https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

Heat wave characteristics: evaluation of regional climate model performances for Germany

Dragan Petrovic[1], Benjamin Fersch[1] and Harald Kunstmann[1,2]

[1]Institute of Meteorology and Climate Research (IMK-IFU), Karlsruhe Institute of Technology, Campus Alpin, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany
[2]Institute of Geography and Center for Climate Resilience, University of Augsburg, Alter Postweg 118, 86159 Augsburg, Germany

Correspondence to: Dragan Petrovic (dragan.petrovic@kit.edu)

Abstract. Heat waves are among the most severe climate extreme events. In this study, we address the impact of increased model resolution and tailored model settings on the reproduction of these events by evaluating different regional climate model outputs for Germany and the near surroundings between 1980–2009. Both, outputs of an ensemble of six EURO-CORDEX models of 12.5 km grid resolution and outputs from a high resolution (5 km) WRF run are employed. The latter was especially tailored for the study region regarding the physics configuration. We analyze the reproduction of maximum temperature, number of heat wave days, heat wave characteristics (frequency, duration and intensity), the 2003 major event and trends in the annual number of heat waves. E-OBS is used as reference and we imply Taylor diagram, Mann-Kendall trend test, spatial efficiency metric and cumulative heat index as a measure for intensity. Averaged over the domain, heat waves occurred about 31 times in the study period with an average duration of 4 days and average heat excess of 10 °C. Maximum temperature was reproduced reasonably well by all models. Despite the same forcing, the models exhibited a large spread in the heat wave reproduction. The domain mean conditions of heat wave frequency and duration were captured reasonably well, but intensity was reproduced weakly. The spread was particularly pronounced for the 2003 event, indicating the difficulty of models to reproduce single major events. All models underestimated the spatial extent of the observed increasing trends. WRF mostly did not perform significantly better than the other models. We conclude that increased model resolution does not add a significant value to heat wave simulation if the base resolution is already relatively high. Tailored model settings seem to play a minor role. The partly distinct differences in performance, however, highlight that the choice of model can be crucial.

1 Introduction

Heat waves are climatological extreme events with severe negative impacts on organisms and ecosystems. Their effects can be illness, large-scale mortality, substantial losses in agricultural production, forest fires and increased energy demand for cooling (Beniston et al., 2007; De Bono et al. 2004; Ciais et al., 2005; Robine et al., 2008; Kyselý et al., 2011; Bastos et al., 2014; Urban et al., 2017). Especially in the midlatitude zones, heat waves are regarded as a major cause of weather-related

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

human mortalities (Luber, 2008; Plavcová and Kyselý, 2019). In Europe, they are a regular part of the summer climate (Vautard et al., 2013), and an increasing trend in the number of summer heat waves was reported for the recent decades with prolonging tendency (Della-Marta et al., 2007; Kyselý, 2010; Valeriánová et al., 2015; Saeed et al., 2017). According to Fischer and Schär (2010), the frequency and length of heat waves have nearly tripled and doubled, respectively, over the period 1880 – 2005 in Western European regions. Especially since the millennium, Europe experienced multiple extraordinary heat waves (Lhotka et al., 2018b). Two of the most prominent events were the heat episodes in 2003 over Western Europe (Fink et al., 2004) and the 2010 event over Eastern Europe and Russia (Schneidereit et al., 2012). In the more recent past, severe summer heat episodes took place in Central Europe in 2013 (Lhotka and Kyselý, 2015b), 2015 (Hoy et al., 2016) and especially in 2018 (e.g., Vogel et al., 2019, Rousi et al., 2022). In the context of climate change, heat waves are expected to become more frequent, more intense and longer lasting in the future around the globe (Meehl and Tebaldi 2004; Lau and Nath 2014; Lemonsu et al. 2014; Seneviratne et al., 2014; Diffenbaugh and Ashfaq 2010). Coumou and Rahmstorf (2012) estimate the probability of severe events like 2003 over Western Europe has increased by a factor of 2 – 4 because of global warming. Shifts of the temperature distribution are considered as the primary drivers of these changes (Ballester et al., 2010; Lau and Nath, 2014). According to Fischer and Schär (2010), larger variance of summer temperature distribution in future climates are also possible drivers. Lhotka et al. (2017) additionally emphasize increased temporal autocorrelation of daily maximum temperature, which leads to more persistent heat waves, as potential drivers.

To analyze and understand such changes at the regional scale and to allow projections for future characteristics of the heat wave events, regional climate models (RCMs) are used. For better interpretation of future climate scenarios and their uncertainties and limitations, it is important to evaluate RCM simulations for the recent or historical climate to detect biases, which are usually present, despite the added value compared to GCMs (Lhotka et al., 2018a; Plavcová and Kyselý, 2019; Lin et al., 2022). The availability and reliability of RCM simulations have rapidly evolved in the last years (Štepánek et al., 2016). This is also due to concerted downscaling projects and initiatives such as PRUDENCE (Christensen and Christensen, 2007), ENSEMBLES (van der Linden and Mitchell, 2009) and most recently CORDEX (Giorgi et al., 2009). In the recent past there have been several heat wave related studies using data from the CORDEX initiative for different parts of the world with focus on both, past periods and future scenarios. For the EURO-CORDEX domain, a focus was put on the evaluation of the RCM's capability in simulating past heat episodes in France (Ouzeau et al., 2016) and all of Europe (Vautard et al., 2013; Lhotka et al., 2018a; Plavcová and Kyselý, 2019; Lin et al., 2022) as well as on the development of future episodes under different scenarios for Portugal (Cardoso et al., 2019), France (Ouzeau et al., 2016), the Mediterranean area (Molina et al., 2020) and all of Europe (Lhotka et al., 2018b; Smid et al., 2019; Machard et al., 2020; Lin et al., 2022). Regarding the rest of the globe and the other CORDEX domains, evaluation studies have been carried out for Africa (Russo et al., 2016), East Asia (Wang et al., 2019b) and South America (Silva et al., 2022). Furthermore, projection studies were performed for Africa (Dosio, 2017), Afghanistan (Aich et al., 2017), South America (Feron et al., 2019), China (Wang et al., 2019a), the MENA region (Varela et al., 2020), the Eastern Mediterranean (Wedler et al., 2023), East Asia (Kim et al., 2023) as well as for the entire globe (Coppola et al., 2021). In the mentioned studies, different horizontal grid resolutions of the models were used and the effects of increased

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

resolution were often analyzed, which led to different findings: Zeng et al. (2016) and Vichot-Llano et al. (2021), for example, found that higher resolution leads to better reproduction of temperature fields, while Di Luca et al. (2013) came to the conclusion that the potential for added value of increased resolution is small. It is important to consider the differences between the compared resolutions in such studies (Petrovic et al., 2022). Besides the model resolution, the model setup regarding the domain configuration and physical parameterizations for the selected target region is a crucial factor for reliable simulations (e.g., Stoelinga et al., 2003; Kumar et al., 2010). For the temperature simulation, Vautard et al. (2013) found that it is primarily sensitive to convection and microphysics schemes. They emphasize that a large part of the model spread in their study can be attributed to different parameterizations. Moreover, they draw a connection between parameterizations and different spatial resolutions. Mooney et al. (2013) found that simulated temperature showed relatively high sensitivity to the land surface models, some sensitivity to the radiation schemes and little sensitivity to the microphysics and planetary boundary layer (PBL) schemes. They concluded a strong dependence on the region and season of the optimal parameterization combination. Kotlarski et al. (2014) state that bias spreads between different configurations of the same model can be similar to those between different models.

To our knowledge, there is no study that presents an evaluation of the EURO-CORDEX RCM's capability to reproduce heat wave characteristics for Germany, which has motivated us to perform this analysis. At the same time this study is the follow-up to the model comparison study for droughts from Petrovic et al. (2022). The thematic proximity is obvious, since heat and drought are often, but not always, related to each other. We analyze a variety of RCM simulations, i.e., a 5 km three-domain Weather Research and Forecasting Model (WRF, Skamarock et al., 2008) run and an ensemble of six EURO-CORDEX realizations at 12.5 km horizontal resolution. The setup for the WRF model was thoroughly determined for Germany. Intuitively, one would expect better performance in simulating hot temperature from WRF due to the higher resolution and the focus on the target region compared to the EURO-CORDEX runs. The WRF model was shown to be capable of simulating spatiotemporal features of heat wave events over a large domain (Wang et al., 2019b). To attribute to a resolution or settings effect of the WRF model performance, we additionally include a 15 km domain configuration in our analysis which is slightly coarser than for the EURO-CORDEX standard. Therefore, following up on the drought study (Petrovic et al., 2022), the objectives of this study are:

1. Evaluation of the performance of regional climate models in reproducing hot temperatures and associated heat wave characteristics by employing a six-member EURO-CORDEX ensemble and a high resolution WRF run. The EURO-CORDEX RCMs and WRF differ in resolution, while the model physics configurations differ among every single RCM.
2. Obtaining insights into the heat wave course for Germany and its near surrounding between 1980 and 2009.

For this purpose, the results are evaluated and compared to observations. Specifically, we analyze reproduction of daily maximum temperature, the number of heat wave days, heat wave characteristics (frequency, duration and intensity), trends in the number of heat waves per year and the heat wave event in 2003.

Moreover, we can compare the new core findings for heat waves with those from the aforementioned drought analysis (Petrovic et al., 2022) in terms of similarities and differences.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

2 Data

In this study, data from the same sources as in Petrovic et al. (2022) is used. While in that study monthly data of precipitation, minimum and maximum temperature was used, here the daily values of the maximum (surface) temperature (T_{max}) are employed.

We use data from an ensemble of six EURO-CORDEX RCM simulations. Each of the experiments were conducted with 0.11° (≈ 12.5 km) horizontal grid resolution and cover the EUR-11 CORDEX-Domain. Outputs from the following RCMs were used: COSMO-CLM, ALADIN 6.3 (hereafter referred to as ALADIN in the text), REMO2015 (REMO), RegCM 4.6 (RegCM), RACMO 2.2e (RACMO) and RCA4. All the runs obtained the boundary conditions from the global ERA-Interim reanalysis (Dee et al, 2011). When this analysis was initiated, these runs were the only ones that cover the study period 1980–2009 and contain the relevant meteorological variables.

In addition to the EURO-CORDEX data, we included the outputs from the ERA-Interim forced reanalysis WRF run for the time period 1980 – 2009 from Warscher et al. (2019). These simulations were preceded by a comprehensive search and final identification of optimal model physics and parameterization configuration for the target region Germany (Wagner and Kunstmann, 2016). This is the first major difference to the EURO-CORDEX outputs, which were aimed at the entire EUR-11 CORDEX-domain (Giorgi et al., 2009). The horizontal grid resolution of the innermost domain, which frames Germany and the near surroundings, is 5 km. This increased resolution is the second major difference to the EURO-CORDEX outputs. As mentioned above, we also used the outputs from the 15 km second WRF domain of the same run. Therefore, we will refer to WRF@5 km and WRF@15 km from here on to distinguish between the two runs.

More detailed information about the EURO-CORDEX RCMs and an overview of the different model physics configurations for all runs can be obtained from Table 1 and 2 in Petrovic et al. (2022). The gridded observational data set from E-OBS (Haylock et al., 2008), version 23.1e, in 0.1° (≈ 11.1 km) horizontal grid resolution serves as reference.

The study region extends from 6° to 15° E and 47° to 55° N, so that it contains Germany and its near surroundings. The WRF and E-OBS data were regridded using bilinear interpolation to match the horizontal grid resolution of the EURO-CORDEX RCMs.

3 Methods

3.1 Analysis of daily maximum temperature reproduction

Since T_{max} is the main variable determining heat waves, the grid cell based summer values (June, July, August) are first analyzed. We use the Taylor diagram (Taylor, 2001), which provides a succinct visual statistical summary in terms of agreement between patterns regarding their correlation, their root-mean-square difference, and the ratio of their variances or standard deviations (Taylor, 2001). Moreover, we calculate the density plot to visualize and compare the distributions of the values of the individual data sets. From spatially and temporally averaged daily values, we calculate the mean bias values to

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

be able to draw conclusions about the role of the model bias for the further results. It must be noted that not only a good simulation of the right tail of the temperature frequency distribution is of importance for the reproduction of heat waves, but also the persistence of the high temperatures (Lhotka et al., 2018a).

3.2 Heat wave definition

There is no universal definition of heat waves. In fact, there are multiple definitions that include different metrics and criteria depending on the region, season and purpose of the study (Feron et al., 2019; Becker et al., 2022). Generally, it describes a period of consecutive days with conditions excessively higher than normal (Perkins et al., 2012). Here, we define a heat wave as an event of at least three consecutive days where the 90th percentile of T_{max} based on each calendar day of the study period is exceeded (Fischer and Schär, 2010). Therefore, the 90th percentile for each calendar day and each grid cell from each data set was calculated first. This was done for each dataset individually to circumvent the T_{max} biases among the different models (Vautard et al., 2013; Lhotka et al., 2018b). We only addressed summer heat waves in this study.

3.3 Analysis of heat wave characteristics

Based on the heat wave definition described above, we calculate the number of heat wave days and number of heat waves for each grid cell for the whole study period 1980 – 2009. Based on the number of heat wave days and heat waves, we determine the mean duration of heat waves for each grid cell from each data set. In order to describe the mean heat wave intensity, we use the cumulative heat index as a measure (e.g. Katavoutas and Founda, 2019; Perkins-Kirkpatrick and Lewis, 2020; Founda et al., 2022). It refers to the integration of heat exceedance over the 90th percentile threshold for all heat wave days during a heat episode or whole season:

$$CumHeat = \sum_{i=1}^{n} \Delta(T_{max,i} - T_{max,P90,i}) \quad (1)$$

where i indicates the calendar day of the heat wave event and $T_{max,P90,i}$ the 90th T_{max} percentile of day i, for each grid cell. To get the mean intensity of heat waves, we integrated all excess values for the whole study period and divided the results by the number of heat waves for each grid cell of each data set. For each of the single aspects (number of heat wave days and heat waves, mean heat wave duration and mean heat wave intensity) we calculate the domain mean value for every data set. In the next step, we subtract the E-OBS reference domain from each RCM domain to get the bias patterns and also calculated the domain mean values. To further evaluate the spatial agreement between reference and each RCM, we utilize the spatial efficiency (SPAEF) metric (Demirel et al., 2018; Koch et al., 2018). The SPAEF is a multiple components performance metric for the comparison of spatial patterns. It is calculated as:

$$SPAEF = 1 - \sqrt{(\alpha - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \quad (2)$$

with the three components α as the Pearson correlation coefficient between observed (obs) and simulated (sim) patterns,

$$\beta = \left(\frac{\sigma_{sim}}{\mu_{sim}}\right) / \left(\frac{\sigma_{obs}}{\mu_{obs}}\right) \quad (3)$$

as the fraction of coefficient of variation representing spatial variability and

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences
Discussions

$$\gamma = \frac{\sum_{i=1}^{n} \min(K_i, L_i)}{\sum_{i=1}^{n} K_i} \quad (4)$$

as the overlap between the histograms of the observed (*K*) and simulated patterns (*L*), both containing the same number *n* of bins. For the calculation of γ the z score of the patterns is used, which allows comparison of two variables with different units. For both histograms of *K* and *L*, the number of values in each bin *i* is counted. Afterwards for each bin the lower (minimum) number of K_i or L_i is picked, which indicates the number of shared values in the same bin. Thereafter these numbers are summed up and divided by the total number *n* of values in *K* or *L*. The SPAEF has a predefined range between -∞ and 1, where 1 corresponds to ideal agreement between two patterns. The three components are independent of each other and typically equally weighted so that they complement each other in a meaningful way and provide holistic pattern information. By this way, instead of exact values on the grid scale, global features such as distribution and variability are evaluated (Koch et al., 2018).

3.4 Heat wave trend analysis

In order to investigate the temporal characteristics of heat waves occurrences, we apply the non-parametric Mann-Kendall trend test approach (Mann, 1945; Kendall, 1975). For this purpose, we first count the number of heat waves per year for each grid cell to obtain the annual development. Then on the resulting time series of each grid cell the test is applied to detect significant monotonic trends at a significance level of 0.05. The Mann-Kendall test is based on the correlation between the ranks of a time series and their time order.

For a time series $x_1, x_2, x_3 \ldots x_n$, the Mann-Kendall test statistic *S* is given by

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(x_j - x_i) \quad (5)$$

with

$$sign(x_j - x_i) = sign(R_j - R_i) = 1 \text{ if } x_j - x_i > 0 \quad (6)$$

$$sign(x_j - x_i) = sign(R_j - R_i) = 0 \text{ if } x_j - x_i = 0 \quad (7)$$

$$sign(x_j - x_i) = sign(R_j - R_i) = -1 \text{ if } x_j - x_i < 0 \quad (8)$$

where *sign* represents an indicator function, *n* the number of data points and R_i and R_j their respective ranks. A positive *S*-statistic indicates an increasing trend, a negative one indicates a decreasing trend.

3.5 The 2003 heat wave event

The heat wave and drought event in the summer months of 2003 in Central Europe is considered as one of the most severe extreme events in the last decades. It has caused 70,000 excess deaths (Poumadere et al. 2005; Robine et al., 2008), distinct decrease of plant productivity and crop failures (Bastos et al., 2014) and record breaking loss of Alpine glaciers' mass (De Bono et al., 2004). This is why the event is also considered a 'mega heat wave' (Barriopedro et al. 2011; Vautard et al., 2013).

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

We investigate the capability of the RCMs to reproduce such a single extreme event in terms of intensity and maximum duration. Therefore, we calculate the cumulative heat for the whole summer period 2003 and determined the maximum duration of the heat episodes during this time for each grid cell. The results in this section are also evaluated based on the domain mean values, mean bias values and the SPAEF.

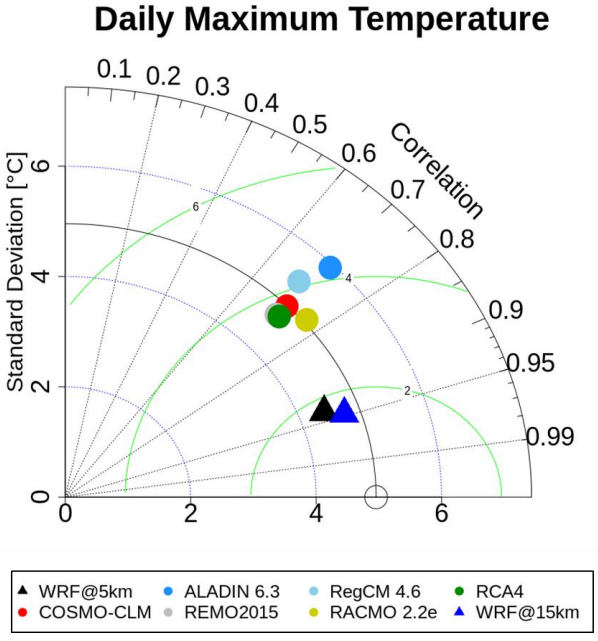
4. Results

4.1 Daily maximum temperature reproduction

Figure 1 shows the Taylor diagram of the grid cell based T_{max} values from each data set for the summer months. There is an obvious difference between the EURO-CORDEX RCMs and the two WRF runs. In terms of correlation with the E-OBS reference, the two WRF runs stand out with values above 0.9. The 15 km WRF run shows a slightly higher value than its 5 km counterpart. The EURO-CORDEX RCMs reach values between 0.65 and 0.8. RACMO holds the highest value, RegCM the lowest. Regarding the root mean square error (RMSE), all RCMs have a value below 5. Again, there is a significant difference between the WRF runs and the EURO-CORDEX RCMs, since the WRF runs have distinctly lower values below 2. The 15 km run is slightly better than its 5 km counterpart. The values of the EURO-CORDEX RCMs are relatively close to each other. RACMO holds the lowest RMSE value, and ALADIN the highest. As far as the agreement of the standard deviation with the reference is concerned, there is a large discrepancy among the EURO-CORDEX RCMs. RACMO comes closest to the reference, ALADIN's standard deviation shows the biggest difference. RACMO's standard deviation shows even a greater match than the WRF runs, the same goes for COSMO-CLM. Once again, the WRF@15 km run is closer to the reference than its 5 km counterpart. This underlines that the temporal variability is better captured in the coarser resolutions. The results suggest that the WRF model settings are determining for the better performance compared to the EURO-CORDEX RCMs. RACMO is the best performing EURO-CORDEX RCM, ALADIN the worst.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions



215 **Figure 1.** Taylor diagram comparing the model performances in reproducing the daily summer T_{max} values in relation to the E-OBS reference for the study period 1980–2009 and the whole study area.

Figure 2 displays the density plot of the summer T_{max} values from each data set. There are pronounced differences between the single distributions in general, but also at the right tail, which is in focus here. Until approx. 10 °C all distributions are relatively similar, the discrepancies begin hereafter. Compared to E-OBS, RCA4 and RACMO are clearly shifted leftwards.
220 Apart from these two models, all other data sets have most of their values in the range between 20 and 25 °C. ALADIN and RegCM are shifted towards the right compared to E-OBS, especially in the right tail area from approx. 30 °C on, in which they clearly have more values than all other runs and E-OBS. In this area, REMO shows high agreement with E-OBS, while WRF@5 km and RCA4 have the least amount of values here. The overall differences between the two WRF runs are distinct. The 15 km run is closer to E-OBS than its 5 km counterpart and it has more values in the right tail of its distribution.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

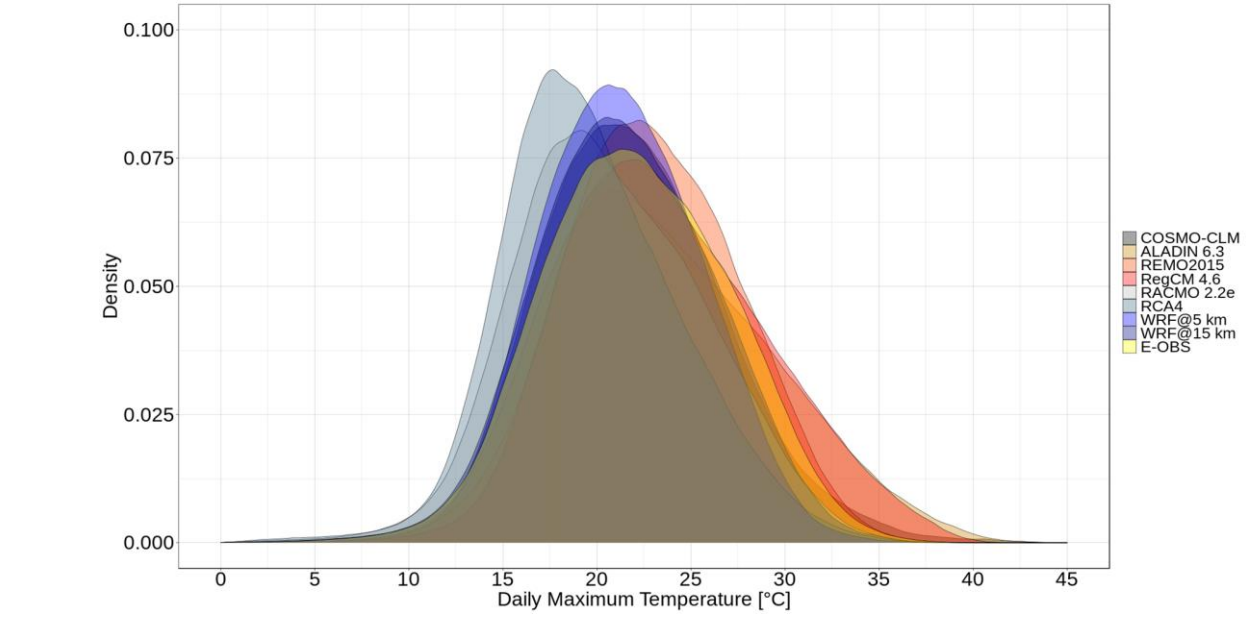


Figure 2. Density plot of the summer T_{max} values from each data set.

Table 1 gives the bias values of the spatially and temporally averaged maximum temperature values for each RCM. Five runs, COSMO-CLM, RACMO, RCA4 and the two WRF runs, show cold mean bias values, the other runs warm ones. The highest negative value is found at RCA4 (-2.40 °C), which is also the highest overall value, followed by RACMO (-1.37 °C). This could be inferred by the density plot in Figure 2. This is also true for the highest positive bias values of ALADIN (1.19 °C) and RegCM (1.62°C). COSMO-CLM (-0.16 °C) holds the lowest bias value. The overall spread is 4.02 °C between RegCM and RCA4.

Table 1. Spatially and temporally averaged daily T_{max} bias values with respect to E-OBS.

Model	T_{max} Bias [°C]
COSMO-CLM	-0.16
ALADIN 6.3	1.19
REMO2015	0.70
RegCM 4.6	1.62
RACMO 2.2e	-1.37
RCA4	-2.40
WRF@5 km	-0.93
WRF@15 km	-0.58

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

4.2 Number of heat wave days

Figure 3 presents the E-OBS pattern of the number of heat wave days for the time period 1980 – 2009 along with the grid cell based differences between each RCM and E-OBS. Table 2 provides more detailed information.

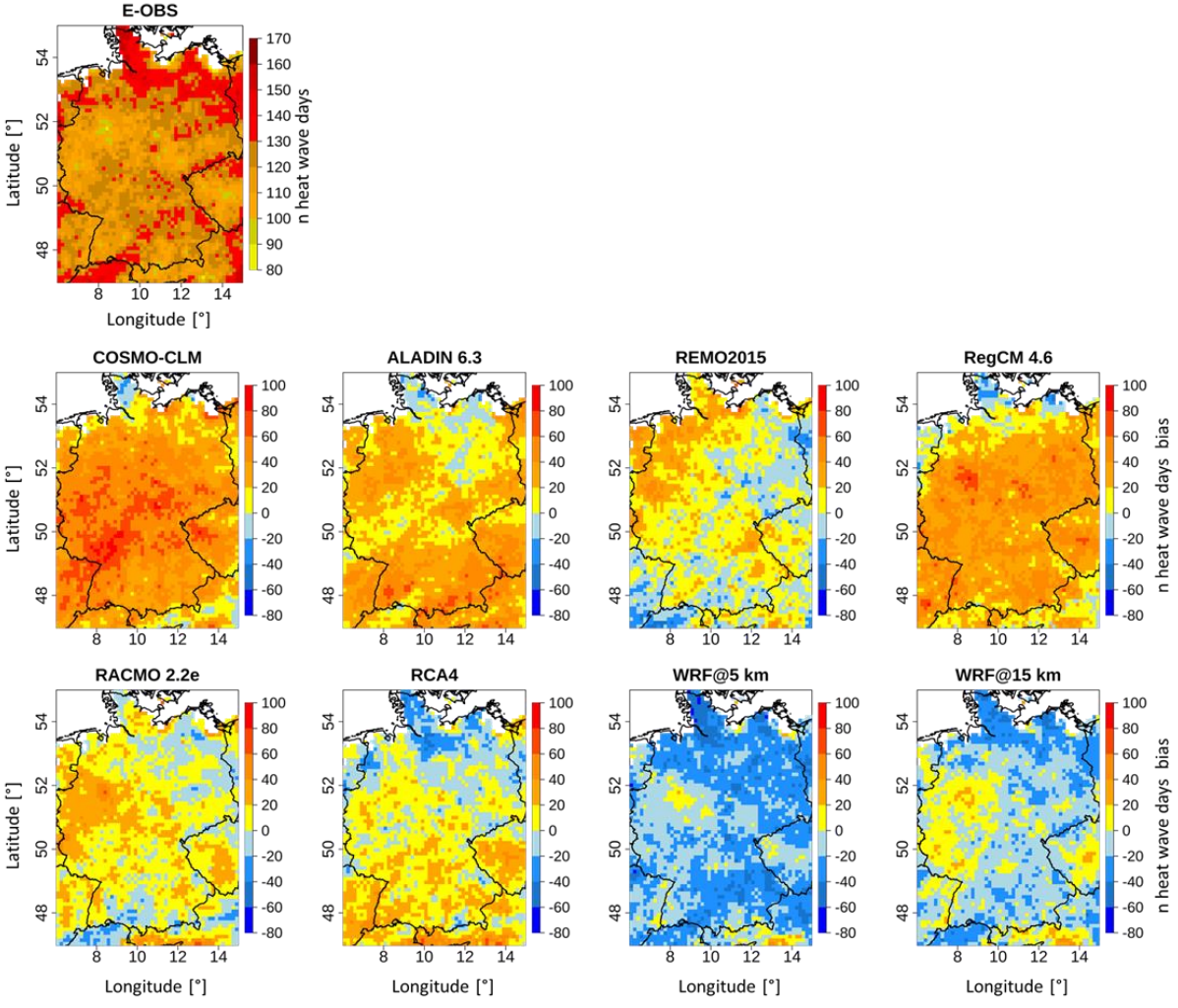


Figure 3. Grid cell based E-OBS number of heat wave days pattern for the summer months between 1980–2009 and differences between each RCM and E-OBS.

In the E-OBS domain the highest values are located in the northern, northeastern and southwestern parts with up to 160 heat wave days. The minimum values range between 80 and 90 days and are distributed punctually all over the domain. There is no clear area characterized by low values. The domain mean value is 122 days (Table 2). The RCM difference patterns show distinct differences among each other. It is noticeable that some of the domains either largely show a “colder” bias (WRF@5 km), which means that they simulated less heat wave days compared to the reference, or a “warm” bias (COSMO-CLM, ALADIN and RegCM), meaning that they simulated more heat wave days. REMO, RACMO and RCA4 have a rather mixed

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

domain. It is striking that the two WRF domains are the only ones predominated by colder values. The majority of the values across all domains range between -60 to 60 days of difference. COSMO-CLM, ALADIN, RegCM and RACMO have relatively similar bias values in the western parts of the domain. Other than that, there are no repeating patterns across a majority of the RCMs. In the WRF@15 km domain there are significantly more warm bias areas compared to its 5 km counterpart, which is also confirmed in Table 2, where the mean bias value of the 15 km run is much closer to 0 (-8.7 compared to -20.9 days). These warm bias areas are mainly located in the western, eastern and southeastern parts of the domain. The values in Table 2 confirm the impressions from Figure 3: the domain mean values from all EURO-CORDEX RCMs are above the E-OBS reference value (122 days), with the COSMO-CLM value showing the biggest difference (42 days). The two WRF runs are below the reference (102 and 114 days). The values of REMO and RCA4 (both 130 days) and WRF@15 km come closest to the reference. Regarding the mean bias values, the inferred negative values from the two WRF runs are visible, while the EURO-CORDEX RCMs all show positive mean bias values. COSMO-CLM has the by far the highest bias value (41.8 days), REMO shows the lowest value (7.3 days). It should be kept in mind here that for RCMs that are not dominated by one bias direction, the values can cancel each other out and provide a small overall mean bias. This is the case for REMO, RACMO and RCA4. The SPAEF values give information about the pattern agreement between the reference and the individual RCMs (not shown here). There is not a single high value. The values are either negative or very low, meaning that there is no good overall spatial agreement from any RCM with the reference. REMO has the highest value (0.19), WRF@ 15 km the lowest (-0.19).

Table 2. Number of heat wave days metrics.

Model	Mean [n days]	Mean Bias [n days]	SPAEF
COSMO-CLM	164	41.8	-0.09
ALADIN 6.3	149	26.2	0.03
REMO2015	130	7.3	0.19
RegCM 4.6	153	31	-0.15
RACMO 2.2e	132	9.4	0.07
RCA4	130	7.8	0.01
WRF@5 km	102	-20.9	-0.11
WRF@15 km	114	-8.7	-0.19
E-OBS	122		

There are no apparent benefits of the WRF runs visible compared to EURO-CORDEX RCMs. This suggests that neither the increased grid resolution nor the model setup have a decisive effect on the reproduction on the number of heat wave days. In fact, the WRF@15 km performed better than its 5 km counterpart, which further underlines that the grid resolution might play

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

a less important role for this aspect. REMO is the RCM with the overall best performance due to the best values in all regards (Table 2), COSMO-CLM performed worst.

4.3 Heat wave characteristics

4.3.1 Heat wave frequency

Figure 4 shows the E-OBS pattern of the number of heat waves for the time period 1980 – 2009 and the grid cell based differences with the RCMs. The relevant scores are given in Table 3.

Table 3. Heat wave frequency metrics.

Model	Mean [n heat waves]	Mean Bias [n heat waves]	SPAEF
COSMO-CLM	29.7	-1.16	-0.14
ALADIN 6.3	31.2	0.31	-0.12
REMO2015	30.5	-0.33	-0.04
RegCM 4.6	33.4	2.57	-0.24
RACMO 2.2e	31.2	0.38	-0.22
RCA4	29.2	-1.71	-0.25
WRF@5 km	26.7	-4.14	-0.25
WRF@15 km	29.5	-1.37	-0.25
E-OBS	30.9		

The E-OBS domain looks mostly uniform with the majority of values ranging between 26 – 34 heat waves. The domain mean value (30.9) in Table 3 underlines this. Only in the north and south there are small concentrations of higher values up to 40 heat waves. The RCM bias domains show rather mixed patterns with both, positive and negative values. RegCM (positive) and WRF@5 km (negative) are the only domains where one direction in bias predominates. This is also reflected in the mean bias values in Table 3, where these two RCMs have the highest values (2.57 and -4.14), while at the other domains the opposite signs tend to cancel each other out, bringing them closer to zero on average. The northern part of all bias domains is predominated by negative bias values. It is also noticeable that in the eastern part of the domain only positive values prevail in RegCM, while negative values prevail in all other bias domains in this area. The WRF@15 km domain looks quite balanced and therefore quite different from its 5 km counterpart. This is also confirmed by the smaller mean bias value (-1.37) in Table 3. The domain mean values in Table 3 show a relatively large range of about seven heat waves between the maximum (33.4 at RegCM) and minimum (26.7 at WRF@5 km) value. The E-OBS reference value of 30.9 means that in average there was approx. one summer heat wave per year in the study period. ALADIN and RACMO (31.2) come closest to this value, WRF@5 km (26.7) shows the biggest discrepancy. COSMO-CLM, REMO, RCA4 and the two WRF runs simulated in average less heat

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

waves than the reference, ALADIN, RegCM and RACMO more. ALADIN has the lowest mean bias value (0.31). The mean bias values of COSMO-CLM, REMO, RCA4 and the two WRF runs are negative, the others are positive. All the SPAEF scores between the reference and the single RCM domains (not shown) are negative here, indicating that there is no good overall spatial agreement at all. The highest score (-0.04) is found at REMO, the lowest (-0.25) at RCA4 and the two WRF runs.

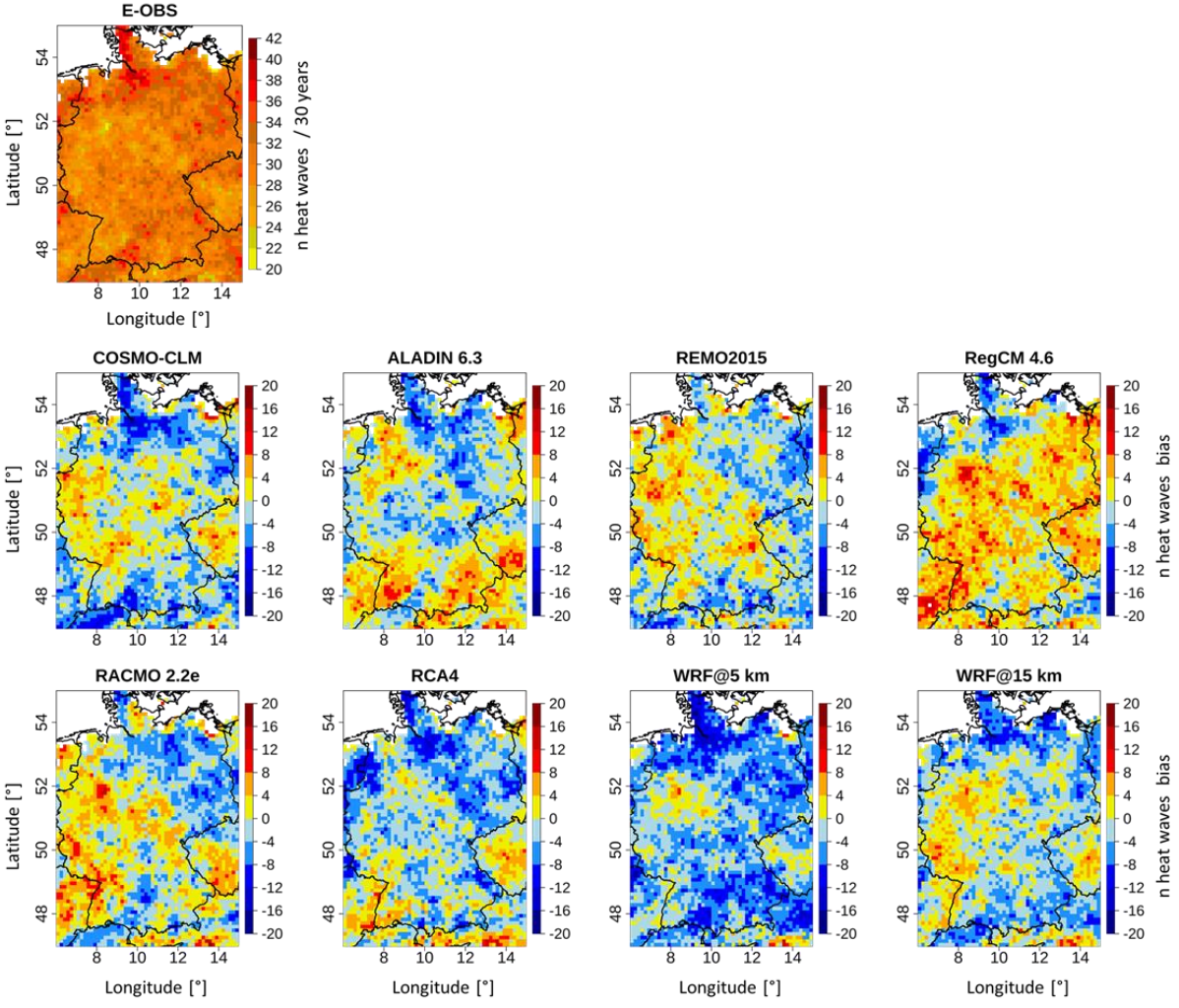


Figure 4. Grid cell based E-OBS summer heat wave frequency pattern between 1980–2009 and differences between each RCM and E-OBS.

There are no recognizable benefits of the two WRF runs either. Here, it is rather the opposite, especially regarding the WRF@5 km run, since it shows the highest mean bias value and the biggest difference to the domain mean value in the number of heat waves compared to the reference. In addition, it has the lowest SPAEF value, which is not very meaningful in this case. The

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

305 model settings seem to have a clearly higher importance than the grid resolution. ALADIN showed the best performance in this section, closely followed by REMO.

4.3.2 Mean heat wave duration

Figure 5 displays the E-OBS pattern of the mean heat wave durations for the time period 1980 – 2009 and the grid cell based differences with the RCMs. The associated scores are shown in Table 4.

310

Table 4. Mean heat wave duration metrics.

Model	Mean [n days]	Mean Bias [n days]	SPAEF
COSMO-CLM	5.46	1.53	-0.65
ALADIN 6.3	4.66	0.72	-0.07
REMO2015	4.17	0.23	-0.13
RegCM 4.6	4.50	0.56	-0.38
RACMO 2.2e	4.14	0.20	-0.21
RCA4	4.37	0.43	-0.24
WRF@5 km	3.78	-0.16	0.06
WRF@15 km	3.84	-0.10	-0.15
E-OBS	3.94		

The E-OBS pattern is very uniform, the majority of the domain is covered by values between 3.75 and 4.25 days. This is also reflected in the domain mean value of 3.94 days (Table 4). This means that the average heat wave duration was quite close to 315 the minimum length (3 days) of a heat wave. All EURO-CORDEX RCM bias domains except REMO and RACMO are prevailed by positive bias values, meaning that the models simulated too long heat episodes. The COSMO-CLM domain even appears to be without any negative bias grid cell. It is also the domain with the highest values. Especially the area in the southwest, where values up to 5 days are reached, is striking. All other bias domains are predominated by values between -1 and 1 day. The two WRF domains are the only ones prevailed by negative bias values. In this case, the WRF@15 km domain 320 is quite close to its 5 km counterpart, but again the negative bias is less pronounced in direct comparison. The patterns of REMO and RACMO are similar. The domain mean values in Table 4 are all close to each other. All EURO-CORDEX RCMs are above the reference value (3.94 days), the WRF runs are below. COSMO-CLM shows the biggest difference (1.52 days), WRF@15 km the smallest (0.1 days). COSMO-CLM is also the RCM with by far the highest mean bias value (1.53 days), which was to be expected from the bias maps (Figure 5). It is the only case where the mean bias is greater than 1 day. The bias 325 value of 1.53 days may seem relatively small, but if it is set in relation to the domain mean values, then it accounts for a fairly large proportion. Only the two WRF runs have negative mean bias values. WRF@15 km holds the smallest (-0.10) mean bias

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

value. Here, the SPAEF values between the reference and the single RCM domains (not shown) are all negative or very low, which is the case for WRF@5 km (0.06). This means that again no RCM was able to satisfactorily reproduce the spatial pattern of the reference. COSMO-CLM holds also in this regard the worst value (-0.65).

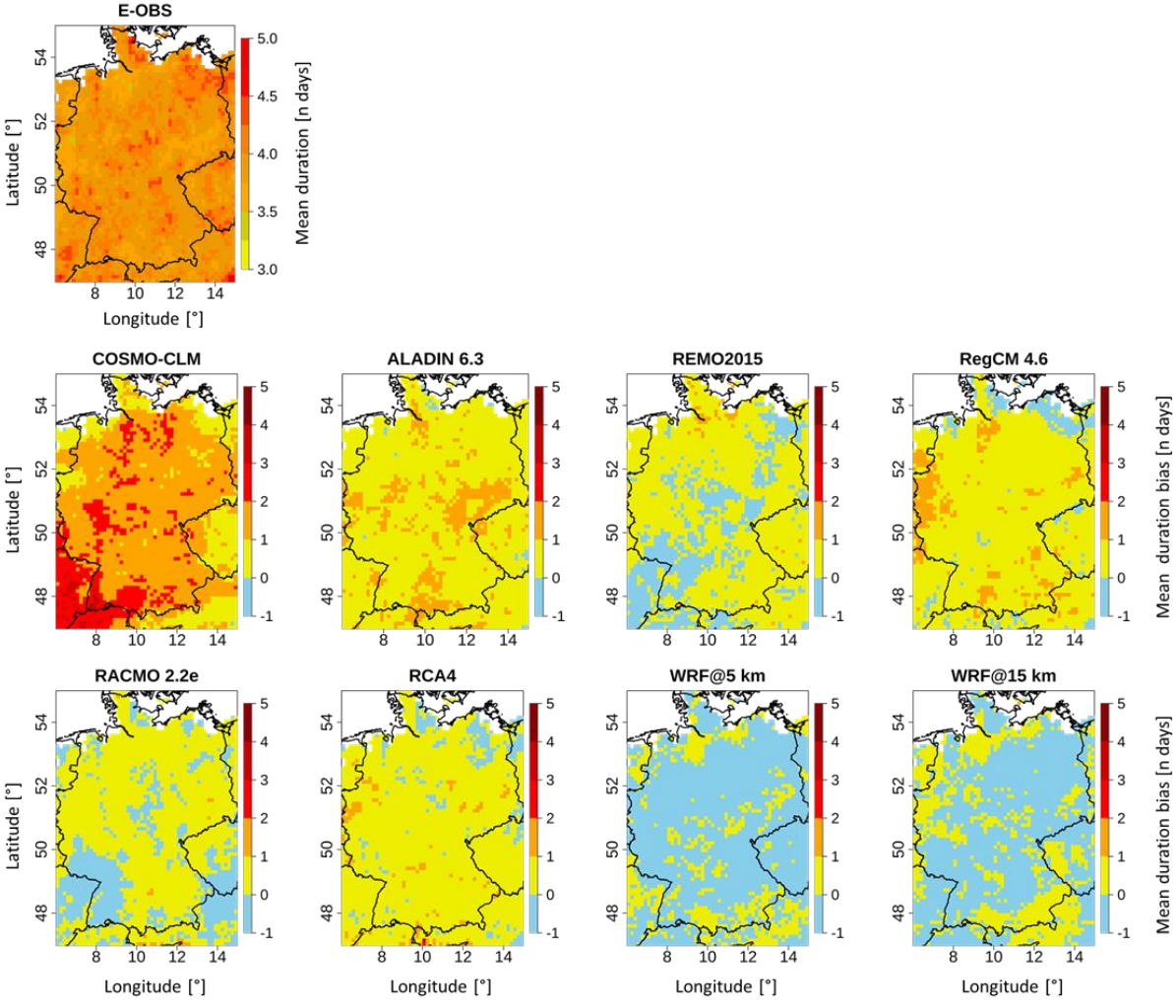


Figure 5. Grid cell based E-OBS summer mean heat wave duration pattern between 1980–2009 and differences between each RCM and E-OBS.

COSMO-CLM is clearly the RCM with the weakest performance due to the weakest scores in each aspect (Table 4). WRF@15 km is the most reliable RCM here, meaning that there are indeed some benefits for this aspect of the analysis, which are rather related to the model settings, since once again, the 15 km performs better than its 5 km counterpart.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

4.3.3 Mean heat wave intensity

Figure 6 provides the E-OBS pattern of the mean heat wave intensity for the time period 1980 – 2009 based on the cumulative heat measure. The grid cell based differences of the RCMs are also included. The relevant scores are given in Table 5.

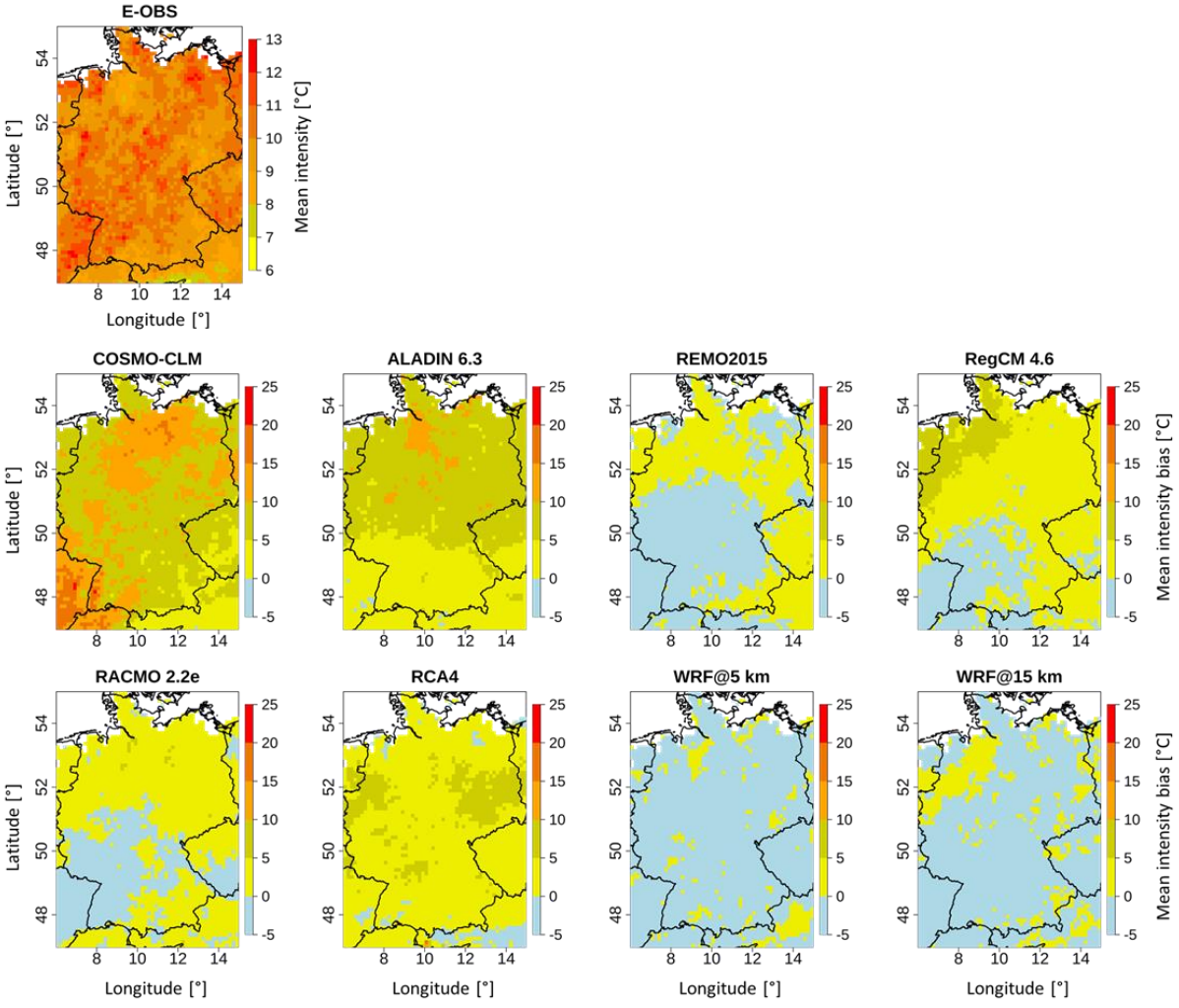


Figure 6. Grid cell based E-OBS summer mean heat wave intensity pattern between 1980–2009 and differences between each RCM and E-OBS.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

Table 5. Mean heat wave intensity metrics.

Model	Mean [°C]	Mean Bias [°C]	SPAEF
COSMO-CLM	18.9	8.88	-0.52
ALADIN 6.3	15.6	5.66	-0.48
REMO2015	9.8	-0.17	-0.19
RegCM 4.6	11.7	1.73	-0.85
RACMO 2.2e	10.9	0.90	-0.49
RCA4	13	3.04	-0.40
WRF@5 km	8.8	-1.17	0.03
WRF@15 km	9.12	-0.85	0.11
E-OBS	9.97		

The E-OBS domain looks quite uniform. A sort of band of higher values extends from the southwest to the northeast. The majority of the values lies within 9 to 11 °C, which is confirmed by the domain mean value of 9.97 °C (Table 5). Accounting for the mean duration (3.94 days) of heat waves from the section above, the average heat excess per day during a heat wave period was 2.53 °C. Regarding the RCM bias maps, the two WRF domains are again predominated by negative values and look quite similar. Areas of positive bias in the 15 km domain are similarly situated as in the 5 km counterpart. Some of them are more extensive like in northwestern part, others are smaller like in southeast. The domains of COSMO-CLM, ALADIN and RCA4 are prevailed by positive values, while those of REMO, RegCM and RACMO show a mixed pattern. In those domains, the areas of cold bias are similar between the models, mainly located in the southwest. They also have in common that the values are mostly between -5 to 5 °C. This leads to relatively small mean bias values (Table 5) due to mutual balancing. Maximum bias values of up to 25 °C are found at the COSMO-CLM domain in the southwestern part. The COSMO-CLM and ALADIN domains are basically covered with comparatively higher values. The domain mean values in Table 5 show distinct differences among themselves with a maximum range of 10.1 °C between COSMO-CLM and WRF@5 km. The two WRF runs and REMO are below the reference value (9.97 °C), all other models are above. REMO's value (9.8 °C) comes closest to the reference, COSMO-CLM (18.9 °C) shows by far the largest difference. Its value is almost double as high as the reference. As inferred from the maps, REMO and the two WRF runs have negative mean bias values, while the other RCMs have positive ones. REMO holds the smallest value (-0.17 °C), COSMO-CLM the highest (8.88 °C), which, compared with the domain mean values, is a very high value. The SPAEF values between the reference and each RCM domain (not shown) are once more all negative (all EURO-CORDEX RCMs) or very low (the two WRF runs), which means that there is no satisfactory reproduction of the reference's spatial structure. RegCM has the lowest value (-0.85), WRF@15 km the highest (0.11). COSMO-CLM is the weakest performing RCM, while REMO showed the overall best performance. This means that there are no real WRF benefits apparent here, except for the possible minor benefits in reproducing the spatial structure. It is striking

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

that the pattern of the WRF domains always being prevailed by negative bias runs through all aspects of the heat wave characteristics as well as the number of heat wave days.

4.4 Heat wave trends

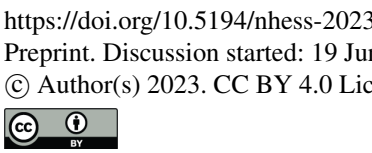
Figure 7 presents the grid cell based results of the Mann-Kendall trend test for the annual number of heat waves in the study period 1980 – 2009 for all RCMs and the E-OBS reference. A summary for each signal and data set is given in Table 6. In this context it should be noted that Mann-Kendall trend test provides information about whether there is a monotonic positive, negative or no trend in a time series at a certain level of significance (here 0.05). It does not give information about exact trend values.

Table 6. Annual number of summer heat wave overall metrics.

Model	negative [%]	neutral [%]	positive [%]
E-OBS	0	88	12
COSMO-CLM	0	99.74	0.26
ALADIN 6.3	0	99.97	0.03
REMO2015	0	99.56	0.44
RegCM 4.6	0	90.35	9.65
RACMO 2.2e	0.03	99.95	0.03
RCA4	0	95.46	4.54
WRF@5 km	0	96.87	3.13
WRF@15 km	0	96.09	3.91

The domains of COSMO-CLM, ALADIN, REMO and RACMO are almost completely covered with no trend signals. This also confirmed by the values in Table 6, where each of these RCMs have more than 99 % of their grid cells in the neutral section. RegCM and RCA4 have distinct areas of positive trends, but in different areas. WRF@5 km has also positive trend areas, but less concentrated. In the E-OBS reference domain there are also concentrated areas of positive trends, mainly in the southwest and in the northern central area. These locations mostly do not agree with those from the RCMs. The WRF@15 km domain also shows positive trend grid cells, but at different areas than the other RCMs. It shows the highest spatial agreement with the reference, especially in the central part. It is striking that there are actually no grid cells with negative i.e. decreasing trend in any domain, which is also confirmed in Table 6. The table further confirms that the majority of the grid cells is covered by no trend signals. By far the largest proportion of positive grid cells can be found in E-OBS (12 %), followed by RegCM (9.65 %), RCA4 (4.54 %) and the WRF runs (3.13 % and 3.91 %). All other runs are in the less than 1 % range. This shows

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023



Natural Hazards and Earth System Sciences
Discussions

that the WRF@15 km run is not only from the locations, but also from the shares of the signals closer to the reference than its 5 km counterpart.

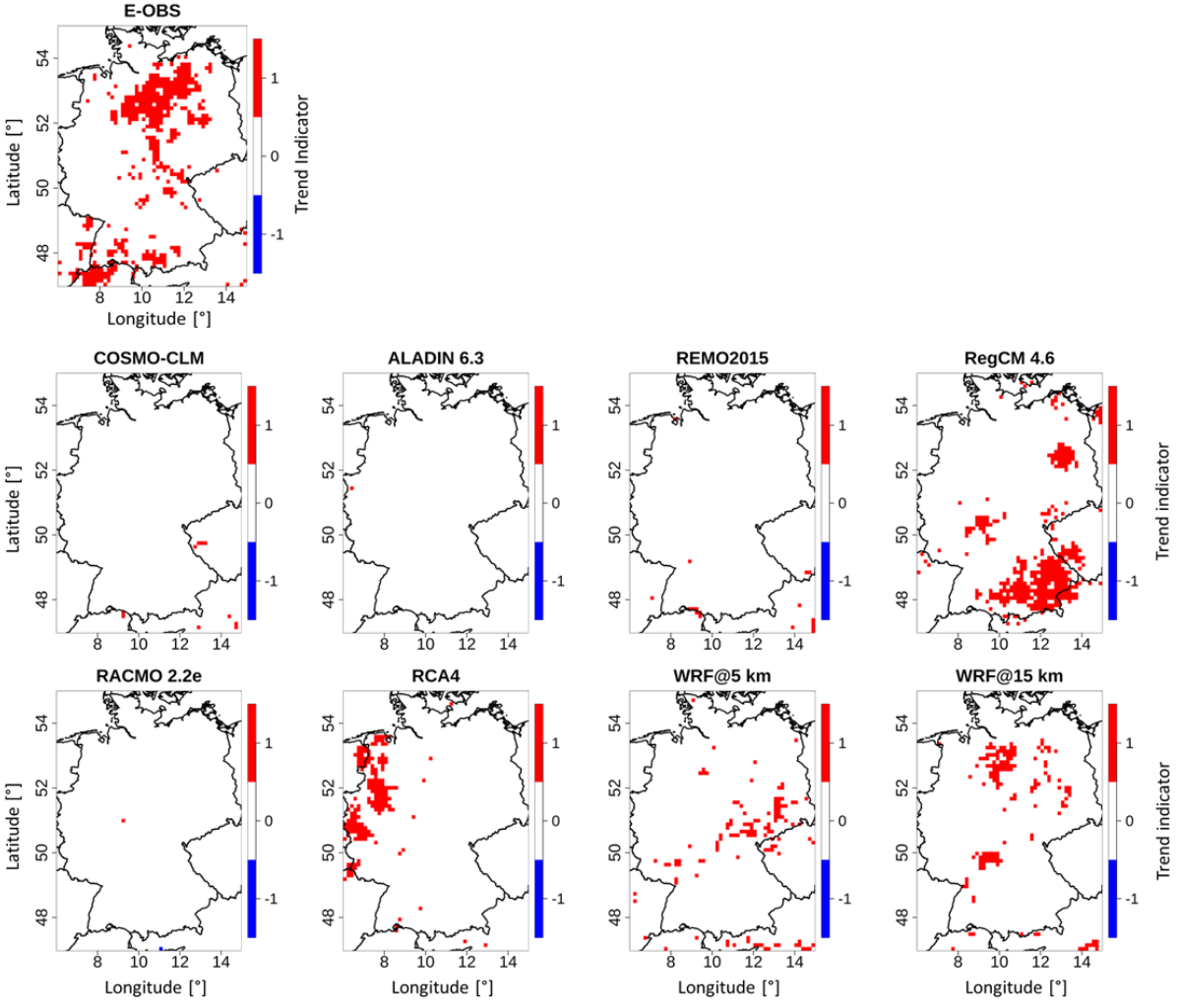
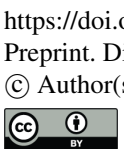


Figure 7. Grid cell based trends of the number of annual summer heat waves for 1980–2009 based on the Mann-Kendall Test for E-OBS and each RCM.

This section reveals that, according to the E-OBS reference, if there is a trend in number of heat waves, it is only positive, meaning that the frequency is increasing with time. But this is not the case everywhere. All models simulate too few pixels with positive trends. Regarding WRF, any possible benefits would be related to the model settings rather than to the grid resolution, since the WRF@15 km is more accurate than its 5 km counterpart.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023



Natural Hazards and Earth System Sciences
Discussions

4.5 The heat wave event 2003

4.5.1 Cumulative heat

Figure 8 shows the E-OBS cumulative heat pattern of the summer season 2003 and the grid cell based differences with each RCM. The associated scores are given in Table 7.

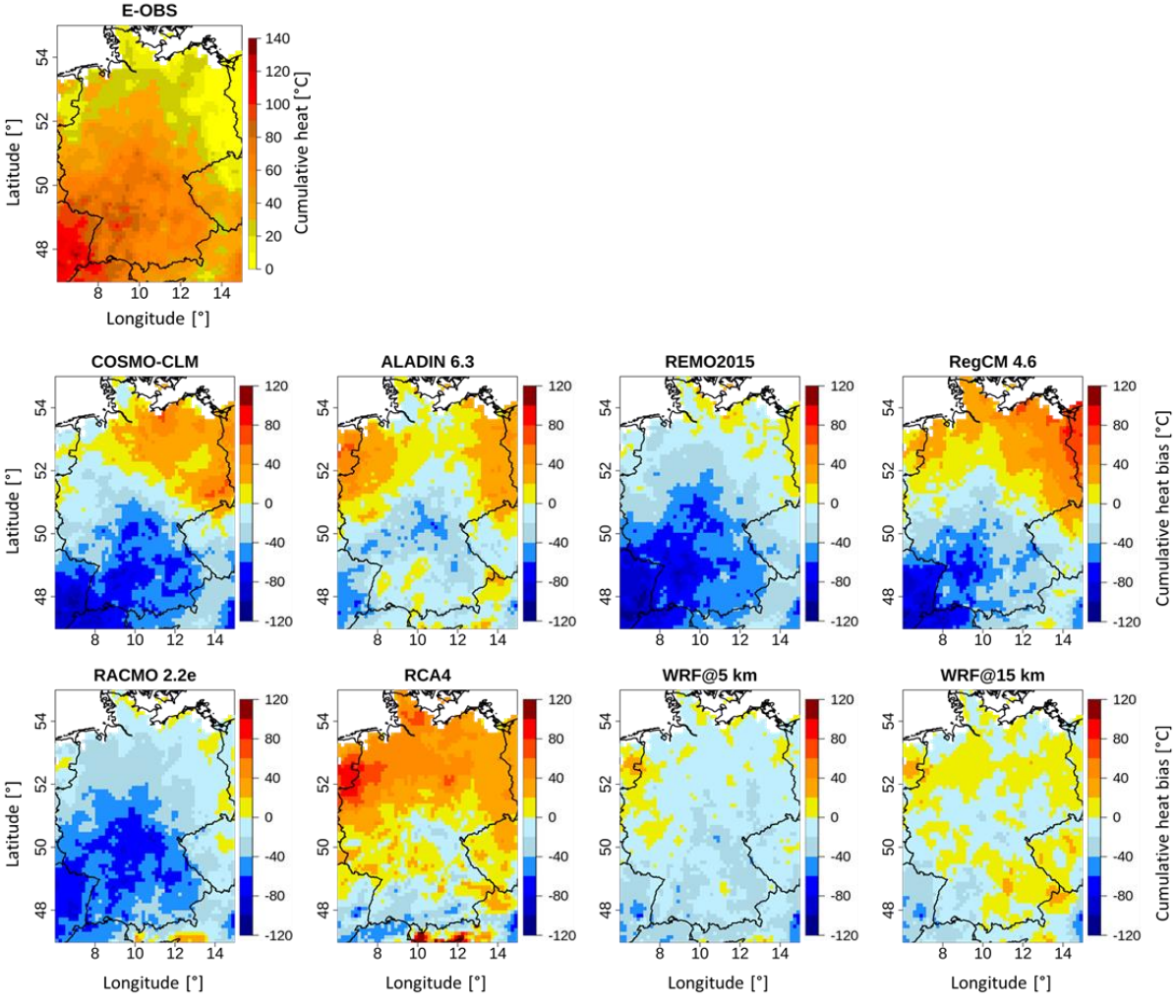


Figure 8. Grid cell based E-OBS summer 2003 cumulative heat pattern and differences between each RCM and E-OBS.

The E-OBS domain shows a quite clear gradient from the southwest to the northeast with decreasing values. The values in the southwest are very high, they have accumulated well above 100 °C during that summer season, making the heat wave most pronounced in this region. These values underline the high intensity of the 2003 heat wave. The mildest values of up to 10 °C heat excess are in the northeastern regions, making them the regions least impacted by the heat wave. COSMO-CLM, REMO, RegCM and RACMO show pronounced cold bias values in the southwest of their domains. This means that they were not able

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

to satisfactorily reproduce the particularly high values of the reference in these regions, showing bias values of up to -120 °C. These bias values indicate that the models simulated only weak or even no heat episodes at all in regions, where the reference showed the most pronounced values. REMO, RACMO and WRF@5 km are prevailed by cold bias values, the remaining RCMs except WRF@15 show mixed patterns, where the northern half is prevailed by warm, the southern half by cold bias, leading to a sort of bipartition. The WRF@15 km is evenly covered with values of both signs. The WRF@5 km pattern is most uniform with relatively low bias values all over the domain. It is striking that in contrast to the previous sections, the WRF domains are not the only ones prevailed by negative respective cold bias values here. In direct comparison with its 5 km counterpart, the WRF@15 km domain has much more areas with warm bias. The domain mean values in Table 7 show a large discrepancy among each other with a range of 52.7 °C between RCA4 (62.3 °C) and REMO (9.6 °C). The reference value of E-OBS is 45.4 °C. Considering the mean heat wave intensity (9.9 °C) for the whole study period from Sect. 4.3.3, this value is more than four times higher and thus very remarkable and it illustrates the great severity of this heat wave. ALADIN (43.2 °C) is closest to the reference value, REMO holds the largest difference. Regarding the mean bias values, REMO (-35.9 °C) and RACMO (-34 °C) have the highest values, ALADIN (-2.2 °C) the lowest. It needs to be considered that in the RCMs with the bipartition pattern mentioned above (COSMO-CLM, ALADIN and RegCM), the values cancel each other out, leading to relatively low mean bias values, depending on the degree of balance. Only RCA4 holds a positive mean bias value (16.9 °C). This underlines that the models rather underestimate the intensity of this heat wave period. The WRF@5 km mean bias (-12.9 °C) is clearly higher than that of its 15 km counterpart (-3.4 °C). Here there are some distinct differences between the SPAEF values. While they are negative for the most EURO-CORDEX RCMs or very low (ALADIN and RCA4), the two WRF runs show relatively high values (0.77 for WRF@5 km and 0.72 for WRF@15 km). This means that the WRF runs reproduced the spatial structure of the reference reasonably well. RegCM holds the lowest (-0.69) SPAEF value.

Table 7. Cumulative heat 2003 metrics.

Model	Mean [°C]	Mean Bias [°C]	SPAEF
COSMO-CLM	25.6	-19.8	-0.38
ALADIN 6.3	43.2	-2.2	0.26
REMO2015	9.6	-35.9	-0.04
RegCM 4.6	35.9	-9.5	-0.69
RACMO 2.2e	11.4	-34	-0.03
RCA4	62.3	16.9	0.09
WRF@5 km	32.5	-12.9	0.77
WRF@15 km	42.1	-3.4	0.72
E-OBS	45.4		

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

435 There are distinct differences between the single models in this section. ALADIN is the RCM with the overall best performance, REMO the one with the worst. There are pronounced benefits of the WRF runs visible in the reproduction of the spatial structure of the reference. In this regard, the 5 km WRF run performs slightly better than its 15 km counterpart. In terms of reproducing the reference domain mean value and of mean bias value, the 15 km WRF run outperforms its 5 km counterpart.

4.5.2 Maximum duration

440 Figure 9 gives the E-OBS pattern of the maximum heat wave duration during the 2003 summer season along with the grid cell based differences with each RCM. The corresponding values are given in Table 8.

Table 8. Maximum heat wave duration 2003 metrics.

Model	Mean [n days]	Mean Bias [n days]	SPAEF
COSMO-CLM	6.91	-1.76	-0.64
ALADIN 6.3	6.34	-2.33	-0.17
REMO2015	4.30	-4.37	-0.42
RegCM 4.6	5.51	-3.15	-0.69
RACMO 2.2e	2.87	-5.80	-0.44
RCA4	8.94	0.27	-0.35
WRF@5 km	7.72	-0.95	0.19
WRF@15 km	8.86	0.19	0.07
E-OBS	8.67		

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

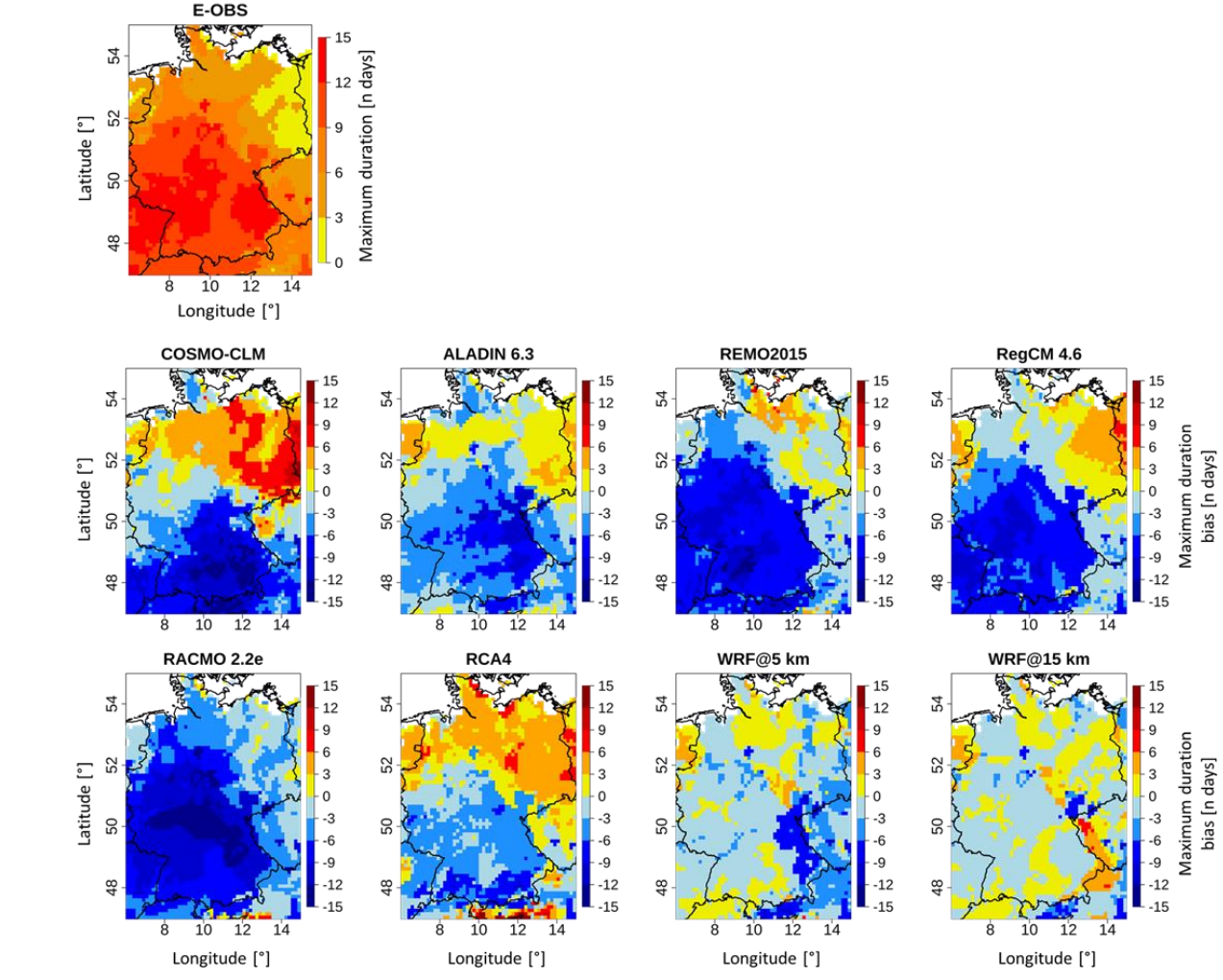


Figure 9. Grid cell based E-OBS summer 2003 heat wave maximum duration pattern and differences between each RCM and E-OBS.

The E-OBS domain shows a sort of bipartition with the highest values in the southwestern and lower values in the northeastern part. This matches the impressions from the section above, where the highest heat excess values were also found in the southwest (Figure 8). Due to the longest durations of up to 16 days, the high heat excess values could accumulate. It should be noted that this is not only a matter of duration, but also of excess values. In the northeastern part there are areas with values ranging between 0 – 3. Since the minimum duration of a heat wave episode was defined to be three days, it means that in these regions no heat episode took place in that summer period. This is also in line with the heat excess pattern (Figure 8), which shows values between 0 and 10 °C in some of these areas. The RCM bias patterns roughly match those in Figure 8. COSMO-CLM, REMO, RegCM and RACMO have strong negative bias values of up to -15 days in the southern and southwestern areas. Since the E-OBS domain shows the highest values of up to 15 days in some parts of those areas, it further confirms that the

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

models actually did not simulate a heat wave episode in some parts of these areas, where the reference shows the most distinct values. It further underlines this big shortcoming of the models. The northern parts in COSMO-CLM and RCA4 are prevailed by positive bias values. COSMO-CLM is the model with the largest areas of high positive bias values of up to 15 days in the eastern parts. This is the region where the E-OBS reference only shows weak or even no heat wave episodes, meaning that the model does the opposite here compared to the southwestern region, simulating relatively strong heat episodes although there were none according to reference. In almost each domain there are high bias values in the Alps region in the south. With the exception of RCA4, these bias values are all negative. In line with the section above, the WRF domains are not the only ones prevailed by negative values. In the southwestern regions, they both show the lowest bias values of all RCMs. The WRF@15 km domain shows the most balanced pattern between positive and negative bias values with most of them ranking in the relatively low range. This is confirmed by the lowest mean bias values (0.19 days) in Table 8. The domain mean values in Table 8 reveal that the E-OBS reference value (8.67 days) is one of the highest, only exceeded by RCA4 (8.94 days) and WRF@15 km (8.86 days). This is also reflected in the mean bias values. It indicates that the models tend to simulate shorter durations. The WRF@15 km value is also the closest to the reference, while RACMO (2.87 days) shows the biggest difference. It is also by far the lowest value. The model also holds the highest mean bias value (-5.80 days). It needs to be remembered that the low mean bias values of WRF@15 km and RCA4 (0.27 days) also result from the contradictory nature of their values. Analogous to the cumulative heat, the two WRF runs hold the best SPAEF values (0.19 and 0.07, respectively), but on a much lower level, making them not very meaningful. The score of the WRF@5 km domain is clearly better than that of its 15 km counterpart. All SPAEF values of the EURO-CORDEX RCMs are negative. The lowest score is found at RegCM (-0.69). Overall, no model achieves to reproduce the spatial structures satisfactorily.

According to the scores, WRF@15 km shows the best overall performance in this section, RACMO the worst. This means that once more WRF@15 km outperformed its 5 km counterpart. The great weakness of COSMO-CLM of simulating oppositely needs also to be considered, since this may not be reflected in the values of the table. The fact that the WRF runs, especially the 5 km run, in this as well as in the previous section were not the only models prevailed by cold respective negative bias values (Figure 8 and Figure 9) like in all the previous sections above, highlights that for a single event, the situation can be very different compared with the overall picture for the complete study period provided in the Sect. 4.2 and 4.3.

5. Discussion

Regarding the T_{max} reproduction in Sect. 4.1, Silva et al. (2022), who compared monthly T_{max} values from six historical runs from different GCM-RCM combinations from CORDEX-CORE with ERA5 as reference for the Pantanal region for the period April - October between 1981 and 2005, found temporal, area averaged correlation values between 0.42 and 0.67. This is distinctly less than in our case. The different RCM forcings and focus on different period and region must be considered though. As for the bias values (Table 1), in previous studies, a negative bias of daily T_{max} was found for the Central European

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

region (Nikulin et al., 2011; Plavcová and Kyselý, 2011). Here, we found both, positive and negative bias, depending on the RCM.

Regarding the heat wave characteristics (Sect. 4.3), there are different reasons discussed in the literature for the over and underestimation by the models. Lhotka et al. (2018b) assume that the overestimation of heat wave frequency and duration of major heat waves, which they found in some RCMs, is related to the large-scale circulation and soil moisture depletion. Underestimation of these events, on the other hand, is associated with too-moist summertime conditions. Vautard et al. (2013), who, like in this study, also found that simulated heat waves from the EURO-CORDEX RCMs were too long and intense (not the case for REMO here), attribute this to biases in the modelled temperature. It must be noted that they used the daily mean instead of maximum temperature for their heat wave definition. Another possible reason they give is a connection between land–atmosphere feedback and overestimated heat wave intensity in models. Lhotka and Kyselý (2015a) also go into this direction, since they found a connection between heat wave intensity and precipitation during and before these events. Vautard et al. (2013) further found that coarser resolution lead to very persistent heat waves. This is also how it looks here at first, if only the WRF@5 km run is compared with the EURO-CORDEX RCMs. The WRF@15 km run shows that this cannot be related to resolution, but to setting effects, since it is coarser then the EURO-CORDEX runs. It must be noted that the resolution differences in Vautard et al. (2013) (50 km vs. 12.5 km) were much more distinct than in our case. According to Plavcová and Kyselý (2019), an overestimation of circulation super types may contribute to the development of too-long heat waves in some simulations. From an overall perspective, the heat wave characteristics are generally quite well captured in terms of spatiotemporal mean values (Table 3 - 5). Lin et al. (2022) came to similar findings, even though they used different heat wave metrics.

As for the heat wave event 2003 (Sect. 4.5), Lin et al. (2022) associate the high bias values in the Alps region for the maximum durations with the representation of orographic effects. Russo et al. (2016) also found a big discrepancy in the RCM's capability to simulate a single major heat wave event. In fact, only one out of 13 RCMs was able to capture the event. They attribute this to some model deficiency simulating really extreme heat waves or to the length of the analyzed time period (1979 – 2005). This shows that increased resolution does not lead to improved reproduction of severe heat waves, which was found in Lhotka et al. (2018a). It must be considered that the difference between the resolutions they compared (12.5 vs. 50 km) was much larger than in our case.

In each section of this study we found an inter-model spread. This is in line with previous heat wave related model employing studies, e.g., Vautard et al. (2013), Gibson et al. (2017), Feron et al. (2019) and Silva et al. (2022). Vautard et al. (2013), who also used ERA-Interim driven EURO-CORDEX outputs, assume several potential sources of spread: the method to process boundary conditions in the model, the convection treatment, the different parameterizations, which are more pronounced in larger domain sizes and specific weather situations and the way in which the interactions between land surfaces and the atmosphere are accounted for in the models. The latter refers to the uncertainty of partitioning between sensible and latent heat flux as well as of radiation fluxes (Lenderink et al., 2007; Stegehuis et al., 2012).

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

In each of the heat wave related sections, no evidence is found that increased resolution leads to better results in reproducing the related metrics. In fact, the WRF@15 km always performed better than its 5 km counterpart. This is partly in line with the results of previous studies. No benefits of increased resolution were also found in Plavcová and Kyselý (2019), who compared 25 km and 50 km resolution and in Molina et al. (2020) (12.5 km vs. 50 km). Cardoso et al. (2019), who also compared 12.5 and 50 km resolution, found a slight benefit of increased resolution. Careto et al. (2022) and Lin et al. (2022) compared 0.11° RCM outputs with the outputs of the driving data in much higher, different resolutions. In both studies benefits of the increased resolution were found, particularly for coastal regions. Vautard et al. (2013) (12.5 vs 50 km) identified partly benefits of increased resolution, depending on the aspects of analysis. They found that increased resolution leads to reduction in biases in the temperature 90th percentile as well as in the heat wave persistence. They suspect that smaller scale and higher frequency variability is better resolved in higher resolutions. Local bias improvements were found in some coastal regions here too. Once more it is noted that in each of the mentioned studies, the difference between the employed resolutions is much higher than in our case here. We assume from our results that a resolution increase from an already relatively high resolution, in this case 12.5 km, has limited to negligible impact. One fact that also needs to be considered is the original resolution difference between E-OBS (12.5 km) and WRF@5 km. For certain aspects, structures from E-OBS may be better represented in resolutions closer to it. This could be a reason why the 15 km WRF runs performs better with otherwise the same settings. Furthermore, it needs to be kept in mind that in the E-OBS data set extreme values tend to be smoothed out due to interpolation processes (Haylock et al., 2008; Hofstra et al., 2009). This can mean that for certain aspects of the analysis, where the models showed cold or negative bias, discrepancy to the true values might be even bigger. Hofstra et al. (2009) emphasize, that this effect is more pronounced for precipitation though.

The three RCMs with the overall best performances are ALADIN, REMO and WRF@15 km. COSMO-CLM showed the weakest overall performance. If this is considered, the choice of the parameterization schemes seems to play a minor role, since ALADIN, REMO and WRF have all different schemes for the individual physics. This finding is opposite to that from Davin et al. (2016) who identified the land surface scheme as highly important for a proper simulation of temperature. The land surface scheme is determining for two crucial factors: the soil moisture and leaf area index (LAI). If the LAI within the land-surface model is based on climatological values instead of dynamical calculations, this can increase evapotranspiration and thus lead to a cooling effect, which reduces the maximum temperature values. The connection between adjacent soil moisture conditions and heat extremes is well documented (e.g., Brabson et al., 2005; Fischer et al., 2007; Lorenz et al., 2010; Hirschi et al., 2011; Jaeger and Seneviratne, 2011; Liu et al., 2013). One possible determiner for bias can further be the microphysics scheme, which is responsible for the cloud processes. In earlier studies, the role of cloud cover for the (maximum) temperature simulation is highlighted. An increased cloud cover leads to a greater fraction of reflected solar radiation, which in turn leads to cooling of T_{max} (Groisman et al., 2000; Sun et al., 2000). Lobell et al. (2007) found that cloud cover is responsible for higher daily T_{max} variability compared to the daily mean values. They consider the cloud cover especially important during the summer period. According to Liang et al. (2008), biases in simulated radiation budgets can lead to errors in surface temperatures. Hamdi et al. (2012) found strong correlations of warm bias with cloud cover representation. However,

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

555 since all models in this study were ran with different microphysics schemes (Table 2 in Petrovic et al. (2022)), relevant conclusions cannot really be drawn. Interestingly, in the study of Lhotka et al. (2018b), COSMO-CLM in combination with a driving GCM was among the RCMs with the best performances in simulating major European heat waves. This could indicate the high importance of the driving data, which is also highlighted by Molina et al. (2020). It is further underlined by the fact that RACMO and RCA4, driven by the same GCM, also showed best performances, while in our study they did not stand out.
560 It is evident that the WRF runs, compared to the EURO-CORDEX RCMs, often showed rather negative/colder bias values. This might be connected with the employed schemes in the WRF simulation (Table 2 in Petrovic et al. (2022)), since there are no commonalities with the other runs at the individual physics. Moreover, a connection between the T_{max} bias (Table 1) and the overall performance does not seem to exist, since COSMO-CLM has the lowest mean bias value (-0.16 °C), but the worst overall performance, while on the other hand RCA4, which holds the highest mean bias value (-2.40 °C), does not show a
565 significant bad performance. There are model output employing heat wave studies in which it has been found that the ensemble mean outperforms the individual models runs, e.g., Russo et al. (2016), Wang et al. (2019a), Lin et al. (2022) and Kim et al. (2023). Here, ensemble mean values were not considered since a focus of the study is about the impact of different model resolution and settings and evaluation of single model performances.

According to Plavcová and Kyselý (2019), biases in the simulations of atmospheric circulation play a crucial role for the
570 simulation of temperature extremes. This is why they claim that an improvement in this field will be among the most important steps towards better reproduction of extreme temperature events and thus also lead to more credibility of future projections.

It is striking that there are some significant differences in the outcomes compared to the drought study (Petrovic et al., 2022) in which the same data sources were used as mentioned above. While it was found that all models performed at similar levels for the drought characteristics, there are some significant differences between the individual performances here for the heat
575 wave characteristics, highlighting that the choice of model can be crucial. In addition, it was shown that the WRF settings and increased resolution were particularly beneficial for reproducing drought trends. This is not the case for heat wave trends. This suggests that these benefits are highly related to the simulation of precipitation, the most important variable for drought, which is not a factor here. The different time scales could also be a factor. Unlike heat waves, where the minimum duration in this study is three days, droughts are prolonged events with a minimum duration of usually at least one month. For this reason,
580 monthly values were considered in the drought study, while daily values were used here. Regarding the reproduction of the 2003 drought and heat wave event, there are pronounced differences between the models in both studies. Interestingly, REMO shows the worst performance in this respect in both cases and significantly underestimates the drought and heat wave conditions, respectively.

6. Conclusions

585 A heat wave analysis for Germany and its near surroundings for the period 1980-2009 was performed. The impact of increased model resolution and appropriate model configuration on the reproduction of heat wave metrics based on the T_{max} simulation

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

is addressed. For this purpose, we employ an ensemble of six ERA-Interim-driven EURO-CORDEX RCMs of 12.5 km horizontal grid resolution and an ERA Interim-driven high-resolution (5 km) WRF run, whose setup was tailored to the target area. The model outputs are evaluated with regard to their ability to reproduce T_{max} and heat wave characteristics based on it, trends and the major event in 2003. E-OBS data is used as reference.

WRF with its increased resolution and tailored model settings is shown to be not necessarily beneficial for improved performance in reproducing heat indices. Benefits are somewhat present only for the reproduction of the mean heat wave durations and for the reproduction of the spatial structures of the cumulative heat values for the heat wave event 2003. In fact, the WRF@15 km run outperforms its 5 km counterpart in each section. Thus, we conclude that (for the selected model configurations) increased resolution does not contribute to better performances regarding heat wave metrics, when both compared resolutions are already relatively high, which is the case here (12.5 km vs. 5 km). Since three models, namely ALADIN, REMO and WRF@15 km show the overall best performances, we further conclude that the tailored model settings of WRF only have limited benefits for the reproduction of the heat wave metrics. The daily T_{max} values are reproduced relatively well by all models, which is also underlined by the rather low mean bias values in Table 1. Regarding the domain mean conditions of the overall characteristics, all models show reasonable performances for the heat wave frequency and mean duration, while this does not go for the mean intensity. The spatial agreement with the reference is not satisfactory for any RCM and section with the exception of the two WRF runs in the reproduction of the cumulative heat pattern for the 2003 event. In general, despite the same forcing by ERA-Interim, the RCMs exhibit a significant spread in their outputs. This is especially pronounced for the 2003 event, which underlines the difficulty of the models to reproduce single major events. Regarding the heat wave trends, the reference shows that, if there is a trend present, it is only increasing, indicating increases in the number of heat waves with time. The RCMs struggle with reproducing these trends. If trends are indicated, they are mostly not spatially accurate. All RCMs underestimate the proportion of grid cells with increasing trends. No specific physics scheme or configuration can be shown to be especially beneficial for the reproduction of the heat wave metrics. Furthermore, there seems to be no correlation between the RCM bias values (Table 1) and the respective RCM performances. According to the E-OBS reference, heat waves occurred about 31 times in the study period with an average duration of about 4 days and average intensity of about 10 °C. This equals an average heat excess per day during a heat wave period of about 2.5 °C.

Our results suggest that a resolution of 12.5 km or even 15 km, as shown with the WRF@15 km run, is sufficient to reach similar findings to those obtained with finer resolutions. The study further reveals that the choice of the RCM can be crucial, due to the partly enormous differences in the performances. Furthermore, it is shown that the adjusted model settings to the specific target region of WRF only have limited impacts, suggesting that for the reproduction of T_{max} and thus heat waves this is a less important factor. The results may guide in the selection of suitable RCMs for certain aspects of heat wave analysis in Germany and similar regions. Not only in a historical context, but also for future projections.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

Data Availability

The EURO-CORDEX data is freely available at the EURO-CORDEX website (https://www.euro-cordex.net/). The E-OBS data is freely available at the ECA&D website (https://www.ecad.eu/). The WRF data and the associated configuration files can be obtained online from Petrovic (2023, https://doi.org/10.5281/zenodo.7998809).

Author contribution

DP, BF and HK developed the methodology for the study. DP carried out the data analysis and drafted the manuscript, with support of BF and HK. HK provided grant funding and supervised the research.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the work of the WRF modeling community, the European Centre for Medium-Range Weather Forecast for the reanalysis data ERA-Interim, the contributors to the EURO-CORDEX projects used in this study, the ECA&D group for the E-OBS data set and Warscher et al. (2019) for providing the WRF simulation data. Big thanks go to Gerhard Smiatek for his support. This work is funded by the ClimXtreme project of the BMBF (German Federal Ministry of Education and Research) under grant “Förderkennzeichen 01LP1903J”.

References

Aich, V., Akhundzadah, N., Knuerr, A., Khoshbeen, A., Hattermann, F., Paeth, H., Scanlon, A., and Paton, E.: Climate Change in Afghanistan Deduced from Reanalysis and Coordinated Regional Climate Downscaling Experiment (CORDEX)—South Asia Simulations, Climate, 2, 38, https://doi.org/10.3390/cli5020038, 2017.

Ballester, J., Rodó, X., and Giorgi, F.: Future changes in Central Europe heat waves expected to mostly follow summer mean warming, Clim Dyn, 7-8, 1191–1205, https://doi.org/10.1007/s00382-009-0641-5, 2010.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The hot summer of 2010: redrawing the temperature record map of Europe, Science, 6026, 220–224, https://doi.org/10.1126/science.1201224, 2011.

Bastos, A., Gouveia, C. M., Trigo, R. M., and Running, S. W.: Analysing the spatio-temporal impacts of the 2003 and 2010 extreme heatwaves on plant productivity in Europe, Biogeosciences, 13, 3421–3435, https://doi.org/10.5194/bg-11-3421-2014, 2014.

Becker, F. N., Fink, A. H., Bissolli, P., and Pinto, J. G.: Towards a more comprehensive assessment of the intensity of historical European heat waves (1979–2019), Atmospheric Science Letters, 11, e1120, https://doi.org/10.1002/asl.1120, 2022.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Natural Hazards and Earth System Sciences
Discussions

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A. T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K.: Future extreme events in European climate: an exploration of regional climate model projections, Climatic Change, S1, 71–95, https://doi.org/10.1007/s10584-006-9226-z, 2007.

Brabson, B. B., Lister, D. H., Jones, P. D., and Palutikof, J. P.: Soil moisture and predicted spells of extreme temperatures in Britain, J. Geophys. Res., D5, https://doi.org/10.1029/2004JD005156, 2005.

Cardoso, R. M., Soares, P. M. M., Lima, D. C. A., and Miranda, P. M. A.: Mean and extreme temperatures in a warming climate: EURO CORDEX and WRF regional climate high-resolution projections for Portugal, Clim Dyn, 1-2, 129–157, https://doi.org/10.1007/s00382-018-4124-4, 2019.

Careto, J. A. M., Soares, P. M. M., Cardoso, R. M., Herrera, S., and Gutiérrez, J. M.: Added value of EURO-CORDEX high-resolution downscaling over the Iberian Peninsula revisited – Part 2: Max and min temperature, Geosci. Model Dev., 6, 2653–2671, https://doi.org/10.5194/gmd-15-2653-2022, 2022.

Christensen, J. H., and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Climatic Change, S1, 7–30, https://doi.org/10.1007/s10584-006-9210-7, 2007.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., Noblet, N. de, Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., and Valentini, R.: Europe-wide reduction in primary productivity caused by the heat and drought in 2003, Nature, 7058, 529–533, https://doi.org/10.1038/nature03972, 2005.

Coppola, E., Raffaele, F., Giorgi, F., Giuliani, G., Xuejie, G., Ciarlo, J. M., Sines, T. R., Torres-Alavez, J. A., Das, S., Di Sante, F., Pichelli, E., Glazer, R., Müller, S. K., Abba Omar, S., Ashfaq, M., Bukovsky, M., Im, E.-S., Jacob, D., Teichmann, C., Remedio, A., Remke, T., Kriegsmann, A., Bülow, K., Weber, T., Buntemeyer, L., Sieck, K., and Rechid, D.: Climate hazard indices projections based on CORDEX-CORE, CMIP5 and CMIP6 ensemble, Clim Dyn, 5-6, 1293–1383, https://doi.org/10.1007/s00382-021-05640-z, 2021.

Coumou, D., and Rahmstorf, S.: A decade of weather extremes, Nature Clim Change, 7, 491–496, https://doi.org/10.1038/nclimate1452, 2012.

Davin, E. L., Stöckli, R., Jaeger, E. B., Levis, S., and Seneviratne, S. I.: COSMO-CLM2: a new version of the COSMO-CLM model coupled to the Community Land Model, Clim Dyn, 9-10, 1889–1907, https://doi.org/10.1007/s00382-011-1019-z, 2011.

De Bono, A., Giuliani, G., Kluser, S., and Peduzzi, P.: Impacts of summer 2003 heat wave in Europe, UNEP/DEWA/GRID-Europe Environ Alert Bull 2:1–4, 2004.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 656, 553–597, https://doi.org/10.1002/qj.828, 2011.

Della-Marta, P. M., Luterbacher, J., Weissenfluh, H. von, Xoplaki, E., Brunet, M., and Wanner, H.: Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability, Clim Dyn, 2-3, 251–275, https://doi.org/10.1007/s00382-007-0233-1, 2007.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

685 Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrol. Earth Syst. Sci., 2, 1299–1315, https://doi.org/10.5194/hess-22-1299-2018, 2018.

Di Liu, Wang, G., Mei, R., Yu, Z., and Yu, M.: Impact of initial soil moisture anomalies on climate mean and extremes over Asia, J. Geophys. Res. Atmos., 2, 529–545, https://doi.org/10.1002/2013JD020890, 2014.

690 Di Luca, A., Elía, R. de, and Laprise, R.: Potential for added value in temperature simulated by high-resolution nested RCMs in present climate and in the climate change signal, Clim Dyn, 1-2, 443–464, https://doi.org/10.1007/s00382-012-1384-2, 2013.

Diffenbaugh, N. S., and Ashfaq, M.: Intensification of hot extremes in the United States, Geophys. Res. Lett., 15, n/a-n/a, https://doi.org/10.1029/2010GL043888, 2010.

695 Dosio, A.: Projection of temperature and heat waves for Africa with an ensemble of CORDEX Regional Climate Models, Clim Dyn, 1-2, 493–519, https://doi.org/10.1007/s00382-016-3355-5, 2017.

Feron, S., Cordero, R. R., Damiani, A., Llanillo, P. J., Jorquera, J., Sepulveda, E., Asencio, V., Laroze, D., Labbe, F., Carrasco, J., and Torres, G.: Observations and Projections of Heat Waves in South America, Scientific reports, 1, 8173, https://doi.org/10.1038/s41598-019-44614-4, 2019.

700 Fink, A. H., Brücher, T., Krüger, A., Leckebusch, G. C., Pinto, J. G., and Ulbrich, U.: The 2003 European summer heatwaves and drought -synoptic diagnosis and impacts, Weather, 8, 209–216, https://doi.org/10.1256/wea.73.04, 2004.

Fischer, E. M., and Schär, C.: Consistent geographical patterns of changes in high-impact European heatwaves, Nature Geosci, 6, 398–403, https://doi.org/10.1038/ngeo866, 2010.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D., and Schär, C.: Soil Moisture–Atmosphere Interactions during the
705 2003 European Summer Heat Wave, J. Climate, 20, 5081–5099, https://doi.org/10.1175/JCLI4288.1, 2007.

Gibson, P. B., Perkins-Kirkpatrick, S. E., Alexander, L. V., and Fischer, E. M.: Comparing Australian heat waves in the CMIP5 models through cluster analysis, J. Geophys. Res. Atmos., 6, 3266–3281, https://doi.org/10.1002/2016JD025878, 2017.

Giorgi, F., Colin Jones, and Ghassem R. Asrar: Addressing climate information needs at the regional level: the CORDEX framework, World Meteorological Organization (WMO) Bulletin, 3, 175–183, 2009.

710 Groisman, P. Y., Bradley, R. S., and Sun, B.: The Relationship of Cloud Cover to Near-Surface Temperature and Humidity: Comparison of GCM Simulations with Empirical Data, Journal of Climate, 11, 1858–1878, https://doi.org/10.1175/1520-0442(2000)013<1858:TROCCT>2.0.CO;2, 2000.

Hamdi, R., van de Vyver, H., and Termonia, P.: New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: application for summer extreme temperature over Belgium, Int. J. Climatol., 13, 2051–2065,
715 https://doi.org/10.1002/joc.2409, 2012.

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., D20, https://doi.org/10.1029/2008JD010201, 2008.

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B.,
720 and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, Nature Geosci, 1, 17–21, https://doi.org/10.1038/ngeo1032, 2011.

Hofstra, N., Haylock, M., New, M., and Jones, P. D.: Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature, J. Geophys. Res., D21, https://doi.org/10.1029/2009JD011799, 2009.

Hoy, A., Hänsel, S., Skalak, P., Ustrnul, Z., and Bochníček, O.: The extreme European summer of 2015 in a long-term perspective, Int. J. Climatol., 2, 943–962, https://doi.org/10.1002/joc.4751, 2017.

Jaeger, E. B., and Seneviratne, S. I.: Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model, Clim Dyn, 9-10, 1919–1939, https://doi.org/10.1007/s00382-010-0780-8, 2011.

Kendall, M.G.: Rank Correlation Methods, Griffin, London, 1975.

Kim, Y.-H., Ahn, J.-B., Suh, M.-S., Cha, D.-H., Chang, E.-C., Min, S.-K., Byun, Y.-H., and Kim, J.-U.: Future changes in extreme heatwaves in terms of intensity and duration over the CORDEX-East Asia Phase Two domain using multi-GCM and multi-RCM chains, Environ. Res. Lett., 3, 34007, https://doi.org/10.1088/1748-9326/acb727, 2023.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models, Geosci. Model Dev., 5, 1873–1886, https://doi.org/10.5194/gmd-11-1873-2018, 2018.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geosci. Model Dev., 4, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014, 2014.

Kumar, R., Samaniego, L., and Attinger, S.: The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics, Journal of Hydrology, 1-2, 54–69, https://doi.org/10.1016/j.jhydrol.2010.07.047, 2010.

Kyselý, J., Plavcová, E., Davídkovová, H., and Kynčl, J.: Comparison of hot and cold spell effects on -cardiovascular mortality in individual population groups in the Czech Republic, Clim. Res., 2, 113–129, https://doi.org/10.3354/cr01014, 2011.

Kyselý, J.: Recent severe heat waves in central Europe: how to view them in a long-term prospect?, Int. J. Climatol., 1, 89–109, https://doi.org/10.1002/joc.1874, 2010.

Lau, N.-C., and Nath, M. J.: Model Simulation and Projection of European Heat Waves in Present-Day and Future Climates, Journal of Climate, 10, 3713–3730, https://doi.org/10.1175/JCLI-D-13-00284.1, 2014.

Lemonsu, A., Beaulant, A. L., Somot, S., and Masson, V.: Evolution of heat wave occurrence over the Paris basin (France) in the 21st century, Climate Research, 1, 75–91, https://doi.org/10.3354/cr01235, 2014.

Lenderink, G., van Ulden, A., van den Hurk, B., and van Meijgaard, E.: Summertime inter-annual temperature variability in an ensemble of regional model simulations: analysis of the surface energy budget, Climatic Change, S1, 233–247, https://doi.org/10.1007/s10584-006-9229-9, 2007.

Lhotka, O., and Kyselý, J.: Hot Central-European summer of 2013 in a long-term context, Int. J. Climatol., 14, 4399–4407, https://doi.org/10.1002/joc.4277, 2015.

Lhotka, O., and Kyselý, J.: Spatial and temporal characteristics of heat waves over Central Europe in an ensemble of regional climate model simulations, Clim Dyn, 9-10, 2351–2366, https://doi.org/10.1007/s00382-015-2475-7, 2015.

Lhotka, O., Kyselý, J., and Farda, A.: Climate change scenarios of heat waves in Central Europe and their uncertainties, Theor Appl Climatol, 3-4, 1043–1054, https://doi.org/10.1007/s00704-016-2031-3, 2018a.

Lhotka, O., Kyselý, J., and Plavcová, E.: Evaluation of major heat waves' mechanisms in EURO-CORDEX RCMs over Central Europe, Clim Dyn, 11-12, 4249–4262, https://doi.org/10.1007/s00382-017-3873-9, 2018b.

Liang, X.-Z., Kunkel, K. E., Meehl, G. A., Jones, R. G., and Wang, J. X. L.: Regional climate models downscaling analysis of general circulation models present climate biases propagation into future change projections, Geophys. Res. Lett., 8, https://doi.org/10.1029/2007GL032849, 2008.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023


Natural Hazards and Earth System Sciences Discussions

Lin, C., Kjellström, E., Wilcke, R. A. I., and Chen, D.: Present and future European heat wave magnitudes: climatologies, trends, and their associated uncertainties in GCM-RCM model chains, Earth Syst. Dynam., 3, 1197–1214, https://doi.org/10.5194/esd-13-1197-2022, 2022.

Lobell, D. B., Bonfils, C., and Duffy, P. B.: Climate change uncertainty for daily minimum and maximum temperatures: A model inter-comparison, Geophys. Res. Lett., 5, https://doi.org/10.1029/2006GL028726, 2007.

Lorenz, R., Jaeger, E. B., and Seneviratne, S. I.: Persistence of heat waves and its link to soil moisture memory, Geophys. Res. Lett., 9, n/a-n/a, https://doi.org/10.1029/2010GL042764, 2010.

Luber, G., and McGeehin, M.: Climate change and extreme heat events, American Journal of Preventive Medicine, 5, 429–435, https://doi.org/10.1016/j.amepre.2008.08.021, 2008.

Machard, A., Inard, C., Alessandrini, J.-M., Pelé, C., and Ribéron, J.: A Methodology for Assembling Future Weather Files Including Heatwaves for Building Thermal Simulations from the European Coordinated Regional Downscaling Experiment (EURO-CORDEX) Climate Data, Energies, 13, 3424, https://doi.org/10.3390/en13133424, 2020.

Mann, H. B.: Nonparametric Tests Against Trend, Econometrica, 3, 245, https://doi.org/10.2307/1907187, 1945.

Meehl, G. A., and Tebaldi, C.: More intense, more frequent, and longer lasting heat waves in the 21st century, Science (New York, N.Y.), 5686, 994–997, https://doi.org/10.1126/science.1098704, 2004.

Molina, M. O., Sánchez, E., and Gutiérrez, C.: Future heat waves over the Mediterranean from an Euro-CORDEX regional climate model ensemble, Scientific reports, 1, 8801, https://doi.org/10.1038/s41598-020-65663-0, 2020.

Mooney, P. A., Mulligan, F. J., and Fealy, R.: Evaluation of the Sensitivity of the Weather Research and Forecasting Model to Parameterization Schemes for Regional Climates of Europe over the Period 1990–95, Journal of Climate, 3, 1002–1017, https://doi.org/10.1175/JCLI-D-11-00676.1, 2013.

Nikulin, G., Kjellström, E., Hansson, U., Strandberg, G., and Ullerstig, A.: Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations, Tellus A 63A:41–55. https://doi:10.1111/j.1600-0870.2010.00466.x, 2011.

Ouzeau, G., Soubeyroux, J.-M., Schneider, M., Vautard, R., and Planton, S.: Heat waves analysis over France in present and future climate: Application of a new method on the EURO-CORDEX ensemble, Climate Services, 1–12, https://doi.org/10.1016/j.cliser.2016.09.002, 2016.

Perkins, S. E., Alexander, L. V., and Nairn, J. R.: Increasing frequency, intensity and duration of observed global heatwaves and warm spells, Geophysical Research Letters, 20, https://doi.org/10.1029/2012GL053361, 2012.

Perkins-Kirkpatrick, S. E., and Lewis, S. C.: Increasing trends in regional heatwaves, Nat Commun, 1, 3357, https://doi.org/10.1038/s41467-020-16970-7, 2020.

Petrovic, D., Fersch, B., and Kunstmann, H.: Droughts in Germany: performance of regional climate models in reproducing observed characteristics, Nat. Hazards Earth Syst. Sci., 12, 3875–3895, https://doi.org/10.5194/nhess-22-3875-2022, 2022.

Plavcová, E., and Kyselý, J.: Evaluation of daily temperatures in Central Europe and their links to large-scale circulation in an ensemble of regional climate models, Tellus A 63A:763–781. https:// doi:10.1111/j.1600-0870.2011.00514.x, 2011.

Plavcová, E., and Kyselý, J.: Temporal Characteristics of Heat Waves and Cold Spells and Their Links to Atmospheric Circulation in EURO-CORDEX RCMs, Advances in Meteorology, 1–13, https://doi.org/10.1155/2019/2178321, 2019.

Poumadère, M., Mays, C., Le Mer, S., and Blong, R.: The 2003 heat wave in France: dangerous climate change here and now, Risk analysis: an official publication of the Society for Risk Analysis, 6, 1483–1494, https://doi.org/10.1111/j.1539-6924.2005.00694.x, 2005.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Robine, J.-M., Cheung, S. L. K., Le Roy, S., van Oyen, H., Griffiths, C., Michel, J.-P., and Herrmann, F. R.: Death toll exceeded 70,000 in Europe during the summer of 2003, Comptes Rendus Biologies, 2, 171–178, https://doi.org/10.1016/j.crvi.2007.12.001, 2008.

Rousi, E., Fink, A. H., Andersen, L. S., Becker, F. N., Beobide-Arsuaga, G., Breil, M., Cozzi, G., Heinke, J., Jach, L., Niermann, D., Petrovic, D., Richling, A., Riebold, J., Steidl, S., Suarez-Gutierrez, L., Tradowsky, J. S., Coumou, D., Düsterhus, A., Ellsäßer, F., Fragkoulidis, G., Gliksman, D., Handorf, D., Haustein, K., Kornhuber, K., Kunstmann, H., Pinto, J. G., Warrach-Sagi, K., and Xoplaki, E.: The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective, Nat. Hazards Earth Syst. Sci., 5, 1699–1718, https://doi.org/10.5194/nhess-23-1699-2023, 2023.

Russo, S., Marchese, A. F., Sillmann, J., and Immé, G.: When will unusual heat waves become normal in a warming Africa?, Environ. Res. Lett., 5, 54016, https://doi.org/10.1088/1748-9326/11/5/054016, 2016.

Saeed, F., Almazroui, M., Islam, N., and Khan, M. S.: Intensification of future heat waves in Pakistan: a study using CORDEX regional climate models ensemble, Nat Hazards, 3, 1635–1647, https://doi.org/10.1007/s11069-017-2837-z, 2017.

Schneidereit, A., Schubert, S., Vargin, P., Lunkeit, F., Zhu, X., Peters, D. H. W., and Fraedrich, K.: Large-Scale Flow and the Long-Lasting Blocking High over Russia: Summer 2010, Monthly Weather Review, 9, 2967–2981, https://doi.org/10.1175/MWR-D-11-00249.1, 2012.

Seneviratne, S. I., Donat, M. G., Mueller, B., and Alexander, L. V.: No pause in the increase of hot temperature extremes, Nature Clim Change, 3, 161–163, https://doi.org/10.1038/nclimate2145, 2014.

Silva, P. S., Geirinhas, J. L., Lapere, R., Laura, W., Cassain, D., Alegría, A., and Campbell, J.: Heatwaves and fire in Pantanal: Historical and future perspectives from CORDEX-CORE, Journal of environmental management, 116193, https://doi.org/10.1016/j.jenvman.2022.116193, 2022.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., Huang, X., Wang, W., and Powers, J. A.: Description of the Advanced Research WRF Version 3; Tech. Rep. NCAR/TN-475+STR, NCAR TECHNICAL NOTE, University Corporation for Atmospheric Research: Boulder, CO, USA, p. 113, 2008.

Smid, M., Russo, S., Costa, A. C., Granell, C., and Pebesma, E.: Ranking European capitals by exposure to heat waves and cold waves, Urban Climate, 388–402, https://doi.org/10.1016/j.uclim.2018.12.010, 2019.

Stegehuis, A. I., Vautard, R., Ciais, P., Teuling, A. J., Jung, M., and Yiou, P.: Summer temperatures in Europe and land heat fluxes in observation-based data and regional climate model simulations, Clim Dyn, 2, 455–477, https://doi.org/10.1007/s00382-012-1559-x, 2013.

Štepánek, P., Zahradnícek, P., Farda, A., Skalák, P., Trnka, M., Meitner, J., and Rajdl, K.: Projection of drought-inducing climate conditions in the Czech Republic according to Euro-CORDEX models, Clim. Res., 2, 179–193, https://doi.org/10.3354/cr01424, 2016.

Stoelinga, M. T., Hobbs, P. V., Mass, C. F., Locatelli, J. D., Colle, B. A., Houze, R. A., Rangno, A. L., Bond, N. A., Smull, B. F., Rasmussen, R. M., Thompson, G., and Colman, B. R.: Improvement of Microphysical Parameterization through Observational Verification Experiment, B. Am. Meteorol. Soc., 12, 1807–1826, https://doi.org/10.1175/BAMS-84-12-1807, 2003.

Sun, B., Groisman, P. Y., Bradley, R. S., and Keimig, F. T.: Temporal Changes in the Observed Relationship between Cloud Cover and Surface Air Temperature, J. Climate, 24, 4341–4357, https://doi.org/10.1175/1520-0442(2000)013<4341:TCITOR>2.0.CO;2, 2000.

https://doi.org/10.5194/nhess-2023-91
Preprint. Discussion started: 19 June 2023

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.

Urban, A., Hanzlíková, H., Kyselý, J., and Plavcová, E.: Impacts of the 2015 Heat Waves on Mortality in the Czech Republic-A Comparison with Previous Heat Waves, International Journal of Environmental Research and Public Health, 12, 1562, https://doi.org/10.3390/ijerph14121562, 2017.

Valeriánová, A., Crhová, L., Holtanová, E., Kašpar, M., Müller, M., and Pecho, J.: High temperature extremes in the Czech Republic 1961–2010 and their synoptic variants, Theor Appl Climatol, 1-2, 17–29, https://doi.org/10.1007/s00704-015-1614-8, 2017.

van der Linden, P., and Mitchell, J. F. B.: ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project, Tech Rep. Met Office Hadley Centre, Exeter, 2009.

Varela, R., Rodríguez-Díaz, L., and deCastro, M.: Persistent heat waves projected for Middle East and North Africa by the end of the 21st century, PloS one, 11, e0242477, https://doi.org/10.1371/journal.pone.0242477, 2020.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernández, J., García-Díez, M., Goergen, K., Güttler, I., Halenka, T., Karacostas, T., Katragkou, E., Keuler, K., Kotlarski, S., Mayer, S., van Meijgaard, E., Nikulin, G., Patarčić, M., Scinocca, J., Sobolowski, S., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V., and Yiou, P.: The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project, Clim Dyn, 9-10, 2555–2575, https://doi.org/10.1007/s00382-013-1714-z, 2013.

Vichot-Llano, A., Martinez-Castro, D., Giorgi, F., Bezanilla-Morlot, A., and Centella-Artola, A.: Comparison of GCM and RCM simulated precipitation and temperature over Central America and the Caribbean, Theor Appl Climatol, 1-2, 389–402, https://doi.org/10.1007/s00704-020-03400-3, 2021.

Vogel, M. M., Zscheischler, J., Wartenburger, R., Dee, D., and Seneviratne, S. I.: Concurrent 2018 Hot Extremes Across Northern Hemisphere Due to Human-Induced Climate Change, Earth's Future, 7, 692–703, https://doi.org/10.1029/2019EF001189, 2019.

Wagner, S., and Kunstmann, H.: High resolution precipitation fields for the planning of urban drainage systems using WRF. SCC Annual Report, 2016.

Wang, P., Hui, P., Xue, D., and Tang, J.: Future projection of heat waves over China under global warming within the CORDEX-EA-II project, Clim Dyn, 1-2, 957–973, https://doi.org/10.1007/s00382-019-04621-7, 2019a.

Wang, P., Tang, J., Sun, X., Liu, J., and Juan, F.: Spatiotemporal characteristics of heat waves over China in regional climate simulations within the CORDEX-EA project, Clim Dyn, 1-2, 799–818, https://doi.org/10.1007/s00382-018-4167-6, 2019b.

Warscher, M., Wagner, S., Marke, T., Laux, P., Smiatek, G., Strasser, U., and Kunstmann, H.: A 5 km Resolution Regional Climate Simulation for Central Europe: Performance in High Mountain Areas and Seasonal, Regional and Elevation-Dependent Variations, Atmosphere, 11, 682, https://doi.org/10.3390/atmos10110682, 2019.

Wedler, M., Pinto, J. G., and Hochman, A.: More frequent, persistent, and deadly heat waves in the 21st century over the Eastern Mediterranean, The Science of the total environment, 161883, https://doi.org/10.1016/j.scitotenv.2023.161883, 2023.

Zeng, X.-M., Wang, M., Zhang, Y., Wang, Y., and Zheng, Z.: Assessing the Effects of Spatial Resolution on Regional Climate Model Simulated Summer Temperature and Precipitation in China: A Case Study, Advances in Meteorology, https://doi.org/10.1155/2016/7639567, 2016.