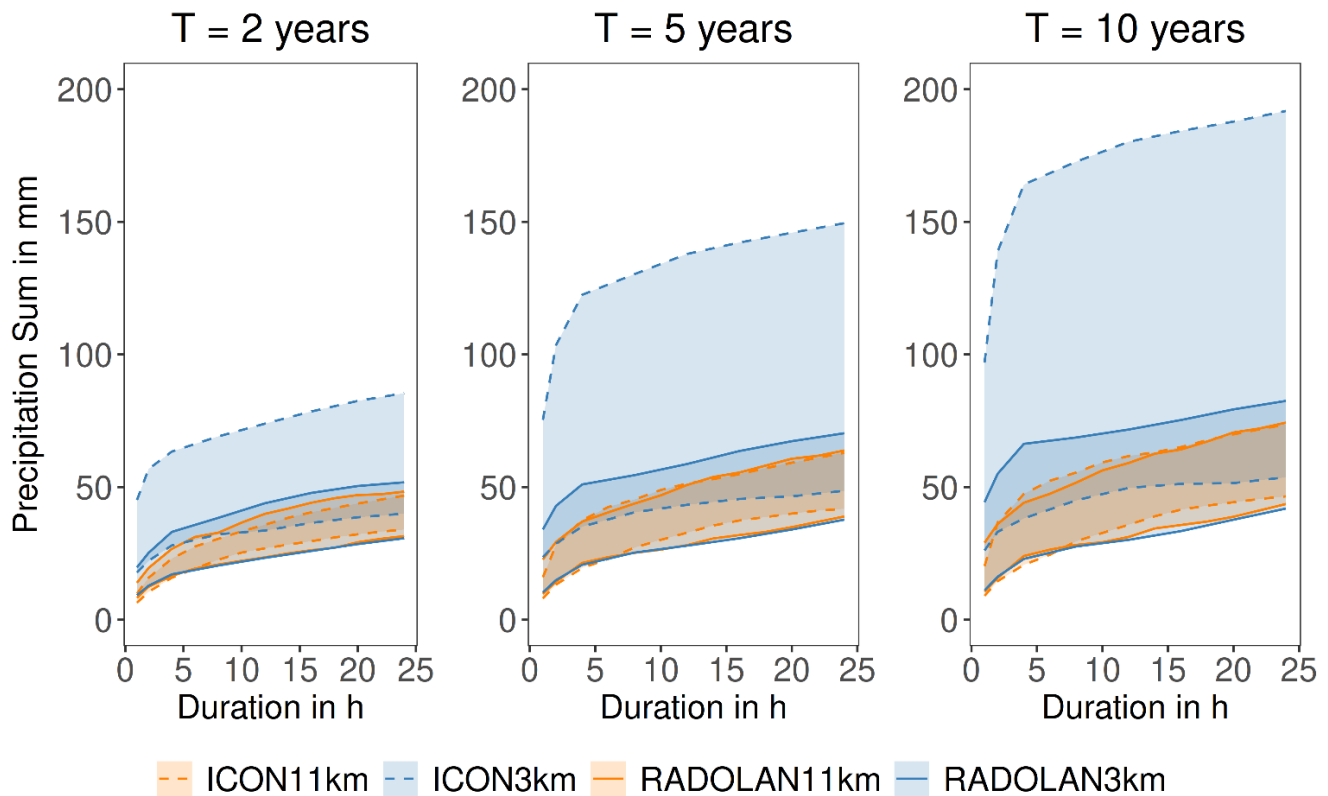


Fig. 8: Minimum-maximum bandwidth of the depth-duration-frequency curves from all cells by ICON11km and ICON3km, together with the corresponding estimates from upscaled adjusted radar data (RADOLAN11km and RADOLAN3km) for return periods T of 2, 5 and 10 years

- 330 The monthly means of the deviations of daily catchment-average precipitation computed by ICON3km and ICON11km to the adjusted radar data estimates upscaled to their respective resolution (Fig. S14 in the appendix) are similar between the climate models during autumn (SON) and winter (DJF) but differ strongly during the peak of summer. In fact, in July the monthly mean bias of daily catchment-average precipitation is 2.5-times as high for ICON3km than for ICON11km (1.0 mm to 0.4 mm), given the overestimation of heavy precipitation by ICON3km.
- 335 Furthermore, through finer grid spacing, ICON3km allows for a better spatial delineation of local precipitation fields, as are particularly distinct over mountainous regions, such as the Ore Mountains in the south-east of the study area. Thereby it offers great advantage for hydrological impact modelling. In fact, coarse climate model cells overlapping the catchment boundaries bear the risk of having their precipitation estimates risen by heavy rainfall that occurs just outside the catchment. As these climate model cells stretch into the catchment, their high precipitation estimates will be attributed to the catchment,
- 340 when in fact the storm occurred slightly outside of the catchment. The finer grid cell structure of ICON3km allows to mitigate this issue.

Precipitation anomalies to the long-term mean are studied to identify how well the climate models are able to depict months with unusual wet conditions (positive precipitation anomalies) and unusual dry conditions (negative precipitation anomalies), as well as potential performance differences between these conditions. Positive monthly precipitation anomalies were found to be predominantly overestimated by the climate models (Fig. 9), especially by ICON3km. This is also reflected in the overestimation of the mean yearly precipitation sums discussed earlier. However, the climate models were still found to be able to reach into the deep negative monthly precipitation anomalies, sometimes even overestimating them. For a considerable number of months, the negative anomalies remain however underestimated, making the picture not as clear cut as for the depiction of positive precipitation anomalies. The same holds true when it comes to the performance comparison of ICON3km and ICON11km.

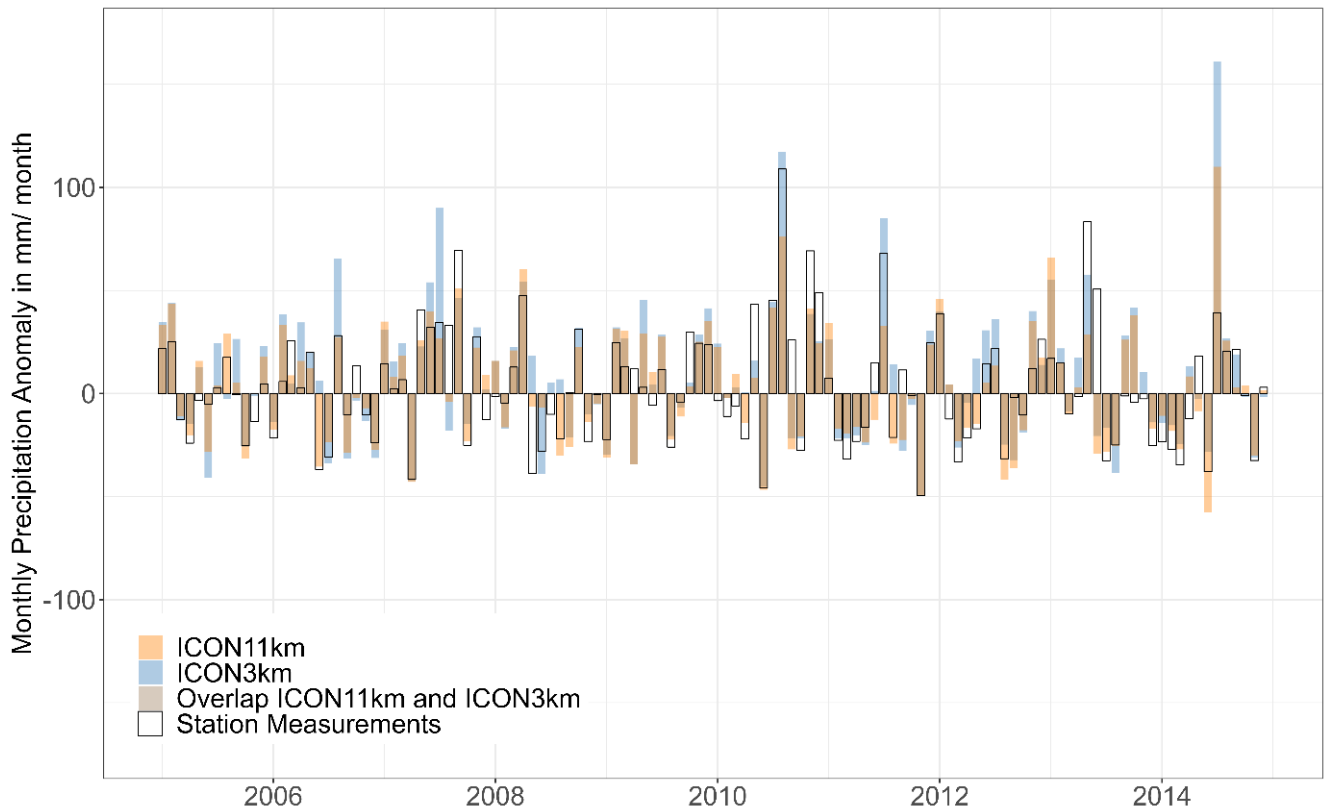


Fig. 9: Monthly precipitation anomalies to the long-term reference (1963 – 2022) for the years 2005 to 2014 for ICON11km, ICON3km and the station measurements

3.2 Hydrological Analysis

3.2.1 Calibration and Validation of the Hydrological Model

The hydrological model was calibrated manually and validated for its hourly discharge on a distinct flood, and for its weekly discharge on a one-year time series. Thiessen-interpolated meteorological point measurements were used to perform the calibration and validation.

For calibration, the largest summer flood in the area during the study period (flood of June 2013, time window: 2013-05-31 00:00 to 2013-06-06 23:00) was considered, looking at the seven catchments for which discharge observations were available. The flood occurred as a result of long-lasting rainfall, brought about by a low-pressure system over East Central Europe, which lead to warm moist air being brought from Southern Europe to Germany (LHW, 2014). At Zeitz, peak discharges were measured of an estimated return period of 100 years (LHW, 2014). Nash-Sutcliffe efficiency (NSE) values range between 0.63 and 0.95 and Kling-Gupta efficiency (KGE) values between 0.76 and 0.96 for six of these catchments, while the catchment of Weida constitutes an outlier with a NSE of only 0.14 and a KGE of 0.60. The calibration results are shown in Fig. S15 of the appendix. The simulated responses generally fell short on depicting the flashiness of the catchments and all but two of them (viz. Eisenhammer and Zeitz) showed a retardation of the flood peak. Furthermore, the baseflow was found to be erroneously modelled as barely responsive. The model was validated on the winter flood of 2011 (time window: 2011-01-07 00:00 to 2011-01-24 23:00), given that no other pronounced summer flood occurred during the study period. The flood was induced by rapid melting of an unusually thick snow layer in conjunction with heavy rainfall (LHW, 2011). The return period of the measured peak discharge at the gauge Zeitz was of 10 years and at the gauge Kleindalzig of > 25 years (LHW, 2011). The model was found to greatly overestimate direct discharge from snowmelt, computing too high and retarded flood peaks, as well as too flashy hydrographs, with a persisting underestimation of the baseflow response.

The model was furthermore calibrated and validated on measured weekly discharge of the years of 2012 and 2008 respectively, as these years are characterised by low monthly precipitation anomalies. The weeks of the year were defined according to ISO 8601. Validation shows that the model overestimates weekly direct discharge, interflow and baseflow.

Assuming the simulated discharges to be systematically distorted by the hydrological model, a qualitative comparison between the concurrent hydrological responses computed based on ICON and RADOLAN data is still possible.

Discharge results from the calibration of the model on the flood of 2013 using Thiessen-interpolated point-based observational data are almost identical to simulations of the same event using RADOLAN precipitation estimates (see Fig. S16 in the appendix). As such, the RADOLAN simulations are considered as reference in the climate model evaluations.

3.2.2 Soil Moisture

Mean hourly relative soil moisture in the hydrological model driven with ICON3km is at $0.144 \text{ m}^3/\text{m}^3$ and thereby significantly higher than in the hydrological model driven with ICON11km, where it is at $0.130 \text{ m}^3/\text{m}^3$. The hourly soil moisture simulations driven with the climate model data are furthermore significantly different in the mean to the results obtained with RADOLAN input data of equal resolution. In fact, the hydrological model driven with ICON3km shows an overestimation of mean hourly soil moisture compared to when driven with RADOLAN3km ($\overline{SM}_{ICON3km} = 0.144 \text{ m}^3/\text{m}^3$, $\overline{SM}_{RADOLAN3km} = 0.135 \text{ m}^3/\text{m}^3$). When driven with ICON11km in turn, mean hourly soil moisture is underestimated by the hydrological model compared to when driven with RADOLAN11km ($\overline{SM}_{ICON11km} = 0.130 \text{ m}^3/\text{m}^3$, $\overline{SM}_{RADOLAN11km} = 0.133 \text{ m}^3/\text{m}^3$). The overestimation, resp. underestimation is not only reflected in the mean over the entire time series, but also shows for the individual months of the year, as seen in the monthly mean biases (with an exception for April, where the bias using the ICON11km driving data is zero, see Fig. S17 in the appendix).

In fact, the soil moisture bias in the hydrological model driven with ICON11km is found to be negative compared to when driven with RADOLAN11km (Fig. 10). Looking at the model driven with ICON3km in relation to when driven with RADOLAN3km, the sign of the soil moisture bias however becomes positive for moderate to high soil moisture.

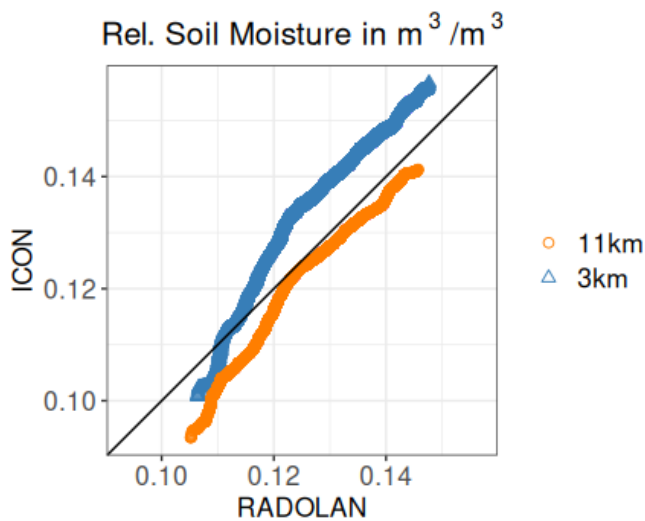


Fig. 10: QQ-plots for hourly soil moisture for the period from 2006 to 2014 for ICON11km to RADOLAN11km and ICON3km to RADOLAN3km

3.2.3 Total Evapotranspiration

Mean hourly total evapotranspiration is higher in the hydrological model driven with ICON3km than when driven with ICON11km ($\overline{ET}_{ICON3km} = 32.96 \cdot 10^{-3} \text{ mm}$, $\overline{ET}_{ICON11km} = 32.56 \cdot 10^{-3} \text{ mm}$). However, both ICON models lead to mean hourly total evapotranspiration values lower than the RADOLAN products of equal resolution ($\overline{ET}_{RADOLAN3km} =$

32.99 · 10⁻³ mm, $\overline{ET}_{RADOLAN11km} = 33.02 \cdot 10^{-3}$ mm). It is however to be noted that none of these differences are significant according to a Welch two sample t-test at 95% confidence level.

The annual cycle of the monthly mean bias of hourly catchment-average total evapotranspiration is shown in Fig. 11. The bias of ICON3km to RADOLAN3km deviates most strongly from the bias of ICON11km to RADOLAN11km in the summer months (JJA). For this time period, the negative bias of the driving model ICON11km is shifted to a positive bias by the nested model ICON3km.

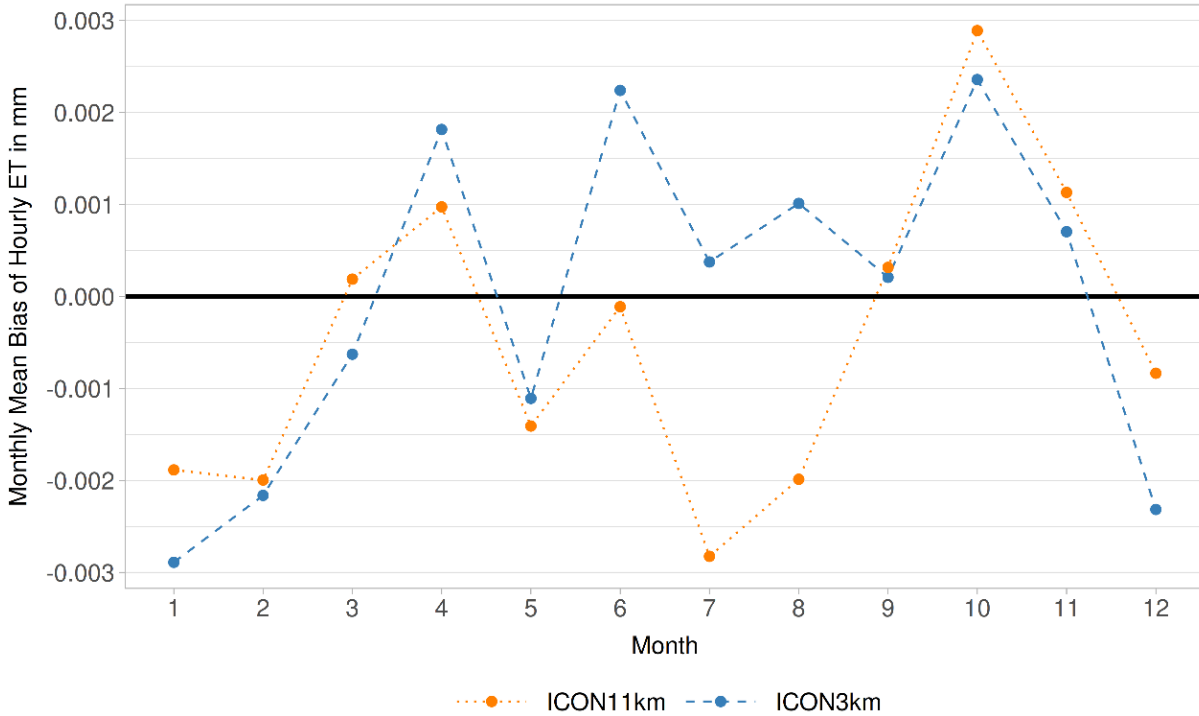
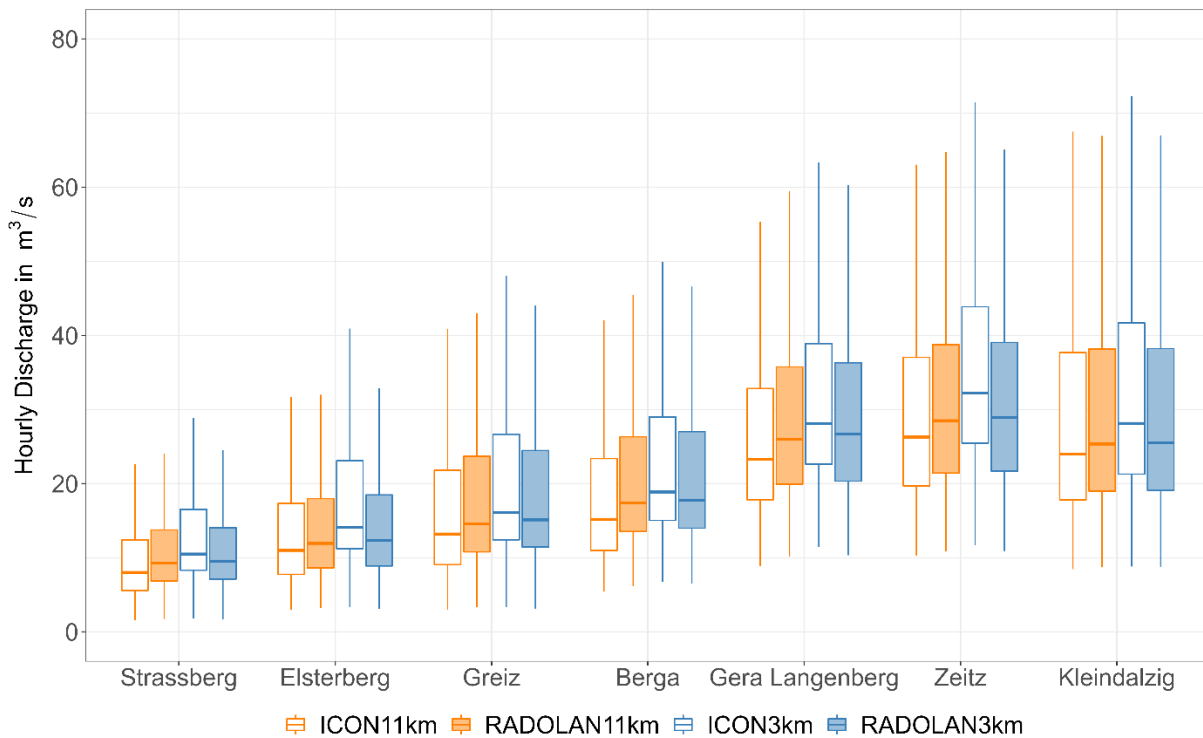


Fig. 11: Monthly mean bias of hourly total evapotranspiration, as calculated by the hydrological model driven with ICON3km and ICON11km for the period from 2006 to 2014

3.2.4 Continuous Discharge Simulation

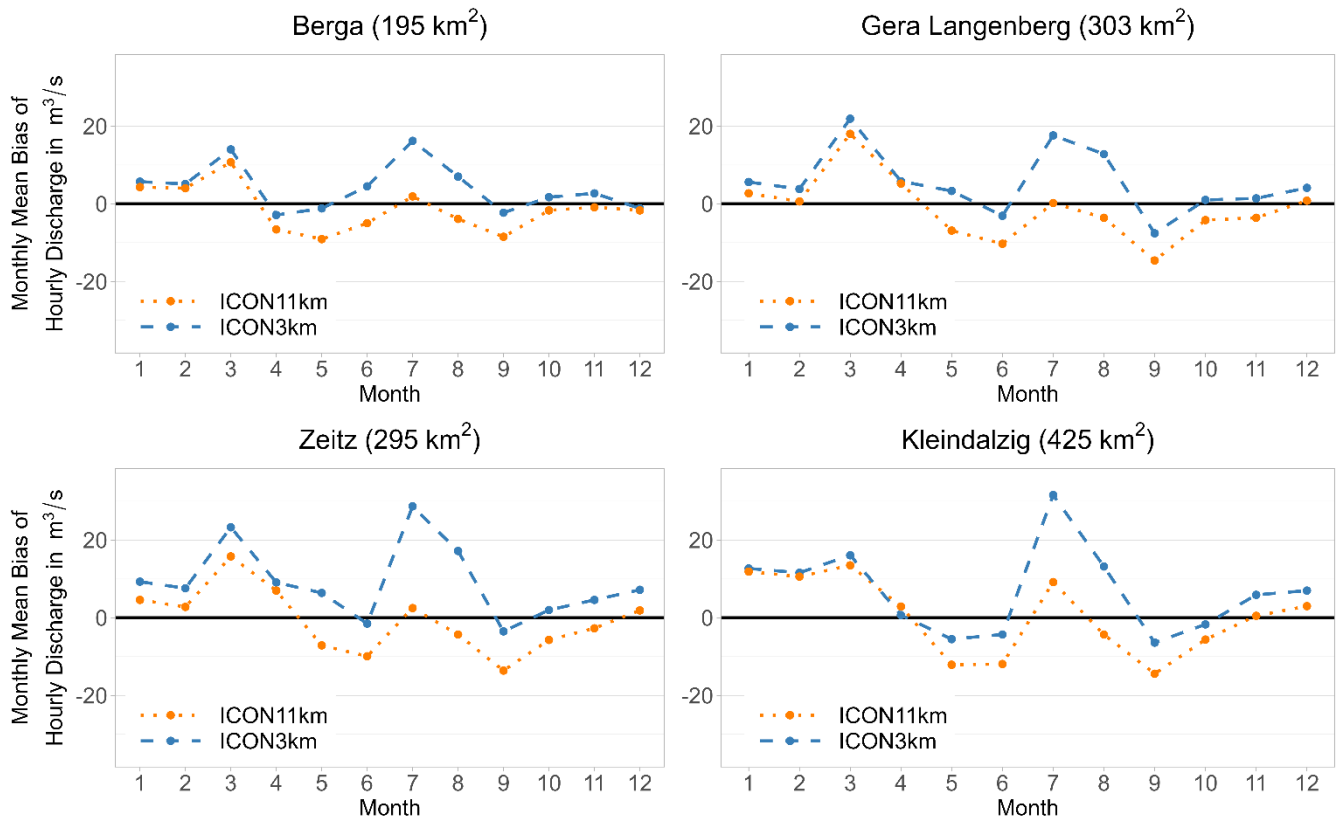
The hydrological model driven with meteorological data from ICON11km was found to compute median hourly discharge values for the study period from 2006 to 2014 that are lower than when driven with adjusted radar data of the same resolution (RADOLAN11km) (Fig. 12). In turn, when driven with data from ICON3km, the hydrological model simulated median hourly discharges that were higher than estimates based on adjusted radar data of equal resolution (RADOLAN3km).



420 **Fig. 12: Boxplots of hourly discharge (period from 2006 to 2014) computed by the WaSiM hydrological model driven with meteorological data from ICON11km and ICON3km, as well as with adjusted radar data of respective equal resolution (RADOLAN11km and RADOLAN3km) for catchments of the main stem of the Weiße Elster river within the study area. Values beyond the whiskers are not represented.**

The frequency distributions show too frequent low flows for the hydrological model driven with ICON11km and too frequent high flows when driven with ICON3km for nine out of the twelve studied catchments (with the exception of Eisenhammer, Dröda and Mylau, all of which are independent feeding catchments).

The temporal variability of the monthly mean bias of hourly discharge is similar for all studied catchments and was found to be strongly shaped by the annual cycle of precipitation bias. During winter (DJF) and early spring (March and April), the hourly discharge estimates based on the climate model data deviate to a similarly strong degree from simulation results obtained based on adjusted radar data of respective resolution (Fig. 13). However, for summer (JJA) the climate models' biases are markedly different to each other due to the pronounced overestimation of convective precipitation by ICON3km.



435 **Fig. 13: Monthly mean bias of hourly discharge from the hydrological model driven with ICON11km/ ICON3km meteorological data compared to when driven with adjusted radar data of equal resolution (RADOLAN11km/ RADOLAN3km) for the four most downstream catchments on the main stem of the Weiße Elster river within the study area for the period from 2006 to 2014**

Driving the hydrological model with ICON11km meteorological data was found to lead to an underestimation of median hourly discharge, but results in an overestimation when looking only at the 99.5th percentile of hourly discharge for six out of the seven catchments on the main stem (see Fig. S18 in the appendix). The use of ICON3km data as input to the hydrological model in turn leads to an overestimation of median hourly discharge among both the complete dataset and the 99.5th percentile.

440 Fig. 14 shows the relation between runoff and the sum of total rainfall volume and 8-h antecedent rainfall for simulations with ICON11km, ICON3km, RADOLAN11km and RADOLAN3km. The segmented regression is highly sensitive to rainfall-runoff events of high return period, such as predominantly recorded from the ICON model simulations. As such, no firm conclusions can be drawn from Fig. 14 on the existence or position of a break point in the relation between runoff and rainfall.

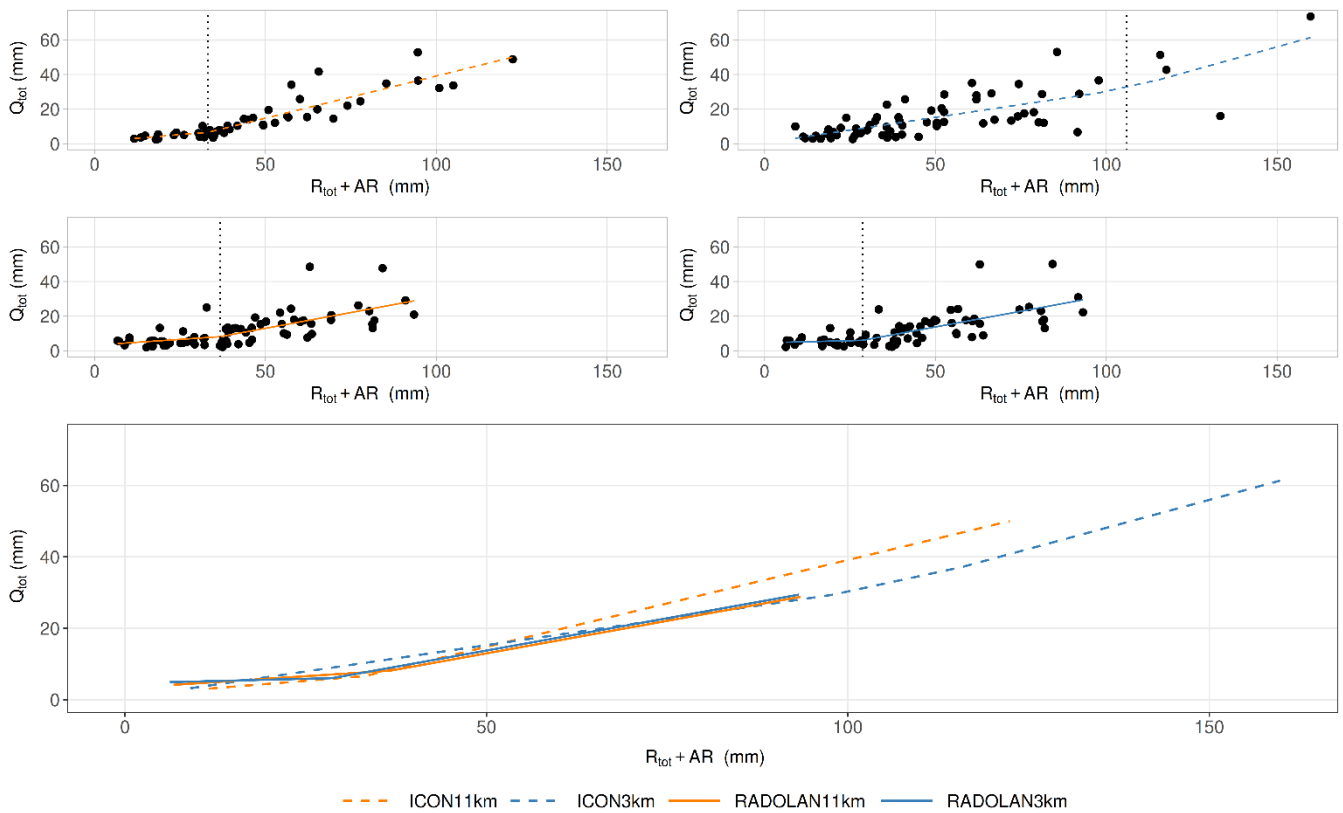
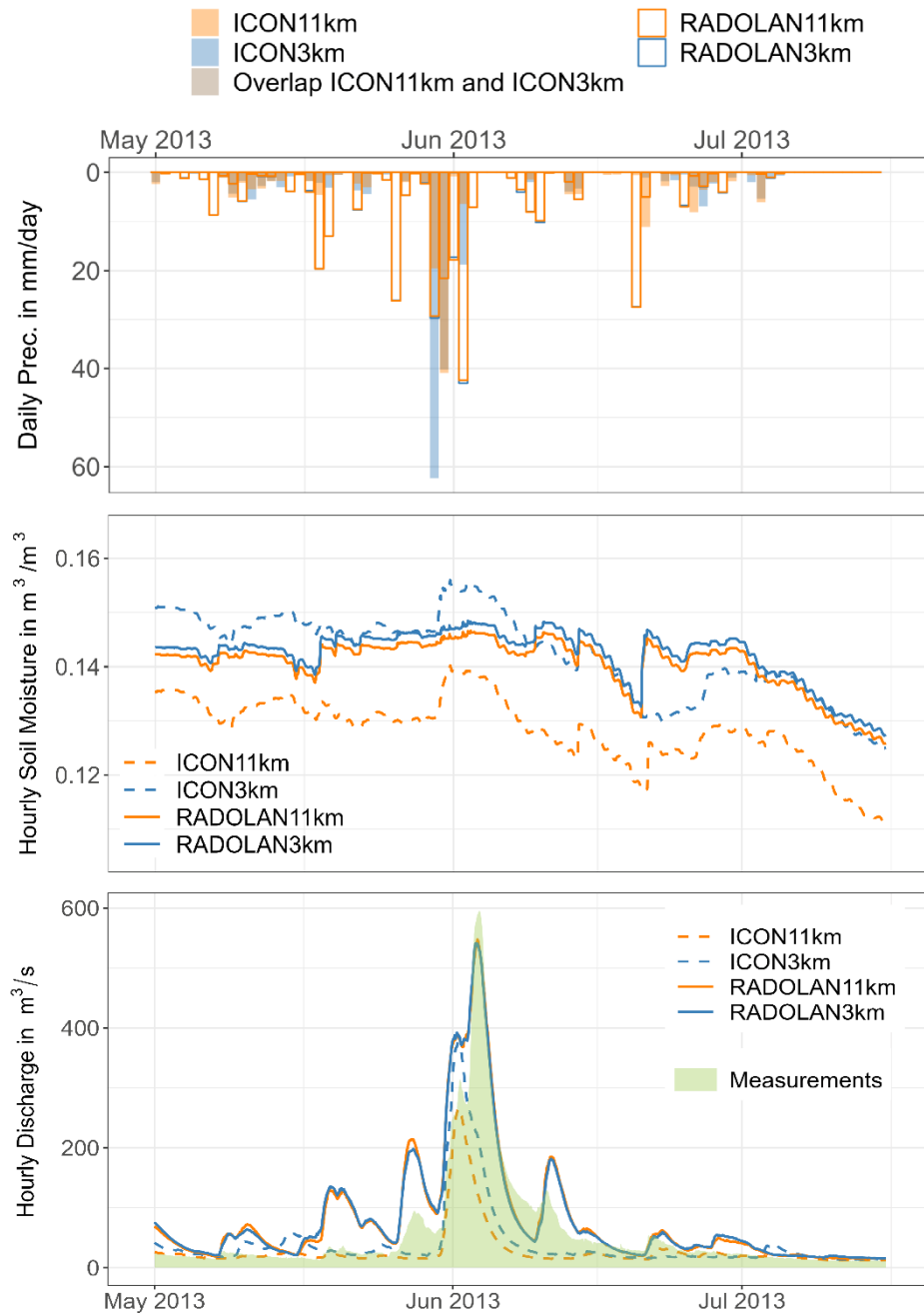


Fig. 14: Upper half: Scatter plots of total runoff (Q_{tot}) to the sum of total rainfall (R_{tot}) and 8-h antecedent rainfall (AR) over the catchment of Kleindalzig and those upstream, fitted with a piecewise linear regression model. The thresholds are indicated with a dotted vertical line. Lower half: Combined piecewise linear regression model results of Q_{tot} to $R_{tot} + AR$

450 3.2.5 Event-based Simulation

The model results for the largest flood that occurred during the study period, namely the flood of June 2013 (which is also the calibration event), are shown in Fig. 15 exemplarily for the catchment furthest downstream in the study area for which discharge measurements are available. Both climate models described a heavy precipitation event occurring too shortly with a peak earlier than observed. As a result, the hydrographs simulated based on ICON3km and ICON11km are too narrow and peak too early. Driven with RADOLAN3km and RADOLAN11km, the soil moisture results suggest that the storm rained off over saturated soils; the precipitation event did not lead to an increase of soil moisture, but to a pronounced runoff peak. In turn, ICON11km as climate input data leads to too low soil moisture, allowing for a buffering of the flood wave by the soils and an attenuation of the flood peak. ICON11km was found to underestimate the peak daily rainfall intensity of the storm. ICON3km in turn strongly overestimated it. While antecedent soil moisture in the hydrological model driven with ICON3km was similar to when driven with RADOLAN3km, however the rainfall by ICON3km likely fell over catchments with not yet fully exhausted storage capacities. In fact, the average soil moisture in the hydrological model driven with ICON3km rose,

leading to a reduction of the flood peak, thereby buffering the large overestimation of daily rainfall intensity seen in ICON3km.



465 **Fig. 15: Top: Daily spatial average precipitation estimates over the catchment of Zeitz and those upstream for the period of the 2013 flood for ICON11km, ICON3km, RADOLAN11km and RADOLAN3km, Centre: the respective hourly relative soil moisture spatial averages gained from continuous simulation, Bottom: the resulting hydrographs (using hourly data) for the catchment of Zeitz together with the discharge measurements**

Throughout the data series, ICON3km shows a number of unrealistically high precipitation intensities, occurring predominantly in summer, and of which many are clustered. This has been discussed in 3.1.5 *Precipitation*, and an example of one of the most severe downpours simulated by ICON3km was presented. In the following, it is shown how this selected heavy precipitation event translates into discharge in the Weiße Elster river basin. Despite showing a maximum hourly precipitation intensity of 143.3 mm/h (2009-07-17 17:00), the resulting peak discharge at Zeitz was of only 198 m³/s, which is somewhat lower than measured during the winter flood of 2011 (peak discharge: 253 m³/s), to which a return period of 10 years was attributed (LHW, 2011). The extreme overestimations of precipitation intensity in ICON3km occur primarily in summer and are strongly confined in space and time. The available soil storage capacity in summer is usually high, the antecedent discharge in the river low, and precipitation volumes are restricted given the limited spatial and temporal extent of the simulated storms. As a result, the discharges simulated from such unrealistically heavy downpours are within the range of recent observations at the catchment outlet.

480 4 Discussion

Here we identify strengths and limitations of the studied CPRCM in depicting hourly air temperature, global radiation, relative humidity, wind speed and precipitation over a chosen medium-size river basin in Germany, and compare these to conclusions from other studies. Previous research efforts have mainly focused on precipitation, with only few discussing other meteorological variables (Lucas-Picher et al., 2021) and seldomly for medium to small basins, but rather on larger subnational scale. Furthermore, it is discussed how the performance results from the use of the analysed CPRCM data for hydrological impact modelling compare to previous studies on other river basins with varying model chains.

4.1 Near-Surface Air Temperature

While both the studied RCM (ICON11km) and the CPRCM (ICON3km) were found to compute mean hourly air temperatures lower than observed, the deviations fall within the measurement uncertainty bandwidth. Hackenbruch et al. (2016) also showed a cold bias for COSMO-CLM at 2.8 km and 7.0 km resolution over southwestern Germany, and so did Tölle et al. (2018) for COSMO-CLM at 1.3 km and 7.0 km resolution over central Germany. These studies found higher bias in summer than in winter, however our results show the opposite, likely due to the predominance of cropland and pasture in our study area (c.f. Fig. S1). Over open land, cold biases are amplified due to higher albedo and shorter roughness lengths, the latter favouring the development of stable atmospheric conditions and allowing for stronger nocturnal cooling (Lind et al., 2020). These results underline the importance of scale and exemplarily show how dominating processes on the catchment scale can differ from those over larger study areas.

The CPRCM (ICON3km) was found to offer improvement for air temperature simulations in the summer months, in keeping with a variety of other studies conducted over Europe (Prein et al., 2015). These improvements have been linked to deep

convection being explicitly resolved, which allows for better estimates of deep-convective clouds (Leutwyler et al., 2017; Hentgen et al., 2019) and a better representation of convective precipitation (Liu et al., 2017).

Furthermore, ICON3km exhibited a slightly lower spatial bias variability throughout the year compared to ICON11km, as reflected by the 95% confidence interval from the station means. This finding is in line with Xie et al. (2025) who saw higher spatial variability in temperature bias by the RCM they used over western Norway compared the use of a CPRCM.

4.2 Global Radiation

Both ICON3km and ICON11km were found to overestimate mean daily global radiation over the study area, however the deviations were not significant under consideration of the measurement uncertainty. Hackenbruch et al. (2016) found an underestimation of half-year mean sums of global radiation over southwestern Germany (1983 – 2000) for COSMO-CLM at 2.8 km and 7.0 km resolution. In our study an underestimation was registered for the highest quantiles of daily global radiation both in intensity and frequency. ICON3km was able to reduce this bias. Through a better estimation of the onset of convection in the summer mornings and resulting reduced stratification, CPRCMs were shown to be able to dampen the RCMs' positive bias in the estimation of the frequency of high clouds (Keller et al., 2016). The simulation of more frequent clear-sky conditions allows for a reduction in the underestimation of global radiation (Keller et al., 2016; Leutwyler et al., 2017; Hentgen et al., 2019).

4.3 Relative Humidity

The skill of CPRCMs in depicting relative humidity is seldomly studied, however it has been shown that an increase in relative humidity constitutes a critical factor in deep convective initiation (Morrison et al., 2022).

No significant differences in mean hourly relative humidity were found for ICON3km or ICON11km compared to the observations under consideration of the measurement uncertainty. Monthly bias means proved negative for winter and autumn, but positive for spring and summer for both the RCM and CPRCM. Hackenbruch et al. (2016) found median relative humidity biases computed by COSMO-CLM at 2.8 km and 7.0 km resolution over southwestern Germany to be positive for both half years, with biases much greater than seen for the ICON-CLM models. Possible explanations for the smaller biases in our study might be improvements to the model physics in the newer ICON-CLM (version 2.6.4) compared to the COSMO-CLM model they used, or differing model parametrisations.

4.4 Wind Speed

ICON3km was found to estimate higher wind speed extremes than ICON11km, thereby slightly improving on the underestimation of extremes shown by both climate models. This result is in line with Iles et al. (2020), who found an increase in intensity and frequency of high winds with resolution, when comparing climate models of 12.5 km and 50 km resolution over the European domain. Similar trends were found by Kunz et al. (2010) when comparing REMO-UBA at 10

km resolution and COSMO-CLM at 18 km resolution over Germany. ICON3km improved on the depiction of the frequency of light winds (0.6 to 1.7 m/s) computed by its driving model ICON11km, however both show an overestimation of the frequency of high wind speeds. As such, the simulation of wind speed remains a challenge for both climate models. Belušić et al. (2018) advance that it takes grid spacings of a few kilometres to accurately simulate small-scale wind systems. Hackenbruch et al. (2016) prove added value for the use of a CPRCM of 2.8 km resolution for the computation of channelled wind flow in the Neckar Valley and underline the importance of orographic structures for the initialisation of local wind systems.

4.5 Precipitation

ICON11km was found to compute too frequent light precipitation, while underestimating the intensity of heavy rainfall. ICON3km in turn overestimates the intensity of the highest quantiles. These results agree with the consensus from literature (Lucas-Picher et al., 2021), e.g. with Ban et al. (2021) and Adinolfi et al. (2021). In their study over the southern United Kingdom, Kendon et al. (2012) showed an RCM of 12 km resolution to compute heavy rainfall as not intense enough, but as too persistent and widespread. A CPRCM of 1.5 km resolution in turn achieved improvements in the frequency distribution, duration and spatial extend, but estimated too intense heavy rainfall (Kendon et al., 2012). Xie et al. (2025) found an overestimation of annual maximum 1-day precipitation and maximum 1-hour precipitation using the HARMONIE climate model in convection-permitting set-up at 3 km resolution compared to running it at coarser resolution (12 km) with parametrised convection. However, they found the CPRCM to outperform the RCM in capturing hourly precipitation above the 95th percentile (Xie et al., 2025).

RCMs with convective parametrisation have shown to underestimate heavy rainfall. In fact, small atmospheric instabilities are sufficient to trigger the convection schemes in the RCMs, leading to the release of light rain and thereby inhibiting a greater computational built-up of the convective cells (Ban et al., 2021). Parametrisation schemes were designed de facto to represent average effects at coarse resolution, however not localised extreme events and individual storms (Kendon et al., 2012).

CPRCMs in turn tend to overestimate the intensity of heavy rainfall, as updrafts are computed as too wide under the coarseness of the model grid (Kendon et al., 2023). Additionally, entrainment is insufficiently captured due to turbulence not being sufficiently resolved in the convective cells (Bryan & Morrison, 2012 cited in Quintero et al., 2022).

ICON3km was found to compute unrealistically high hourly rainfall intensities of up to 157.3 mm/h over the study region, which is around twice as high as the maximum seen in RADOLAN3km (73.6 mm/h). These extremely high rainfall intensities were primarily estimated in the summer, and most were seen to cluster spatially and extend over several consecutive time steps. They are consequently likely the result of shortcomings in the parametrisation of the microphysics, rather than randomness or numerical instability.

560 ICON11km was found to compute too high mean annual precipitation totals, likely because of an overestimation of wet-hour frequency. This result is in agreement with findings by Strandberg & Lind (2020) for a 12.5 km resolution RCM over the European domain. The underestimation of average daily precipitation sums by RADOLAN compared to ground-truth rain gauge data (Kreklow et al., 2020) further exacerbates this positive bias.

The deviations of ICON3km and ICON11km to the adjusted radar data estimates upscaled to their respective resolution were found to be similar between the climate models during autumn and winter but differing most strongly during the peak of summer. In fact, synoptic weather systems, which are dominant in winter, are well represented even in coarse climate models (Strandberg & Lind, 2020) and the CPRCM is likely not providing much additional information or added value to their depiction. For summer however, when small-scale convective processes are at the forefront, the explicit representation of convection in the CPRCM likely leads to pronounced differences compared to estimates by its driving RCM where convection is parametrised (Strandberg & Lind, 2020).

Monthly negative precipitation anomalies were not found to be better represented by ICON3km than by ICON11km. They are predominantly shaped by large-scale atmospheric circulations, handed over to the CPRCM as lateral boundary conditions by its driving RCM. Taylor et al. (2013) however found a change from positive to negative in the sign of the soil-moisture precipitation feedback over their semi-arid domain when switching from an RCM with parametrised convection to a CPRCM. A negative feedback makes a drought less likely, whereas a positive feedback may prolong or intensify it (Taylor et al., 2013). Kendon et al. (2019) further show over an Africa-wide domain, that the use of a CPRCM (4.5 km resolution) can be of added value for the estimation of the length of dry spells through a more accurate depiction of the triggering of diurnal rainfall and an improved representation of the dying out of westward propagating systems.

4.6 Soil Moisture

580 Reszler et al. (2018) looked at soil moisture on a monthly basis and found estimates from a hydrological model driven with CPRCM data to be similar in winter and spring to those when using RCM input data, however the estimates deviate considerably in summer and autumn. In fact, during summer and autumn both climate models underestimate soil moisture, but the CPRCM is able to reduce the negative bias seen in the hydrological model driven with RCM data (Reszler et al., 2018). In contrast, ICON3km as input to the hydrological model WaSiM lead to an overestimation of hourly soil moisture throughout the year, while ICON11km lead to an underestimation. However, the results agree with the study of Reszler et al. (2018) on higher soil moisture in summer and autumn by the hydrological model driven with CPRCM data than when driven with RCM data.

4.7 Total Evapotranspiration

Biases in observed and modelled meteorological variables can greatly influence evapotranspiration estimates. Meyer et al. (1989) introduced random and systematic errors to the input data of the Penman method to study the effect of such errors on the estimation of daily potential evapotranspiration. They found that random errors distort potential evapotranspiration estimates in a similar manner than systematic errors. Standard errors in potential evapotranspiration were fairly low when introducing random errors on wind run data, but three times as high when applying random errors of the same scale on temperature data, and four to five times as high when adding the random errors on relative humidity or solar radiation data (Meyer et al., 1989). In climate modelling, however the uncertainty is not equally distributed among variables. Lai et al. (2022) investigated the missing-variable estimation approaches by FAO and how the selection or exclusion of certain climate variables from global climate models shape the uncertainty of projected reference crop evapotranspiration using the Penman-Monteith method. They concluded that air temperature and shortwave radiation should be prioritised given the low projection uncertainty of these variables and high influence on reference evapotranspiration. Kay & Davies (2008) found the uncertainty introduced by the climate model structure to be higher than that by the potential evaporation formulation.

Mendoza et al. (2016) found higher basin-averaged mean annual evapotranspiration estimates in hydrological modelling when using meteorological forcing data from the WRF model at 4 km resolution in convection-permitting set-up than at coarser resolution of 12 km or 36 km with parametrised convection. These results are in keeping with the study presented here, where higher mean hourly total evapotranspiration is registered for the hydrological model driven with ICON3km compared to when driven with ICON11km. During summer (JJA), the monthly mean biases of hourly catchment-average total evapotranspiration computed by the hydrological model driven with ICON3km and ICON11km deviate most strongly, with a shift from a negative bias by the driving model ICON11km to a positive bias by the driving model ICON3km. Higher total evapotranspiration by ICON3km, especially in summer, is likely the result of higher water availability through higher precipitation intensities in ICON3km, as well as a reduction of the negative summer bias in temperature and global radiation by ICON3km. Mendoza et al. (2016) also saw a wet bias in the CPRCM they studied. They furthermore analysed the partitioning of basin-averaged mean annual precipitation into basin-averaged mean annual evapotranspiration and found partitioning to be lower when the hydrological model is driven with CPRCM data than when driven with data from the RCMs.

4.8 Discharge Simulations

All the discussed meteorological variables are used as input to the hydrological model and shape through their biases the accuracy of the discharge simulations. The hydrological model WaSiM driven with ICON3km simulates higher discharge than when driven with ICON11km. These findings are in agreement with the consensus from literature. Kay et al. (2015) report higher low flow volumes, median flows and high flow volumes when using CPRCM data compared to when using

620 data from an RCM with parametrised convection. Mendoza et al. (2016) also saw higher basin-averaged mean annual runoff when driving their hydrological model with a CPRCM than with an RCM. Reszler et al. (2018) simulated higher maximum annual flood peaks using data from a CPRCM of 3 km resolution than from an RCM of 12.5 km resolution. Higher peak streamflows were also found by Schaller et al. (2020) using CPRCM data to simulate a selected autumn flood event on the West Coast of Norway compared to the use of GCM data. Davis et al. (2022) simulated higher (maximum) water levels in an urban catchment in India driving their hydrological model with data from a CPRCM ensemble than with data from the
625 Global Forecasting System. Ascott et al. (2023) in turn computed greater daily river flows in the Lake Victoria basin when using CPRCM data compared to the use of RCM data. In keeping, Poncet et al. (2024) and Xie et al. (2025) found higher peak-over-threshold historical flood peak discharges using CPRCM data than when using RCM data to drive their hydrological models. Archer et al. (2024) found larger relative flooded areas from the historical intensity-duration-frequencies derived from a CPRCM ensemble than from interpolated rainfall observations.

630 Given its strong overestimation of precipitation intensity, ICON3km did not allow for improved discharge simulation compared to ICON11km. Kay et al. (2015), Reszler et al. (2018) and Xie et al. (2025) attributed no added value to the use of their studied CPRCM for hydrological impact modelling. In fact, Kay et al. (2015) found worse performance in discharge computation using the 1.5 km CPRCM compared to using its driving 12 km RCM as input to the gridded hydrological model CLASSIC-GB, given the strong positive precipitation bias of the CPRCM when left uncorrected. Reszler et al. (2018) found
635 deficiencies in the depiction of the temporal distribution of rainfall intensities to be another potential source of error. In a later study, Kay (2022) did however find added value in the use of a 2.2 km CPRCM compared to a 12 km RCM (each 12 member ensembles) for the simulation of daily low flow volume, median flow, and high flow volume across Britain. They attribute these improvements to changes in the physics of the used CPRCM compared to the earlier model employed in their 2015 study (c.f. Kay et al., 2015). Also Mendoza et al. (2016), Schaller et al. (2020), Davis et al. (2022) and Poncet et al.
640 (2024) found various added value in the use of CPRCMs for hydrological modelling (c.f. Table 1). Xie et al. (2025) only found added value for the simulation of extreme flood peaks, characterised by return periods above 10 years.

The 99.5th percentile of hourly discharge was overestimated in the median by the hydrological model driven with ICON3km, as a result of the overestimation of the intensity of heavy precipitation. The input of ICON11km data to the hydrological model also resulted in an overestimation of hourly discharge of the 99.5th percentile in the median for six out of
645 the seven catchments on the main stem, despite an underestimation of precipitation intensity. An overestimation of high discharges by the hydrological model driven with RCM data is also apparent in the study of Reszler et al. (2018). For four out of the six catchments they studied, the return periods from the hydrological model driven with 12.5 km RCM data overtake or at least come to match the return period curves from the observations for high return periods, while undercutting them for lower return periods.

650 We showed exemplarily for the flood event of 2013 in the Weiße Elster catchment, that a simulated overestimation in precipitation does not always translate into an overestimation in flood peak discharge, as the storm cell might be shifted in space in the climate model simulation and rain off over soils with greater available storage capacity. In fact, in simulations run over vast domains, such as the Central European (CEU) domain in this study, the general events and patterns can be expected to match, but the climate models are unlikely to perfectly capture the actual geographic position of a local storm.

655 To conclude, no added value could be found in the use of uncorrected climate model data from ICON3km for hydrological impact modelling, neither across its complete range, nor for the 99.5th percentile. The results presented here are drawn from a short time series (10 years) obtained from a single CPRCM run and reflect the hydroclimatic simulations over a single cluster of catchments. Longer time series, climate model ensembles and additional study areas are needed to draw firm conclusions on the general performance of CPRCMs for hydrological impact modelling.

660 5 Conclusion

Climate change projections are key for the development of mitigation and adaptation strategies (IPCC, 2023). Policy decisions rely on climate model results that are not only to be of low bias and uncertainty, but also of high spatial and temporal resolution. CPRCMs bear major potential for advancement towards these requirements, as they run on fine model grids and no longer rely on deep-convection parametrisation schemes, which come with great uncertainty (Ban et al., 2021).
665 Yet to this day, their performance in depicting meteorological variables on the catchment scale and their potential added value for hydrological impact modelling is scarcely studied (Lucas-Picher et al., 2021).

In the work presented here, the skill of estimating meteorological variables on the catchment scale is analysed for ICON-CLM 2.6.4 in its convection-permitting setup at 3 km resolution (ICON3km), and for its driving model ICON-CLM 2.6.4 at 11 km resolution with parametrised convection (ICON11km, forced with ECMWF-ERA5). Analyses are conducted
670 exemplarily over the Weiße Elster basin in East Central Germany for the historical period from 2005 to 2014. It is further studied, how well the given uncorrected climate model data is suited for hydrological impact modelling in the catchment over the period from 2006 to 2014 using the distributed physically based hydrological model WaSiM, and whether the CPRCM provides added value.

While ICON3km offers some improvements in the depiction of meteorological variables on the catchment scale, however for
675 a variety of other aspects no added value is to be found. ICON3km allows for an improved depiction of wind speed characteristics and shows improvements in the estimates of mean hourly air temperature and summer daily global radiation. However, the strong overestimation of the intensity and frequency of heavy rainfall by ICON3km impedes its suitability for hydrological impact modelling, as it translates into a pronounced overestimation of discharge in the hydrological model. With only a decadal-long time series from one CPRCM over a single basin studied, these results are however only able to

680 give a hint at possible strengths and limitations of CPRCMs. The current body of literature shows a promising evolution of
CPRCM results for hydrological impact modelling, while also highlighting persisting shortcomings and showcasing
individual CPRCM runs with no apparent overall added value for hydrological impact modelling, as presented in this paper
and e.g. in Xie et al. (2025). For robust generalised performance indications of CPRCMs, large ensembles and time spans,
over varying study areas would be required. Coordinated research efforts are key to achieving this and have started to gain
685 momentum in recent years (Lucas-Picher et al., 2021).

While biases induced by the climate models could be corrected through adjustment of the meteorological data or through
calibration of the hydrological model, doing so comes with great limitations. On the one hand, bias correction of high spatial
and temporal resolution meteorological data comes with major challenges (e.g. Haerter et al., 2011; Addor & Seibert, 2014).
Furthermore, the bias correction approaches commonly rely on the assumption of stationary bias, making it questionable for
690 climate projections (Huang et al., 2014). On the other hand, in conceptual lumped models, systematic uni- or multivariate
errors in the input data can to a certain degree be straightened out during the calibration process (Xie et al., 2025). This is
more difficult in process-based deterministic models (Xie et al., 2025). However, using hydrological model calibration to
overcome shortcomings in the meteorological data comes with the risk that parameters are calibrated outside of their
plausible range. Furthermore, parameters obtained from calibration on historical discharge may not be valid under climate
695 change projections (Merz et al., 2011; Osuch et al., 2015). The question of the optimal choice of the hydrological modelling
structure for one-way coupling with CPRCMs remains a topic of ongoing research. While some studies show process-based
deterministic models to outperform conceptual lumped models in the simulation of rainfall-generated flood peaks, flood
frequency and flood seasonality from CPRCMs (Xie et al., 2025), other studies show better estimates of the number of peaks
over threshold using a lumped model (Poncet et al., 2024). More studies are needed to draw firm conclusions on the most
700 advantageous modelling structure for coupling with CPRCMs, also under consideration of the high computational costs of
process-based fully distributed hydrological models.

Under consideration of expected model improvements over the next few years (Kendon et al., 2021), the results from this
study suggest particular potential of CPRCMs for the simulation of hydroclimatic conditions in summer and for hydrological
impact modelling in regions of highly mountainous or urban character. It is advised to direct research efforts to these areas.

705 **6 Code availability**

The code scripts are available from the corresponding author upon request.

7 Data availability

Currently the climate model data is only available for partners of the RegIKlim project via the Free Evaluation System Framework (Freva). Open access to the data is planned for the near future.

710 8 Competing interests

The authors declare that they have no conflict of interest.

9 Author contribution

OW worked on the methodology, formal analysis, software, visualisation and writing (original draft preparation and editing). VM contributed to the conceptualization, data curation, resources, supervision and review. LB was involved in the funding acquisition, supervision and review.
715

10 Acknowledgments

We thank Klemens Barfus for insightful comments during the development of this paper, as well as Jeewanthi Thotapitiya and Laura Detjen for valuable feedback on the manuscript. Oakley Wagner was financed by the Helmholtz Institute for Climate Service Science (HICSS), a cooperation between Climate Service Center Germany (GERICS) and University of Hamburg, Germany. The climate model data was provided in the frame of the NUKLEUS (Actionable Local Climate Information for Germany) project, which is part of the larger RegIKlim project funded by the Federal Ministry of Education and Research (BMBF; grant number: 01LR2002A-G). The climate models were run by Klaus Keuler and Michael Woldt at BTU Cottbus-Senftenberg in collaboration with the CLM-Community and EURO-CORDEX. We are grateful for their support. The hydrological results constitute part of the KlimaKonform project at TU Dresden, a sub-project of RegIKlim.
720
725 Observational meteorological data was provided by Deutscher Wetterdienst (DWD). The catchment-specific setup and calibration of the hydrological model was built upon an earlier version provided by Sherifdeen Olamilekan Babalola (Babalola, 2023). We thank Deutsches Klimarechenzentrum (DKRZ) and the Centre for Information Services and High-Performance Computing at TU Dresden (ZIH) for generously providing computational resources for this study. We extend our thanks to three anonymous reviewers for their constructive feedback, which allowed to improve the manuscript.

730

References

- Addor, N. and Seibert, J.: Bias correction for hydrological impact studies – beyond the daily perspective, *Hydrol. Process.*, 28, 4823–4828, <https://doi.org/10.1002/hyp.10238>, 2014.
- Adinolfi, M., Raffa, M., Reder, A., and Mercogliano, P.: Evaluation and Expected Changes of Summer Precipitation at
735 Convection Permitting Scale with COSMO-CLM over Alpine Space, *Atmosphere*, 12, 54, <https://doi.org/10.3390/atmos12010054>, 2021.
- Archer, L., Hatchard, S., Devitt, L., Neal, J. C., Coxon, G., Bates, P. D., Kendon, E. J., and Savage, J.: Future Change in Urban Flooding Using New Convection-Permitting Climate Projections, *Water Resour. Res.*, 60, <https://doi.org/10.1029/2023WR035533>, 2024.
- 740 Ascott, M. J., Christelis, V., Lapworth, D. J., Macdonald, D., Tindimugaya, C., Iragena, A., Finney, D., Fitzpatrick, R., Marsham, J. H., and Rowell, D. P.: On the application of rainfall projections from a convection-permitting climate model to lumped catchment models, *J. Hydrol.*, 617, 129097, <https://doi.org/10.1016/j.jhydrol.2023.129097>, 2023.
- Babalola, S. O.: Adoption of the Meteorological Conditions of the Ahr River Valley Flood of July 2021 to the Upper Weiße Elster River Catchment, Germany for Flood Hazard Analysis, Master Thesis, TU Dresden, 2023.
- 745 Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., Belušić, D., Berthou, S., Brisson, E., Cardoso, R. M., Chan, S. C., Christensen, O. B., Fernández, J., Fita, L., Frisius, T., Gašparac, G., Giorgi, F., Goergen, K., Haugen, J. E., Hodnebrog, Ø., Kartsios, S., Katragkou, E., Kendon, E. J., Keuler, K., Lavin-Gullon, A., Lenderink, G., Leutwyler, D., Lorenz, T., Maraun, D., Mercogliano, P., Milovac, J., Panitz, H.-J., Raffa, M., Remedio, A. R., Schär, C., Soares, P. M. M., Srnec, L., Steensen, B. M., Stocchi, P., Tölle, M. H., Truhetz, H., Vergara-
750 Temprado, J., Vries, H. de, Warrach-Sagi, K., Wulfmeyer, V., and Zander, M. J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution: Part I: evaluation of precipitation, *Clim. Dyn.*, 57, 275–302, <https://doi.org/10.1007/s00382-021-05708-w>, 2021.
- Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, *Quart J Royal Meteorol Soc*, 134, 1337–1351, <https://doi.org/10.1002/qj.289>, 2008.
- 755 Belušić, A., Prtenjak, M. T., Güttler, I., Ban, N., Leutwyler, D., and Schär, C.: Near-surface wind variability over the broader Adriatic region: insights from an ensemble of regional climate models, *Clim. Dyn.*, 50, 4455–4480, <https://doi.org/10.1007/s00382-017-3885-5>, 2018.

BGR & SGD: Bodenübersichtskarte der Bundesrepublik Deutschland 1:200.000, 2018.

760 Brinckmann, S. and Dirksen, R.: Eine Methode zur Korrektur und Unsicherheitsbestimmung der Temperaturmessungen an den Klimareferenzstationen, *MOL-RAO Aktuell*, 1, 2020.

Brutsaert, W.: *Evaporation into the atmosphere*, Kluwer Academic Publishers, Dordrecht, 1982.

Bryan, G. H. and Morrison, H.: Sensitivity of a Simulated Squall Line to Horizontal Resolution and Parameterization of Microphysics, *Mon. Weather Rev.*, 140, 202–225, <https://doi.org/10.1175/MWR-D-11-00046.1>, 2012.

765 Buttle, J. M., Webster, K. L., Hazlett, P. W., and Jeffries, D. S.: Quickflow response to forest harvesting and recovery in a northern hardwood forest landscape, *Hydrol. Process.*, 33, 47–65, <https://doi.org/10.1002/hyp.13310>, 2019.

Copernicus Land Monitoring Service, <https://doi.org/10.2909/960998c1-1870-4e82-8051-6485205ebbac>, 2020.

Davis, S., Pentakota, L., Saptarishy, N., and Mujumdar, P. P.: A Flood Forecasting Framework Coupling a High Resolution WRF Ensemble With an Urban Hydrologic Model, *Front. Earth Sci.*, 10, <https://doi.org/10.3389/feart.2022.883842>, 2022.

770 DWD: Pyranometer: Messung von Global- und Diffusstrahlung, <https://www.dwd.de/DE/leistungen/solarenergie/pyranometer.html?nn=16102>, last access: 29 October 2023, n.d.

DWD: Data Set Description: Historical hourly RADOLAN grids of precipitation depth (binary), 2021.

Fantini, A., Raffaele, F., Torma, C., Bacer, S., Coppola, E., Giorgi, F., Ahrens, B., Dubois, C., Sanchez, E., and Verdecchia, M.: Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-CORDEX and EURO-CORDEX
775 experiments against high resolution observations, *Clim. Dyn.*, 51, 877–900, <https://doi.org/10.1007/s00382-016-3453-4>, 2018.

Friedlander, S. K. and Topper, L.: *Turbulence: Classic Papers on Statistical Theory*, Interscience, 187, 1961.

Giorgi, F.: Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next?, *J. Geophys. Res. Atmos.*, 124, 5696–5723, <https://doi.org/10.1029/2018JD030094>, 2019.

780 Gutowski, W. J., Ullrich, P. A., Hall, A., Leung, L. R., O'Brien, T. A., Patricola, C. M., Arritt, R. W., Bukovsky, M. S., Calvin, K. V., Feng, Z., Jones, A. D., Kooperman, G. J., Monier, E., Pritchard, M. S., Pryor, S. C., Qian, Y., Rhoades, A. M., Roberts, A. F., Sakaguchi, K., Urban, N., and Zarzycki, C.: The Ongoing Need for High-Resolution Regional Climate

- Models: Process Understanding and Stakeholder Information, *Bull. Am. Meteorol. Soc.*, 101, E664-E683, <https://doi.org/10.1175/BAMS-D-19-0113.1>, 2020.
- 785 Hackenbruch, J., Schädler, G., and Schipper, J. W.: Added value of high-resolution regional climate simulations for regional impact studies, *Meteorol. Z.*, 25, 291–304, <https://doi.org/10.1127/metz/2016/0701>, 2016.
- Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C.: Climate model bias correction and the role of timescales, *Hydrol. Earth Syst. Sci.*, 15, 1065–1079, <https://doi.org/10.5194/hess-15-1065-2011>, 2011.
- 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, 790 <https://doi.org/10.1029/2008JD010201>, 2008.
- Hentgen, L., Ban, N., Kröner, N., Leutwyler, D., and Schär, C.: Clouds in Convection-Resolving Climate Simulations Over Europe, *J. Geophys. Res. Atmos.*, 124, 3849–3870, <https://doi.org/10.1029/2018JD030150>, 2019.
- Huang, S., Krysanova, V., and Hattermann, F. F.: Does bias correction increase reliability of flood projections under climate change? A case study of large rivers in Germany, *Intl Journal of Climatology*, 34, 3780–3800, 795 <https://doi.org/10.1002/joc.3945>, 2014.
- Iles, C. E., Vautard, R., Strachan, J., Joussaume, S., Eggen, B. R., and Hewitt, C. D.: The benefits of increasing resolution in global and regional climate simulations for European climate extremes, *Geosci. Model Dev.*, 13, 5583–5607, <https://doi.org/10.5194/gmd-13-5583-2020>, 2020.
- 800 IPCC: Summary for Policymakers, in: *Climate Change 2021: The Physical Science Basis: Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 3–32, 2021.
- 805 IPCC: *Climate Change 2023: Synthesis Report: Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 35-115, <https://doi.org/10.59327/IPCC/AR6-9789291691647>, 2023.
- Jiang, L. and O’Neill, B. C.: Global urbanization projections for the Shared Socioeconomic Pathways, *Glob. Environ. Change*, 42, 193–199, <https://doi.org/10.1016/j.gloenvcha.2015.03.008>, 2017.

- Kam, P. M., Aznar-Siguan, G., Schewe, J., Milano, L., Ginnetti, J., Willner, S., McCaughey, J. W., and Bresch, D. N.:
810 Global warming and population change both heighten future risk of human displacement due to river floods, *Environ. Res. Lett.*, 16, 44026, <https://doi.org/10.1088/1748-9326/abd26c>, 2021.
- Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., and Brumby, S. P.: Global land use / land cover with Sentinel 2 and deep learning, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11/07/2021 - 16/07/2021, 2021.
- 815 Kay, A. L. and Davies, H. N.: Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts, *J. Hydrol.*, 358, 221–239, <https://doi.org/10.1016/j.jhydrol.2008.06.005>, 2008.
- Kay, A.: Differences in hydrological impacts using regional climate model and nested convection-permitting model data, *Clim. Chang.*, 173, <https://doi.org/10.1007/s10584-022-03405-z>, 2022.
- Kay, A. L., Rudd, A. C., Davies, H. N., Kendon, E. J., and Jones, R. G.: Use of very high resolution climate model data for
820 hydrological modelling: baseline performance and future flood changes, *Clim. Chang.*, 133, 193–208, <https://doi.org/10.1007/s10584-015-1455-6>, 2015.
- Keller, M., Fuhrer, O., Schmidli, J., Stengel, M., Stöckli, R., and Schär, C.: Evaluation of convection-resolving models using satellite data: The diurnal cycle of summer convection over the Alps, *Meteorol. Z.*, 25, 165–179, <https://doi.org/10.1127/metz/2015/0715>, 2016.
- 825 Kendon, E. J., Fischer, E. M., and Short, C. J.: Variability conceals emerging trend in 100yr projections of UK local hourly rainfall extremes, *Nat. Commun.*, 14, 1133, <https://doi.org/10.1038/s41467-023-36499-9>, 2023.
- Kendon, E. J., Prein, A. F., Senior, C. A., and Stirling, A.: Challenges and outlook for convection-permitting climate modelling, *Philos Trans A Math Phys Eng Sci*, 379, 20190547, <https://doi.org/10.1098/rsta.2019.0547>, 2021.
- Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J.: Realism of Rainfall in a Very High-Resolution Regional
830 Climate Model, *J. Clim.*, 25, 5791–5806, <https://doi.org/10.1175/JCLI-D-11-00562.1>, 2012.
- Kendon, E. J., Stratton, R. A., Tucker, S., Marsham, J. H., Berthou, S., Rowell, D. P., and Senior, C. A.: Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale, *Nat. Commun.*, 10, 1794, <https://doi.org/10.1038/s41467-019-09776-9>, 2019.

- Körner, P.: Radarbasierte Niederschlagsdaten 2001 - 2020: Radarbasierte Niederschlagsdaten im Zeitraum 2001-2020, Beschaffung, Aufbereitung und Bereitstellung im ReKIS, <https://publikationen.sachsen.de/bdb/artikel/39692/documents/60999>, 2022.
- Koutsoyiannis, D., Kozonis, D., and Manetas, A.: A mathematical framework for studying rainfall intensity-duration-frequency relationships, *J. Hydrol.*, 206, 118–135, 1998.
- Kreklow, J., Tetzlaff, B., Burkhard, B., and Kuhnt, G.: Radar-Based Precipitation Climatology in Germany—Developments, Uncertainties and Potentials, *Atmosphere*, 11, 217, <https://doi.org/10.3390/atmos11020217>, 2020.
- Kronenberg, R., Franke, J., Neumann, T., Struve, S., Bernhofer, C., and Sommer, W.: Das Regionale Klimainformationssystem ReKIS – eine gemeinsame Plattform für Sachsen, Sachsen-Anhalt und Thüringen, in: *Umweltinformationssysteme: Grundlagen einer angewandten Geoinformatik/Geo-IT*, 3rd ed., edited by: Fischer-Stabel, P., Herbert Wichmann-Verlag, 2021.
- Kunz, M., Mohr, S., Rauthe, M., Lux, R., and Kottmeier, C.: Assessment of extreme wind speeds from Regional Climate Models: Part 1: Estimation of return values and their evaluation, *Nat. Hazards Earth Syst. Sci.*, 10, 907–922, <https://doi.org/10.5194/nhess-10-907-2010>, 2010.
- Kyrouac, J. and Theisen, A.: Biases of the MET Temperature and Relative Humidity Sensor (HMP45) Report, 2017.
- Lai, C., Chen, X., Zhong, R., and Wang, Z.: Implication of climate variable selections on the uncertainty of reference crop evapotranspiration projections propagated from climate variables projections under climate change, *Agric. Water Manag.*, 259, 107273, <https://doi.org/10.1016/j.agwat.2021.107273>, 2022.
- Leutwyler, D., Lüthi, D., Ban, N., Fuhrer, O., and Schär, C.: Evaluation of the convection-resolving climate modeling approach on continental scales, *J. Geophys. Res. Atmos.*, 122, 5237–5258, <https://doi.org/10.1002/2016JD026013>, 2017.
- LHW: Bericht über das Hochwasser Januar 2011, Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt, https://klimawandel.geo.uni-halle.de/sites/default/files/abschlussbericht_2011.pdf, 2011.
- LHW: Bericht über das Hochwasser im Juni 2013 in Sachsen-Anhalt: Entstehung, Ablauf, Management und statistische Einordnung, Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt, https://lhw.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/Landesbetriebe/LHW/neu_PDF/4.0/SB_Hochwasserschutz/Hochwasserbericht_2013.pdf, 2014.

860 Lind, P., Belušić, D., Christensen, O. B., Dobler, A., Kjellström, E., Landgren, O., Lindstedt, D., Matte, D., Pedersen, R. A., Toivonen, E., and Wang, F.: Benefits and added value of convection-permitting climate modeling over Fenno-Scandinavia, *Clim. Dyn.*, 55, 1893–1912, <https://doi.org/10.1007/s00382-020-05359-3>, 2020.

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., Chen, F., Chen, L., Clark, M., Dai, A., Dudhia, J., Eidhammer, T., Gochis, D., Gutmann, E., Kurkute, S., Li, Y., Thompson, G., and Yates, D.: Continental-scale
865 convection-permitting modeling of the current and future climate of North America, *Clim. Dyn.*, 49, 71–95, <https://doi.org/10.1007/s00382-016-3327-9>, 2017.

Lucas-Picher, P., Argüeso, D., Brisson, E., Trambly, Y., Berg, P., Lemonsu, A., Kotlarski, S., and Caillaud, C.: Convection-permitting modeling with regional climate models: Latest developments and next steps, *WIREs Clim. Change*, 12, <https://doi.org/10.1002/wcc.731>, 2021.

870 McGill, R., Tukey, J. W., and Larsen, W. A.: Variations of Box Plots, *Am. Stat.*, 32, 12, <https://doi.org/10.2307/2683468>, 1978.

Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: Effects of different regional climate model resolution and forcing scales on projected hydrologic changes, *J. Hydrol.*, 541, 1003–1019, <https://doi.org/10.1016/j.jhydrol.2016.08.010>, 2016.

875 Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, *Water Resour. Res.*, 47, <https://doi.org/10.1029/2010WR009505>, 2011.

METEK: 2D Ultrasonic Anemometer, <https://www.th-friedrichs.de/en/products/wind/ultrasonic-anemometer/2d-ultrasonic-anemometer/>, last access: 30 October 2023.

Meyer, S. J., Hubbard, K. G., and Wilhite, D. A.: Estimating potential evapotranspiration: the effect of random and
880 systematic errors, *Agric. For. Meteorol.*, 46, 285–296, [https://doi.org/10.1016/0168-1923\(89\)90032-4](https://doi.org/10.1016/0168-1923(89)90032-4), 1989.

Miera, J. J., Schmidt, K., Suchodoletz, H. von, Ulrich, M., Werther, L., Zielhofer, C., Ettel, P., and Veit, U.: Large-scale investigations of Neolithic settlement dynamics in Central Germany based on machine learning analysis: A case study from the Weiße Elster river catchment, *PloS one*, 17, e0265835, <https://doi.org/10.1371/journal.pone.0265835>, 2022.

Monteith, J. L.: *Vegetation and the atmosphere*, 1, Academic Press, London, 1975.

- 885 Morrison, H., Peters, J. M., Chandrakar, K. K., and Sherwood, S. C.: Influences of Environmental Relative Humidity and Horizontal Scale of Subcloud Ascent on Deep Convective Initiation, *J. Atmos. Sci.*, 79, 337–359, <https://doi.org/10.1175/JAS-D-21-0056.1>, 2022.
- Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global
900 warming, *Sci. Rep.*, 9, 16063, <https://doi.org/10.1038/s41598-019-52277-4>, 2019.
- Oke, T. R.: *Boundary Layer Climates*, 2nd ed., Routledge, London and New York, 1987.
- Osuch, M., Romanowicz, R. J., and Booij, M. J.: The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics, *Hydrol. Sci. J.*, 60, 1299–1316, <https://doi.org/10.1080/02626667.2014.967694>, 2015.
- 895 Philipona, R.: Underestimation of solar global and diffuse radiation measured at Earth's surface, *J. Geophys. Res.*, 107, 4654, <https://doi.org/10.1029/2002JD002396>, 2002.
- Poncet, N., Lucas-Picher, P., Trambly, Y., Thirel, G., Vergara, H., Gourley, J., and Alias, A.: Does a convection-permitting regional climate model bring new perspectives on the projection of Mediterranean floods?, *Nat. Hazards Earth Syst. Sci.*, 24, 1163–1183, <https://doi.org/10.5194/nhess-24-1163-2024>, 2024.
- 900 Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E., Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., and Jacob, D.: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits?, *Clim. Dyn.*, 46, 383–412, <https://doi.org/10.1007/s00382-015-2589-y>, 2016.
- Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., and Leung, R.: A review on regional convection-permitting
905 climate modeling: Demonstrations, prospects, and challenges, *Rev. Geophys.*, 53, 323–361, <https://doi.org/10.1002/2014RG000475>, 2015.
- Prein, A. F., Rasmussen, R. M., Wang, D., and Giangrande, S. E.: Sensitivity of organized convective storms to model grid spacing in current and future climates, *Philos Trans A Math Phys Eng Sci*, 379, 20190546, <https://doi.org/10.1098/rsta.2019.0546>, 2021.
- 910 Quintero, F., Villarini, G., Prein, A. F., Krajewski, W. F., and Zhang, W.: On the role of atmospheric simulations horizontal grid spacing for flood modeling, *Clim. Dyn.*, 59, 3167–3174, <https://doi.org/10.1007/s00382-022-06233-0>, 2022.

- Raschendorfer, M.: The new turbulence parameterization of LM, in: COSMO News Letter No. 1, edited by: Doms, G. and Schättler, U., 25–28, 2008.
- Reszler, C., Switanek, M. B., and Truhetz, H.: Convection-permitting regional climate simulations for representing floods in small- and medium-sized catchments in the Eastern Alps, *Nat. Hazards Earth Syst. Sci.*, 18, 2653–2674, <https://doi.org/10.5194/nhess-18-2653-2018>, 2018.
- Richards, L. A.: Capillary conduction of liquids through porous mediums, *J. Appl. Phys.*, 1, 318–333, <https://doi.org/10.1063/1.1745010>, 1931.
- Rode, M., Klauer, B., Petry, D., Volk, M., Wenk, G., and Wagenschein, D.: Integrated nutrient transport modelling with respect to the implementation of the European WFD: The Weiße Elster Case Study, Germany, *Water SA*, 34, 490–496, <https://doi.org/10.4314/wsa.v34i4.183661>, 2008.
- Ross, C. A., Ali, G. A., Spence, C., and Courchesne, F.: Evaluating the Ubiquity of Thresholds in Rainfall-Runoff Response Across Contrasting Environments, *Water Resour. Res.*, 57, <https://doi.org/10.1029/2020WR027498>, 2021.
- Rudd, A. C., Kay, A. L., Wells, S. C., Aldridge, T., Cole, S. J., Kendon, E. J., and Stewart, E. J.: Investigating potential future changes in surface water flooding hazard and impact, *Hydrol. Process.*, 34, 139–149, <https://doi.org/10.1002/hyp.13572>, 2020.
- Sanchez, G., Serrano, A., Cancillo, M. L., and Garcia, J. A.: Pyranometer Thermal Offset: Measurement and Analysis, *J. Atmos. Oceanic Tech.*, 32, 234–246, <https://doi.org/10.1175/JTECH-D-14-00082.1>, 2015.
- Schaller, N., Sillmann, J., Müller, M., Haarsma, R., Hazeleger, W., Hegdahl, T. J., Kelder, T., van den Oord, G., Weerts, A., and Whan, K.: The role of spatial and temporal model resolution in a flood event storyline approach in western Norway, *Weather Clim. Extrem.*, 29, 100259, <https://doi.org/10.1016/j.wace.2020.100259>, 2020.
- Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpiilloz, C., Di Girolamo, S., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdüsühli, S., Schlemmer, L., Schulthess, T. C., Sprenger, M., Ubbiali, S., and Wernli, H.: Kilometer-Scale Climate Models: Prospects and Challenges, *Bull. Am. Meteorol. Soc.*, 101, E567–E587, <https://doi.org/10.1175/BAMS-D-18-0167.1>, 2020.
- Schulla, J.: Model Description WaSiM (Water balance Simulation Model), 2021.
- Schulz, E. W. and Sanderson, B. G.: Stationarity of turbulence in light winds during the maritime continent thunderstorm experiment, *Bound.-Lay. Meteorol.*, 111, 523–541, <https://doi.org/10.1023/B:BOUN.0000016546.42602.0a>, 2004.

- Scott, D. W.: Sturges' rule, *WIREs Computational Stats*, 1, 303–306, <https://doi.org/10.1002/wics.35>, 2009.
- 940 Seifert, A.: A revised cloud microphysical parameterization for COSMO-LME, in: *COSMO News Letter No. 7: Proceedings from the 8th COSMO General Meeting in Bucharest, 25–28, 2008*.
- Steppeler, J., Hess, R., Schättler, U., and Bonaventura, L.: Review of numerical methods for nonhydrostatic weather prediction models, *Meteor. Atmos. Phys.*, 82, 287–301, <https://doi.org/10.1007/s00703-001-0593-8>, 2003.
- Strandberg, G. and Lind, P.: The importance of model resolution on simulated precipitation in Europe – from global to
945 regional model, 2020.
- Sturges, H. A.: The Choice of Class Interval, *J. Am. Stat. Assoc.*, 21, 65–66, 1926.
- Tamm, O., Kokkonen, T., Warsta, L., Dubovik, M., and Koivusalo, H.: Modelling urban stormwater management changes using SWMM and convection-permitting climate simulations in cold areas, *J. Hydrol.*, 622, 129656, <https://doi.org/10.1016/j.jhydrol.2023.129656>, 2023.
- 950 Tang, W. and Carey, S. K.: HydRun: A MATLAB toolbox for rainfall–runoff analysis, *Hydrol. Process.*, 31, 2670–2682, <https://doi.org/10.1002/hyp.11185>, 2017.
- Taylor, C. M., Birch, C. E., Parker, D. J., Dixon, N., Guichard, F., Nikulin, G., and Lister, G. M. S.: Modeling soil moisture-precipitation feedback in the Sahel: Importance of spatial scale versus convective parameterization, *Geophys. Res. Lett.*, 40, 6213–6218, <https://doi.org/10.1002/2013GL058511>, 2013.
- 955 Tiedtke, M.: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, *Mon. Weather Rev.*, 117, 1779–1800, [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2), 1989.
- Tölle, M. H., Schefczyk, L., and Gutjahr, O.: Scale dependency of regional climate modeling of current and future climate extremes in Germany, *Theor. Appl. Climatol.*, 134, 829–848, <https://doi.org/10.1007/s00704-017-2303-6>, 2018.
- USGS: Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global, <https://doi.org/10.5066/F7PR7TFT>, 2018.
- 960 van Pham, T., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., Rieger, D., Zängl, G., and Früh, B.: ICON in Climate Limited-area Mode (ICON release version 2.6.1): a new regional climate model, *Geosci. Model Dev.*, 14, 985–1005, <https://doi.org/10.5194/gmd-14-985-2021>, 2021.

Weigl, E. and Winterrath, T.: Radargestützte Niederschlagsanalyse und –vorhersage (RADOLAN, RADVOR-OP), *promet*, 35, 78–86, 2010.

- 965 Winterrath, T., Rosenow, W., and Weigl, E.: On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management, in: *Weather radar and hydrology proceedings of a symposium held in Exeter, UK, April 2011*, 323–329, 2012.

WMO: *WMO Guidelines on the Calculation of Climate Normals*, 2017.

- Xie, K., Li, L., Chen, H., and Xu, C.-Y.: Assessing the performance of convection-permitting climate model in reproducing basin-scale hydrological extremes: A western Norway case study, *J. Hydrol.*, 656, 132989, <https://doi.org/10.1016/j.jhydrol.2025.132989>, 2025.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M Description of the non-hydrostatic dynamical core, *Quart J Royal Meteorol Soc*, 141, 563–579, <https://doi.org/10.1002/qj.2378>, 2015.

- 975 Zorn, A., Tafel, J., and Schäfer, S.: Klimapolitisches Engagement auf kommunaler Ebene in Deutschland: Entwicklung eines Index zu Bewusstsein, Konzeption und Durchführung von Klimaschutz und Klimawandelanpassung, *Raumforschung und Raumordnung*, <https://doi.org/10.14512/rur.2225>, 2024.