Present and future European heat wave magnitudes: climatologies, trends, and their associated uncertainties in GCM-RCM model chains

Changgui Lin[1,2], Erik Kjellström[1,3], Renate Anna Irma Wilcke[1], and Deliang Chen[2]

[1]Rossby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[2]Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
[3]Department of Meteorology and the Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

Correspondence: Changgui Lin (mapulynn@gmail.com)

Abstract. This study investigates present and future European heat wave magnitudes, represented by the Heat Wave Magnitude Index-daily (HWMId), for regional climate models (RCMs) and their driving global climate models (GCMs) over Europe. A subset of the large EURO-CORDEX ensemble is employed to study sources of uncertainties related to the choice of GCMs, RCMs, and their combinations.

We initially compare the evaluation runs of the RCMs driven by ERA-interim reanalysis to E-OBS, finding that the RCMs can capture most of the observed spatial and temporal features of HWMId. With their higher resolution compared to GCMs, RCMs can reveal spatial features of HWMId associated with small-scale processes (e.g., orographic effects); moreover, RCMs represent large scale features of HWMId in a satisfactory way (e.g., by reproducing the general west-east gradient). Our results indicate a clear added value of the RCMs compared to their driving GCMs. Forced with the emission scenario RCP8.5, all the GCM and RCM simulations consistently project a rise in HWMId at an exponential rate. However, the climate change signals projected by the GCMs are generally attenuated when downscaled by the RCMs, with the spatial pattern also altered.

The uncertainty in a simulated future change of heat wave magnitudes following global warming can be attributed almost equally to the difference in model physics (as represented by different RCMs) and to the driving data associated with different GCMs. Regarding the uncertainty associated with RCM choice, a major factor is the different representation of the orographic effects. No consistent spatial pattern in the ensemble spread associated with different GCMs is observed between the RCMs, suggesting GCMs' uncertainties are transformed by RCMs in a complex manner due to the nonlinear nature of model dynamics and physics.

In summary, our results support the use of dynamical downscaling for deriving regional climate realization regarding heat wave magnitudes.

1 Introduction

Heat waves can have adverse effects on health and have been reported to lead to excessive mortality rates among people in many regions of the world (Guo et al., 2017). In Europe, the heat waves in summer 2003 (Robine et al., 2008) and 2010

(Barriopedro et al., 2011) are prominent examples. Even in high-latitude areas, such as Scandinavia, heat waves can lead to excess mortality as reported by Åström et al. (2019) for the long and warm Scandinavian summer 2018 (Wilcke et al., 2020).

25 In addition to health problems atmospheric heat waves are often related to water shortages, a decline in agricultural production, and increased risk of forest fires or dieback, all of which can have severe impacts both on natural ecosystems and human society (IPCC, 2014).

The Intergovernmental Panel on Climate Change (IPCC) concluded that the frequency and duration of heat wave lengths have increased in the observed past on a global scale (IPCC, 2018; Seneviratne et al., 2021). Part of the observed changes in

30 frequency and intensity of daily temperature extremes on the global scale since the mid-20th century has been attributed to increasing anthropogenic forcing of the climate system. Furthermore, projections for the future under scenarios of increasing greenhouse gas forcing indicate a continued increase in both intensity and in duration of heat waves (Molina et al., 2020). To alleviate future problems, adaptive measures are needed, even under scenarios with strong mitigation (IPCC, 2022). Climate change adaptation, in turn, requires relevant information about changes both in geographical and temporal extent as well as in

35 intensity and duration of heat waves.

Climate model projections constitute the most prominent information about future climate. International coordinated Climate Model Intercomparison Projects (CMIP), such as CMIP5 (Taylor et al., 2012) and CMIP6 (Eyring et al., 2016), provide large ensembles of global climate model (GCM) projections. GCMs, however, are most often run at coarse horizontal resolutions, implying that they have a crude representation of relevant local and regional processes and often come with strong

40 biases at the regional scale (Luo et al., 2020). As a remedy, to improve the quality of the simulated climate and add value compared to GCMs, empirical statistical downscaling employing a wide range of approaches (Benestad et al., 2008, 2018; Hertig et al., 2019; Soares et al., 2019) and dynamical downscaling using a regional climate model (RCM) (Torma et al., 2015; Rummukainen, 2016; Strandberg and Lind, 2021) are used, each with its relative strengths. This study focuses on the dynamical downscaling. Operated at higher resolution than their driving GCMs, RCMs can more realistically represent orog-

45 raphy, land-sea contrasts, and atmospheric processes such as mid-latitude cyclones. Large ensembles of RCM climate change projections have been produced for different continents under the auspices of CORDEX (the CoOrdinated Regional climate Downscaling EXperiment, Giorgi et al., 2009). Specifically, for Europe, joint efforts in EURO-CORDEX (Jacob et al., 2020) have resulted in more than 100 RCM projections at 0.11° grid spacing under a range of different future scenarios that are now being used for building climate services (Sørland et al., 2020; Rennie et al., 2021).

50 Even if RCMs add value compared to GCMs, they do not come without biases of their own. Notably, when given input from coarse-resolution GCMs that have biases in their representation of the large-scale circulation and sea-surface conditions, RCMs also show biases (e.g., Jacob et al., 2007; Vautard et al., 2020). As RCMs often describe physical processes differently, not only biases in the historical climate, but also future climate change signals can differ from those of the underlying GCM (Coppola et al., 2021). As an example, most EURO-CORDEX RCMs have a more rudimentary treatment of aerosols and their

55 interaction with radiative fluxes and clouds than the CMIP5 GCMs they are forced with. Consequently, discrepancies in future trends for GCM-RCM chains as reported by (Sørland et al., 2018) have been suggested to be related to differences in aerosols and their impact on downwelling shortwave radiation (Jerez et al., 2021).

A univocal and optimal definition of heat wave can be for debate depending on impacts of interest (Perkins and Alexander, 2013; Horton et al., 2016). Most existing heat wave-related indicators describe only a single characteristic of heat waves. The Heat Wave Magnitude Index-daily (HWMId, Russo et al., 2015), a dimensionless magnitude that was designed to take into account both heat wave duration and intensity, represents an integrative approach in classifying heat waves. Indeed, it has been successfully used in a growing number of heat wave studies (e.g., Russo et al., 2016; Zampieri et al., 2016; Ceccherini et al., 2017; Dosio et al., 2018; Molina et al., 2020). Being fairly established, therefore, the HWMId is utilized here for representing heat wave magnitudes.

The aim of this study is fourfold. First, we examine GCM-RCM combinations for Europe in the EURO-CORDEX collection to investigate to what degree the RCMs can represent heat waves in the historical climate when forced by reanalysis and GCMs, respectively. Second, we investigate if there is any added value in the representation of heat waves in the RCMs compared to the driving GCMs. Third, we investigate to what extent the RCMs modify the climate change signal of HWMId compared to the driving GCMs. Fourth, we explore the sources of uncertainties related to the choice of GCMs, RCMs, and their combinations.

2 Data and methods

2.1 Heat Wave Magnitude Index-daily (HWMId)

The HWMId is described as the maximum magnitude of heat waves occurring in a year, where a heat wave is defined as a period of at least three consecutive days with $T_{\max}$ above a percentile-based daily threshold for a reference period. Specifically, for a given day-of-year d (from 1 to 366), the threshold is the 90th percentile of the set of data A_d defined by

$$A_d = \cup_{y \in Y_{\mathrm{ref}}} \cup_{i \in W_d^{-15,15}} T_{\max,y,i},\tag{1}$$

where $\cup$ denotes the union of sets; Y_{ref} represents the years within the reference period; $W_d^{-15,15}$ is the 31-day window centered at day d; and, $T_{\max,y,i}$ is the daily $T_{\max}$ of day i in year y.

For each day in an identified heat wave the daily magnitude, M_d, is calculated following:

$$M_d = \begin{cases} \dfrac{T_{\max,d} - T_{\max,\mathrm{ref},25p}}{T_{\max,\mathrm{ref},75p} - T_{\max,\mathrm{ref},25p}} & \text{if } T_{\max,d} > T_{\max,\mathrm{ref},25p} \\ 0 & \text{otherwise} \end{cases},\tag{2}$$

where $T_{\max,d}$ is the daily $T_{\max}$ of day d, and $T_{\max,\mathrm{ref},25p}$ and $T_{\max,\mathrm{ref},75p}$ are the 25th and 75th percentiles, respectively, of the $T_{\max}$ time series over the reference period. Here, $T_{\max,\mathrm{ref},25p}$ and $T_{\max,\mathrm{ref},75p}$ (utilized for normalization) differ slightly from Russo et al. (2015) who used annual data. This modification hardly influences the usage of HWMId, and somehow makes calculation more stable. M_d is calculated at each grid point. According to the definition (Eq. 2), a daily magnitude M_d equal to n indicates that the temperature anomaly of day d with respect to $T_{\max,\mathrm{ref},25p}$ is n times the climatological interquartile range (IQR) within the reference period.

Table 1. The RCMs and the driving GCMs that form the simulation matrix in the study.

Model	Institute	Resolution	Realization
RCM:			
HIRHAM5	Danish Meteorological Institute	EUR-11[*]	
RACMO22E	Royal Netherlands Meteorological Institute	EUR-11	
RCA4	Swedish Meteorological and Hydrological Institute	EUR-11	
REMO2015	Climate Service Center Germany	EUR-11	
GCM:			
EC-EARTH	Irish Centre for High-End Computing	T159L62	r12i1p1
HadGEM2-ES	Met Office Hadley Centre	N96L38	r1i1p1
NorESM1-M	Norwegian Climate Centre	F19L26	r1i1p1

[*] EURO-CORDEX rotated 0.11° (about 12.5 km) grid.

For each heat wave, the magnitude is calculated as the sum of the daily magnitudes of the constituent days. Finally, the annual maximum is identified from the individual heat waves in a year. This number is the HWMId analyzed here. A demonstration with specific examples is available in the supplementary information for further assistance in understanding the HWMId.

2.2   Climate model simulations and other data

The large number of GCM-RCM combinations available from EURO-CORDEX allows us to examine RCMs' behaviors (e.g., signal modification and uncertainty transformation) within the downscaling process for simulating heat wave magnitudes. For that we used only a subset of the available ensemble to gain a full GCM-RCM matrix without gaps, to ensure a fair comparison after aggregating along either the GCM dimension or the RCM dimension. That means, we chose three GCMs that have been downscaled by four RCMs returning a 3×4 matrix of climate simulations, as listed in Table 1. The driving GCM simulations, from CMIP5, include runs with historical forcing (to 2005) forced with historical forcings and projection runs (since 2006) forced with representative concentration pathways (RCPs). Here, we focus only on RCP8.5 because there are fewer EURO-CORDEX simulations forced by other RCPs, hardly available for formulating such a full simulation matrix. Three periods were defined for the calculation of the climate change signals in simulated heat wave magnitudes (ΔHWMId), termed "recent past" (1981–2020), "nearest decades" (2021–2060), and "end of the century" (2061–2100). The uncertainty across simulations is roughly described here by the spread of values (maximum − minimum) considering the limited size of the simulation matrix. The spread along the RCM dimension represents the uncertainty associated with the RCM model physics while the spread along the GCM dimension corresponds to the uncertainty associated with the driving GCM simulations.

In addition to the GCM-RCM combinations that compose the matrix, the EURO-CORDEX initiative also provides a set of evaluation runs, for which the participating RCMs are forced with boundary conditions from the ERA-Interim reanalysis (Dee et al., 2011). This type of simulation, which is often referred to as perfect-boundary conditions run, allows for an in-depth

comparison with the observed climate including also its temporal evolution. In the first part of the study, we compared these evaluation runs to observations, as well as to ERA-Interim. The observational data used in the study are from the E-OBS daily gridded data set version 20.0e of the European Climate Assessment & Data (Cornes et al., 2018, ECA&D, www.ecad.eu), which covers Europe at a 0.1° regular grid spacing for the period from 1950 to July 2019. When RCMs are driven by ERA-Interim the reference period for HWMId is 1989–2008, and when driven by GCMs it is 1981–2010. For each dataset, HWMId was calculated upon the original grid points. The area of investigation is bounded by 10°W–30°E and 35°N–70°N and is land-only.

Mean bias error (MBE), root mean square error (RMSE), and Pearson correlation coefficient (r) were adopted as performance indicators for a model in simulating HWMId compared to E-OBS. They quantify the degree of an overall overestimation/underestimation, the degree of closeness in values, and the association in variations, respectively. When these indicators were applied for spatial patterns, they were calculated after HWMId values of all the datasets accounted for other than ERA-Interim were conservatively remapped to the ERA-Interim grids. RMSE and r were also used to determine the similarity in spatial patterns between simulations.

In order to better understand the underlying processes of the simulated ΔHWMId, a preliminary effort was made by exploring the simulated climate change signals in dry conditions and their connection with HWMId. We investigated atmospheric and soil dry conditions, as represented by dry days (with precipitation < 1 mm) and precipitation $-$ evaporation ($P - E$), respectively. As a background of warming, the simulated climate change signals in annual mean $T_{\max}$ were also examined.

3 Results

3.1 Evaluation of reanalysis-driven RCMs in simulating historical heat waves

Among various aspects of heat waves, we want to answer the following questions: i) Does the spatial pattern of the climatological mean reveal observed local information? and ii) Does the regional mean show the same signal/response of large-scale climate variations as observations?

Figure 1a shows the comparison of the spatial pattern of the climatological mean of HWMId over 1989–2008 for the evaluation runs of the RCM considered. E-OBS indicates a clear west-east gradient, with HWMId values in western coastal areas on average 1.2 higher than those in eastern areas. The same pattern is confirmed by reanalysis data (ERA-Interim), which is used as boundary and initial conditions for the evaluation RCM simulations. The RCM evaluation runs can reproduce the observed west-east gradient. Though all RCM simulations agree on the general spatial pattern, they do differ from each other in representing details of some local features. For example, among the four RCMs, RCA4 has particularly high values over central Europe, while REMO2015 is the only RCM that captures the observed level of HWMId values (7–9) over eastern Europe. As the RCMs were driven by the same reanalysis, the differences among the RCM simulations should be related to differences in model physics which is an important source of uncertainty in simulating extreme heat waves. Ignoring the spatial pattern and focusing on the percentage of the land area exceeding certain HWMId values (Fig. 1b), we find a close agreement between the RCM simulations, although a slight overestimation was found at the high-end tail of the HWMId distribution (HWMId$\geq$9 or
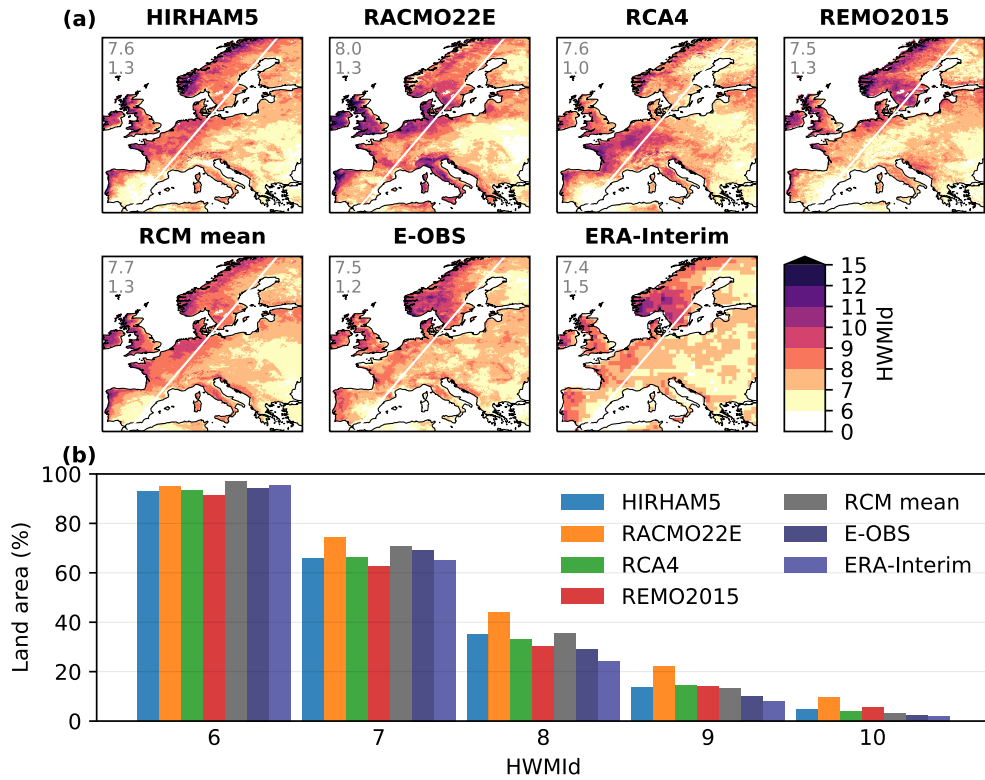
Figure 1. Climatological mean HWMId over 1989–2008 in the four RCMs' evaluation runs, the multi-RCM mean, E-OBS, and ERA-Interim: (**a**) spatial pattern, and (**b**) percentage of the land area exceeding certain HWMId levels (HWMId≥6, 7, 8, 9, or 10). The two numerals in each map show the area-weighted average over the entire domain and the difference between western and eastern parts divided by the white line, respectively.

10) compared to E-OBS and ERA-Interim. The overestimation for RACMO22E is apparent for most of HWMId values (7–10).
140 According to RMSE and r, REMO2015 has the best performance among the selected four RCMs in representing the spatial pattern of the climatological mean HWMId (Table 2). We also note that the multi-RCM mean has a generally smaller RMSE and a higher r than most of the individual RCMs, probably due to a compensation of deficiencies of each RCM in representing different processes.

Averaged in space, the RCM evaluation runs reproduce generally, but not perfectly, the temporal evolution of the E-OBS
145 HWMId as shown in Fig. 2a. Years with high HWMId values (1994, 2003, 2006, and 2007) are captured by all the RCMs. However, the RCMs fail in reproducing the ranking of these years by occasionally overestimating or underestimating the HWMId values. For example, RACMO22E overestimates HWMId excessively in 1995 as does RCA4 in 1997. Reflecting on the distribution function as shown in Fig. 2b, the HWMId values simulated by the RCMs show an overestimation of the 75th percentile and the median, while deviations for the 25th percentile are smaller with REMO2015 showing an underestimation
150 and HIRHAM5 and RACMO22E overestimations. Both HIRHAM5 and the multi-RCM mean show a better representation

Table 2. Spatial MBE, RMSE, and r for the climatological mean HWMId over 1989–2008 in ERA-Interim, the four RCMs' evaluation runs, and the multi-RCM mean (corresponding to Fig. 1), with E-OBS as reference.

	ERA-Interim	HIRHAM5	RACMO22E	RCA4	REMO2015	RCM mean
MBE	−0.13	0.12	0.45	0.08	0.00	0.17
RMSE	0.55	0.89	1.17	1.03	0.76	0.74
r^*	0.86	0.70	0.60	0.53	0.79	0.77

* All are statistically significant at $p < 0.01$.
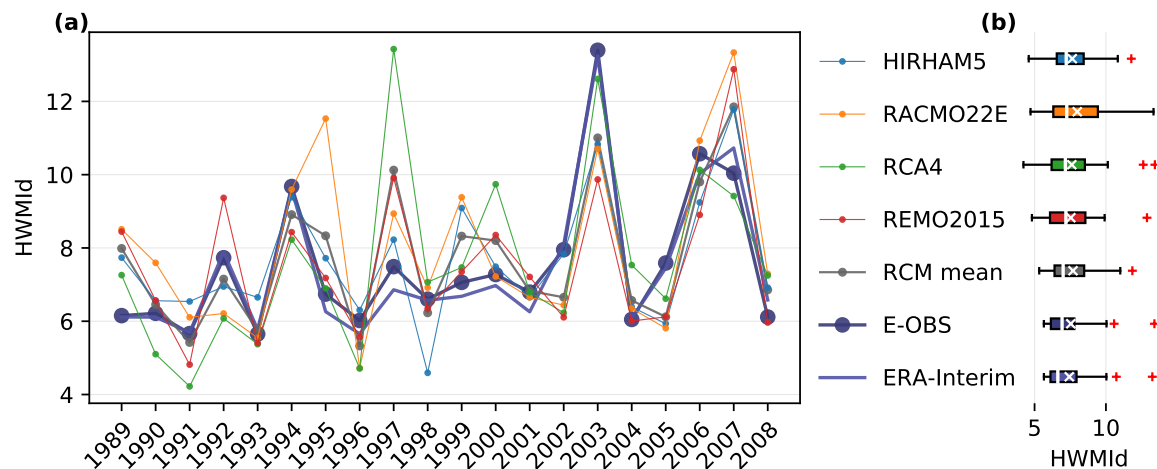
Figure 2. Regional mean HWMId from 1989 to 2008 in the four RCMs' evaluation runs, the multi-RCM mean, E-OBS, and ERA-Interim: (**a**) temporal evolution, and (**b**) corresponding box-plot. In each box-plot, the box limits represent lower and upper quartiles; the white line and 'x' marker inside the box are the median and the mean, respectively; and the outliers marked with red plus signs outside of the whisker lie at least 1.5 times the IQR away from the box limits.

of the observed IQR but shifted to higher values and, in addition, with a shorter right tail, which means too weak heat wave extremes. At the same time, and in contrast to E-OBS and ERA-Interim, all the RCMs have individual years with HWMId values lower than the observations, indicating that there are also underestimations of heat waves. Among the selected RCMs, HIRHAM5 outperforms the others in reproducing the temporal evolution of the regional mean (Table 3). Like the case of spatial measures, the multi-RCM mean has also a generally smaller RMSE and a higher r of temporal measures than most of the individual RCMs.

3.2 Evaluation of heat waves in GCM-driven RCMs under the recent past climate

A similar comparison, as for the ERA-Interim-driven runs discussed above, was conducted for the GCM-driven runs but focusing only on the spatial patterns. The influence of the shift of driving data from ERA-interim to GCM simulations can

Table 3. Similar to Table 2 but of temporal measures for the time series of regional mean HWMId from 1989–2008 (corresponding to Fig. 2).

	ERA-Interim	HIRHAM5	RACMO22E	RCA4	REMO2015	RCM mean
MBE	−0.11	0.10	0.45	0.07	−0.00	0.16
RMSE	0.35	1.22	1.84	1.75	1.53	1.25
r^*	0.99	0.78	0.64	0.68	0.68	0.78

* All are statistically significant at $p < 0.01$.

be seen by comparing Table 4 with Table 2, where the reference period as the basis of HWMId is identically 1989–2008. The GCMs perform poorly in simulating spatial characteristics of HWMId compared to ERA-Interim (RMSE: 1.02–1.26 vs. 0.55, and r: 0.17–0.43 vs. 0.86), whereas GCM-driven RCM simulations show similar results as the ERA-Interim-driven ones, which becomes in particular obvious for RMSE (0.97–1.37 vs. 0.76–1.17). This implies an improvement in the representation of the spatial characteristics of HWMId introduced by the dynamical downscaling with RCMs. Moreover, the improvement seems to depend strongly on the model (e.g., RCA4 has a relatively weak spatial r).

Figure 3 presents the spatial pattern of climatological mean HWMId under the recent past climate. With reference to Fig. 1, a shift in time for both the reference period as the basis of HWMId and the period of investigation should be noted. As expected, due to the relatively large overlap in time, the observed spatial distribution under the recent past climate, as revealed by E-OBS (Fig. 3), is similar to that in the evaluation period (Fig. 1); but the shift in time has led to an increase from on average about 7.5 to about 8.0 on a continental basis with an increase mainly in Mediterranean parts of Europe. The GCMs capture some of the observed spatial pattern but miss out both in detailed structure and amplitude (Fig. 3). The observed west-east gradient is hardly seen in the GCM simulations: EC-EARTH and HadGEM2-ES report no difference between the western and eastern parts (divided by the white line on the map), while NorESM1-M shows a west-east gradient of 1.3, which is even higher than for E-OBS but simulates excessively high HWMId values in the easternmost part of the domain. The downscaling with RCMs improves the representation of the observed spatial pattern by reproducing the west-east gradient, as well as revealing small-scale processes such as orographic effects which are in no way represented by the GCMs due to their coarse resolution. In general, the RCMs (other than a few cases of RACMO22E and RCA4) show smaller RMSE and higher r compared to their driving GCMs (Table 5). Similar to the evaluation runs (Table 2), the multi-RCM (row) means driven by a GCM simulation show a better performance compared to most of the individual RCMs. This is also true for the GCM dimension; i.e., for each RCM, the ensemble mean across the three driving GCMs outperforms the individual ensemble members.

Furthermore, we also looked into the differences along the GCMs in representing HWMId and how the RCMs respond to them when driven by these GCMs. We want to point out that there is no accordance on which GCM that performs best among the three in reproducing the spatial pattern of HWMId in E-OBS when considering different performance indicators; e.g., EC-EARTH has the smallest MBE/RMSE, whereas HadGEM2-ES has slightly higher r (Table 5). Meanwhile, the "best" GCM according to a performance indicator may not necessarily secure a best downscaling among the different driving GCMs for an RCM. We also observed a large spread in the regional average across the GCMs, which is reduced by all RCMs. Moreover, for

Table 4. Spatial MBE, RMSE, and r for the climatological mean HWMId over 1989–2008 (with reference period also as 1989–2008) in the selected GCM simulations and the GCM-driven simulations that compose the RCM simulation matrix (columns for RCMs and rows for their driving GCMs), with E-OBS as reference. For each statistic, an additional row is given for the column mean (i.e., along the GCM dimension) of the driving GCM simulations and the RCM simulation matrix.

	GCM	HIRHAM5	RACMO22E	RCA4	REMO2015	RCM mean
MBE:						
EC-EARTH	−0.21	−0.09	0.19	0.17	−0.21	0.02
HadGEM2-ES	0.17	0.13	0.86	0.36	−0.40	0.25
NorESM1-M	0.49	0.08	0.36	0.54	−0.06	0.24
Model mean	0.02	0.04	0.47	0.37	−0.22	
RMSE:						
EC-EARTH	1.15	1.01	1.12	1.31	1.07	0.94
HadGEM2-ES	1.02	0.97	1.35	1.26	0.97	0.88
NorESM1-M	1.26	1.05	1.37	1.34	1.03	0.98
Model mean	0.95	0.85	1.08	1.10	0.87	
r^*:						
EC-EARTH	0.17	0.51	0.43	0.27	0.45	0.50
HadGEM2-ES	0.43	0.62	0.40	0.34	0.62	0.59
NorESM1-M	0.42	0.48	0.30	0.13	0.47	0.46
Model mean	0.36	0.61	0.45	0.33	0.60	

[*] All are statistically significant at $p < 0.01$.

each RCM driven by the three GCMs, the downscaling behaves consistently despite the large difference in the spatial pattern of HWMId between the driving GCMs. This can be reflected by the fact that, for each RCM, we observed lower RMSE and higher r (of spatial measures) between the RCM simulations than between its driving GCMs (Table S1). Likewise, when driven by one GCM, the simulations of different RCMs also most often tend to be more similar to each other than to the driving GCM (i.e., higher RMSE and lower r within the "GCM" column compared to other columns in Table S2). Thus, it is interesting to explore the influences of driving data versus model physics on the uncertainty for the RCMs in simulating heat wave magnitudes under the recent past climate.

Along the RCM dimension of the matrix (i.e., RCMs with a same driving GCM), the ensemble spread of HWMId values is on average close to one fifth of the ensemble mean on a continental basis (Fig. 4). A slightly lower spread/mean ratio is observed along the GCM dimension (Fig. 5), indicating that the uncertainties associated with driving data are of similar magnitude as those associated with model physics. However, different features can be observed in the spatial pattern of the ensemble spread along the two dimensions. As presented in Fig. 4, the ensemble spread along the RCM dimension shows
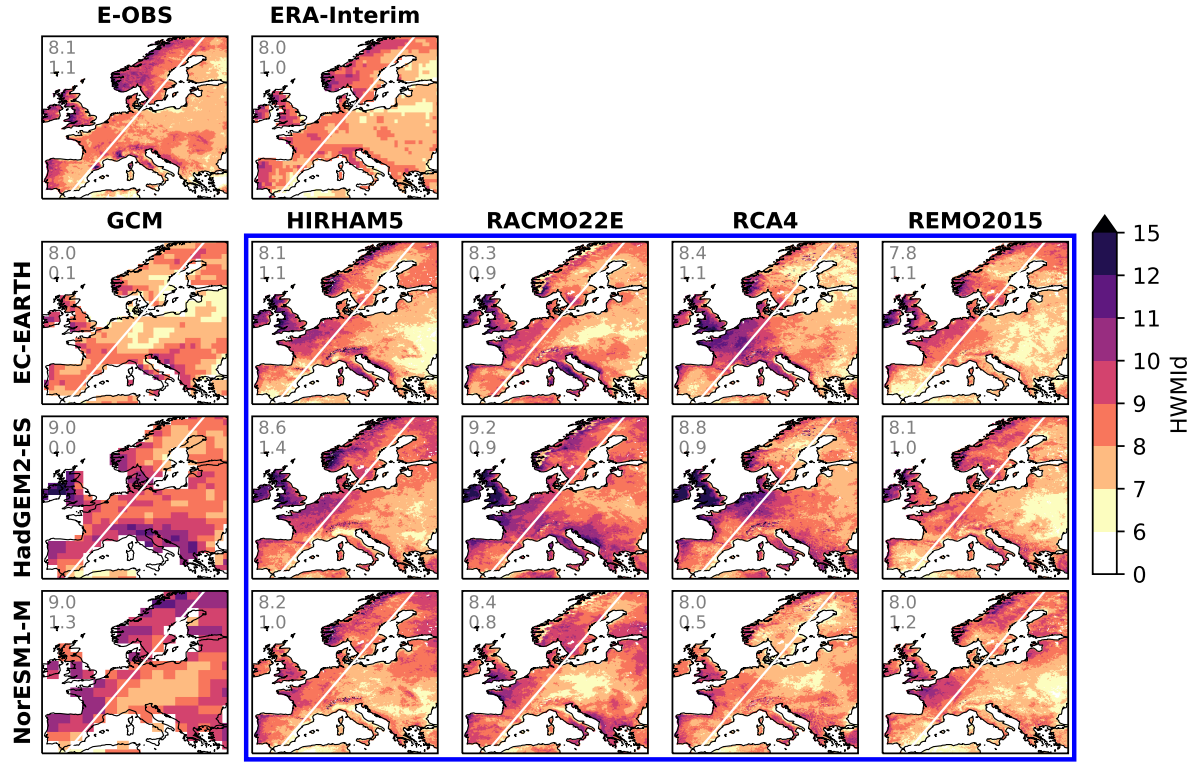
Figure 3. Climatological mean HWMId under the recent past climate (1981–2020) in the RCM simulation matrix (inside the red rectangle; columns for RCMs and rows for their driving GCMs) as well as the driving GCM simulations. Data from E-OBS and ERA-Interim are also shown for comparison. The two numerals in each map show the area-weighted average over the entire domain and the difference between western and eastern parts divided by the white line, respectively.

high values mostly in mountainous areas such as the Scandinavian Mountains and the Alps, suggesting the disagreement in
200 the orographic effects on heat wave magnitudes across the RCMs as one of the major sources of uncertainty that is associated
with model physics. Aggregating on the GCM dimension (i.e., calculating mean and spread for each RCM with different
GCMs), the ensemble means (first row of Fig. 5) reveal a similar spatial pattern, whereas the spreads (second row of Fig. 5)
show considerable differences in the spatial pattern. A quantitative analysis regarding the similarity based on RMSE and r
is presented in Table S3. Higher r values between RCMs than between any RCM and its driving GCMs, in combination
205 with lower RMSE scores (in all but one), indicate that the RCMs tend to converge in the ensemble mean and spread along
the GCM dimension. Compared to the ensemble mean, however, the r values are much lower for the ensemble spread. This
observation implies that the uncertainties of GCMs in simulating heat wave magnitudes are not simply inherited by RCMs but
are transformed in a nonlinear manner due to the complex model dynamics and physics.

Table 5. Similar to Table 4 but for the recent past climate (1981–2020, with the reference period of HWMId as 1981–2010).

	GCM	HIRHAM5	RACMO22E	RCA4	REMO2015	RCM mean
MBE:						
EC-EARTH	−0.14	0.03	0.19	0.26	−0.31	0.05
HadGEM2-ES	0.91	0.47	1.07	0.71	−0.31	0.55
NorESM1-M	0.95	0.06	0.31	−0.08	−0.15	0.04
Model mean	0.40	0.18	0.52	0.30	−0.17	
RMSE:						
EC-EARTH	1.05	1.03	1.07	1.19	1.02	0.91
HadGEM2-ES	1.40	1.15	1.48	1.46	0.97	1.09
NorESM1-M	1.51	1.00	1.24	1.08	1.08	0.95
Model mean	1.05	0.93	1.09	1.06	0.92	
r^*:						
EC-EARTH	0.37	0.61	0.51	0.50	0.59	0.62
HadGEM2-ES	0.45	0.57	0.55	0.38	0.59	0.59
NorESM1-M	0.35	0.48	0.27	0.34	0.44	0.48
Model mean	0.42	0.61	0.52	0.47	0.60	

[*] All are statistically significant at $p < 0.01$.

3.3 Simulated future change in heat wave magnitudes

Change in the climatological mean HWMId relative to the recent past climate was investigated for two future periods within the century: the nearest decades (2021–2060) and the end of the century (2061–2100). Figure 6 presents the spatial pattern in each of the simulations. The HWMId increases strongly at the end of the century, with approximately more than a factor of three compared the nearest decades. Otherwise no notable difference is observed in the spatial pattern according to the spatial r (Fig. 6b) for the two periods. Similar to the results of the recent past climate (Sect. 3.2), a large spread in ΔHWMId is seen across the GCMs (5.8–12.8 and 21.6–48.4 for the nearest decades and the end of the century, respectively; Fig. 6). The RCMs decrease the spread (4.9–8.7 and 15.0–31.9 for the nearest decades and the end of the century, respectively; Fig. 6), reflecting less change compared to the driving HadGEM2-ES and NorESM1-M. The spatial pattern of ΔHWMId among the RCM simulations is similar both along the RCM dimension and the GCM dimension: indicating a stronger rise in northern and southern Europe but comparatively moderate for the central domain. This partly differs from the GCM simulations, which tend to show a more pronounced south-north gradient in the increase in HWMId. It also differs from the climatological mean HWMId under the recent past climate, which displays a generally west-east gradient (Fig. 1). The ensemble spread, along either the RCM dimension (Fig. 7) or the GCM dimension (Fig. 8), has a magnitude most often exceeding half of the ensemble
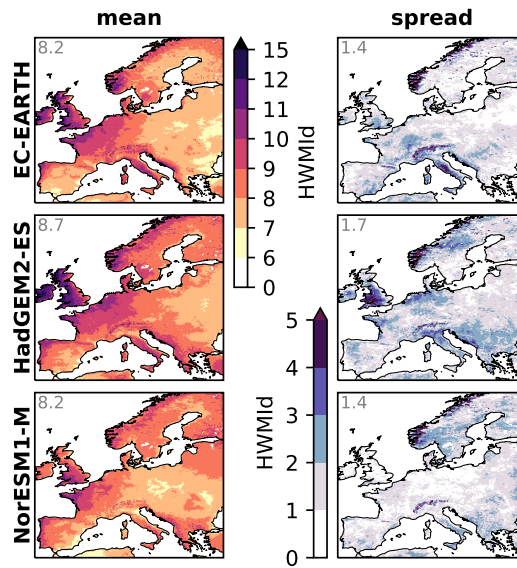
Figure 4. Ensemble mean and spread (maximum − minimum) of simulated HWMId under the recent past climate (1981–2020) for each row in the RCM simulation matrix indicating the uncertainty associated with different RCMs. The numeral in each map shows the area-weighted average.
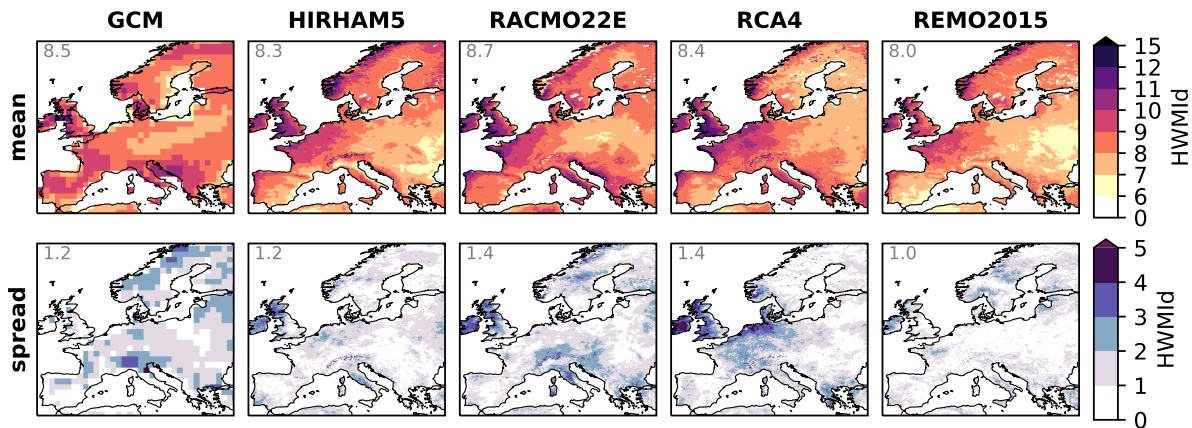
Figure 5. As Fig. 4 but for the columns of the RCM simulation matrix indicating the uncertainty associated with driving data (boundary and initial conditions). Ensemble mean and spread of the driving GCM simulations are also presented for comparison.

mean on a continental basis and to some extent follows the spatial distribution of the ensemble mean (with spatial r from 0.50 to 0.92; not shown). This suggests that the driving GCMs and the RCMs contribute about equally to the uncertainty in simulating climate change in heat wave magnitudes. For the ensemble spread, however, high values in mountainous areas are

225
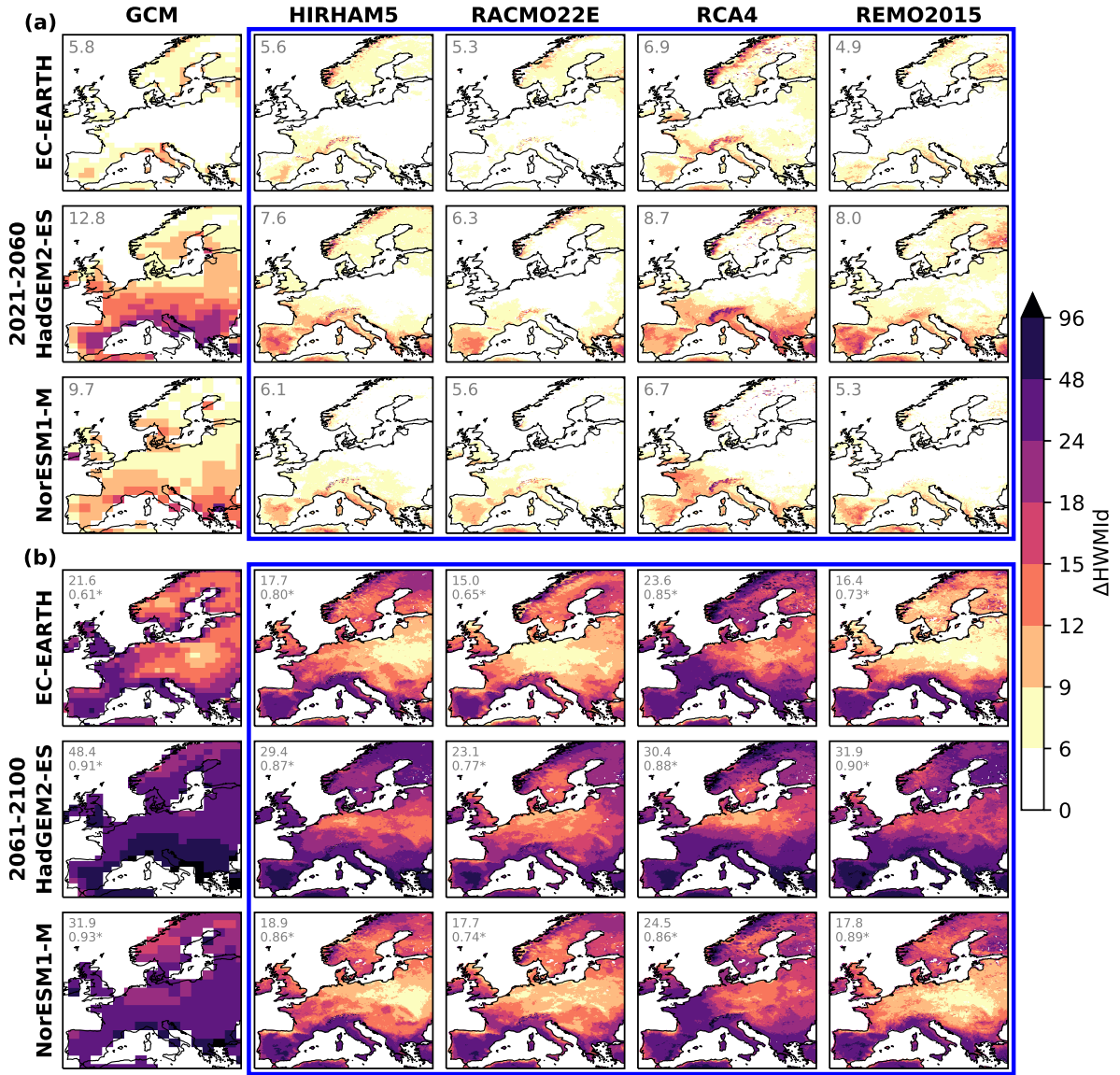
Figure 6. Change in climatological mean HWMId in the nearest decades (a: 2021–2060) and in the end of the century (b: 2061–2100) relative to the recent past climate (Fig. 3). The first numeral in each map shows the area-weighted average, and the second (only presented in b) is the spatial r between the two periods (with '*' indicating statistically significant at $p < 0.05$)

more common along the RCM dimension than along the GCM dimension. This underlines again that processes related to orography are a major source of uncertainty related to future changes in heat wave magnitudes in Europe.

The percentages of the land area exceeding certain HWMId levels (Fig. 9) provide us with a special form of survival function (complementary cumulative distribution function) to investigate the probability distribution of climatological mean HWMId
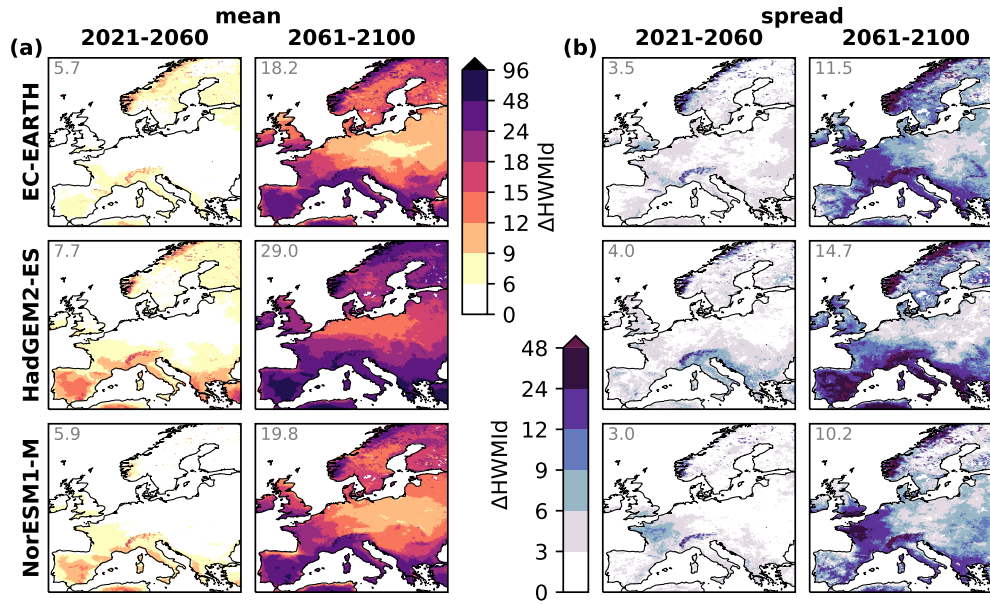
Figure 7. As Fig. 4 but for the climate change in HWMId.

230 values across the domain regardless of geophysical location. All model simulations can reproduce the E-OBS data probability distribution of HWMId values (solid lines) within ±15 %. A considerable rise in the future is projected by all the simulations. In the end of the century (2061–2100), the entire domain is projected to reach HWMId values higher than those experienced until now. The rise occurs across the whole probability range at an approximately exponential rate; i.e., increases between every two neighbor periods are approximately the same as mapped to the logarithmic scale applied to Fig. 9. The rise is even stronger for

235 the tail of HWMId distribution. As indicated in Fig. 6 the RCMs show a tendency to smoothen the change signal in the driving GCM simulations, particularly when driven by HadGEM2-ES and NorESM1-M. The only exception is RCA4 downscaling EC-EARTH. Again, we note that the RCMs generally are more alike than their driving GCMs, though RCA4 shows a pattern closest to the GCMs. Finally, we observe that the spread between the RCMs, and hence the uncertainty, increase a lot over time.

240 Probability distributions are also investigated for the region-wide annual HWMId values in the defined three periods, as shown in Fig. 10. As revealed by E-OBS and ERA-Interim, the region-wide HWMId for the recent past climate is represented by a positively-skewed distribution, or rather a quasi-log-normal distribution (as demonstrated by the quasi-normal shape of the distribution in Fig. 10b with a logarithmic-scale x-axis). As discussed for the ERA-Interim driven simulations (Fig. 2), also the GCM-driven RCMs can to some extent capture the shape of the distribution as well as the median, but with a wider value range

245 for each simulation. Most of the simulations keep the shape of distribution as the values increase in the future (especially for the nearest decades), whereas some (e.g., HadGEM2-ES, RCA4 driven by EC-EARTH, and RACMO22E driven by NorESM1-M) display a distribution even with a negative skewness for the end of the century (mostly visible with logarithmic axis presented
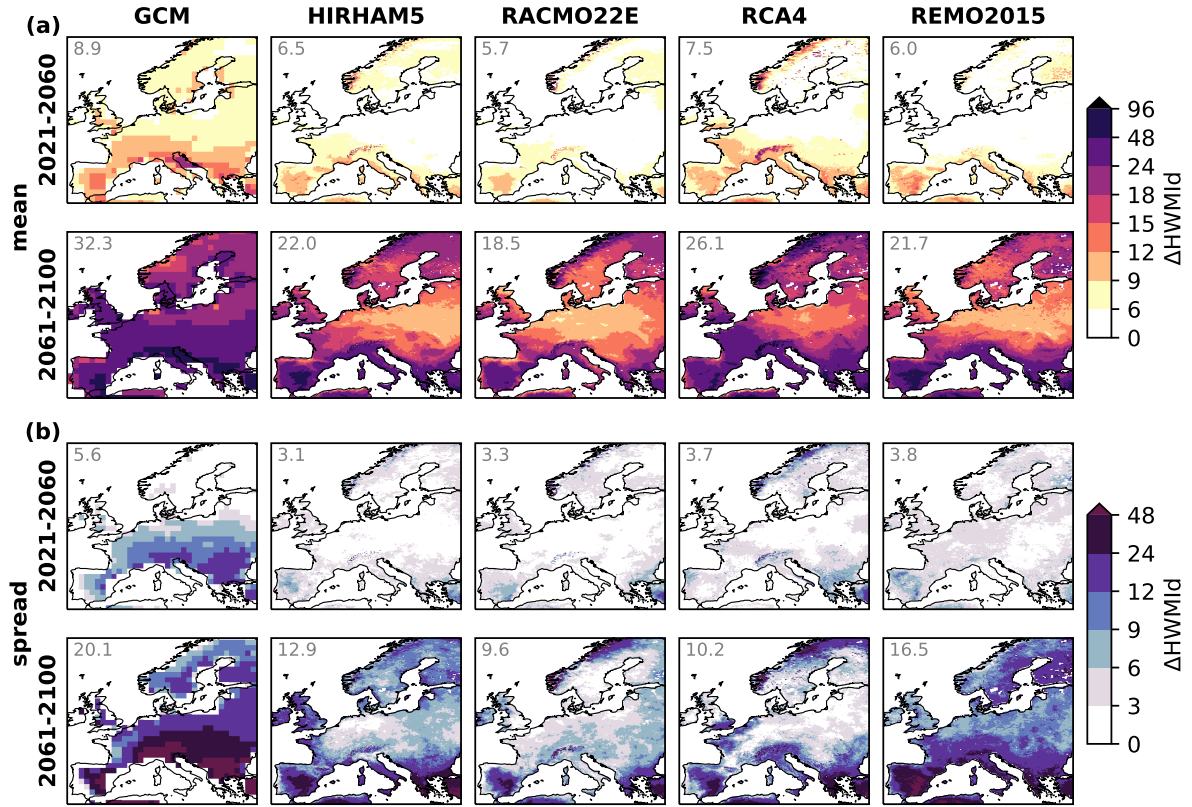
Figure 8. As Fig. 5 but for the climate change in HWMId.

in Fig. 10b), which indicates that high HWMId values become more common. In addition to increasing levels of HWMId, also the spread defined by the ranges increase in the future. With a logarithmic-scale x-axis, we observe interesting features in the change signal: i) The mean and the median increase approximately at an exponential rate; ii) the width of the logarithmically-transformed distribution does not change much from one period to another, i.e., a rise approximately at an exponential rate for both the low and high ends of HMWId values.

4 Discussion

4.1 Added value of RCMs compared to GCMs

With more small-scale processes being resolved, added value is expected from dynamical downscaling with an RCM compared to its driving GCM on the regional scale. Indeed, a large number of previous studies (e.g., Torma et al., 2015; Rummukainen, 2016; Strandberg and Lind, 2021) have reported such added value by RCMs in many respects. Thus, it is of great interest if RCMs create added value also for heat wave magnitudes. To realistically represent HWMId, a model must be able to capture
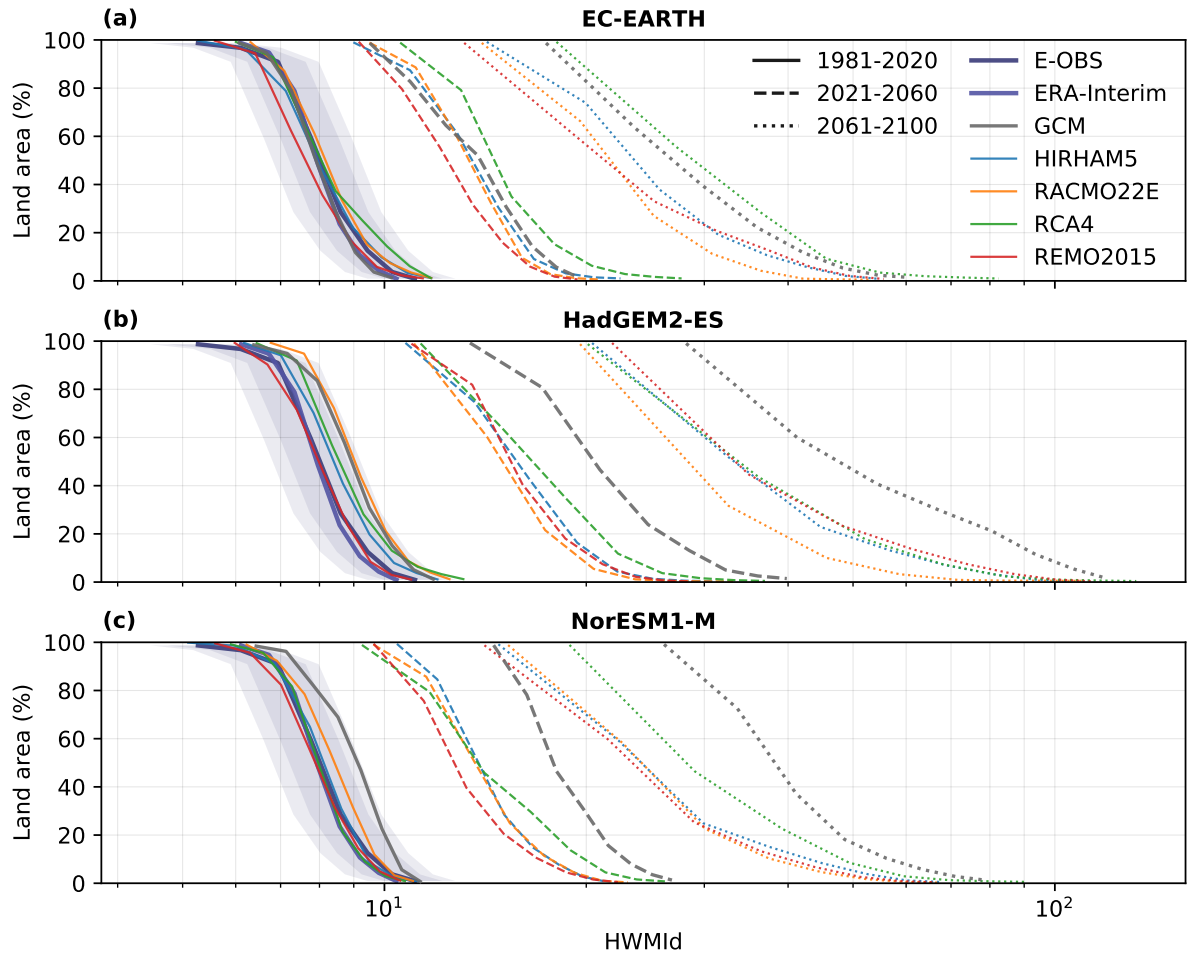
Figure 9. Percentage of the land area exceeding certain HWMId levels during the defined three periods (identified by the line styles), grouped by the driving GCMs: (**a**) EC-EARTH, (**b**) HadGEM2-ES, and (**c**) NorESM1-M. Colors represent different RCMs and gray the GCMs themselves. Data from E-OBS and ERA-Interim are also shown for comparison under the recent past climate. The shaded areas indicate within ±10 % (relatively dark) and ±15 % (light) of E-OBS data. Note that a logarithmic-scale x-axis is used.

not only the mean magnitude but also the intra-annual temporal evolution of daily T_{max}. Before discussing any potential added

260 value we assess to which degree the studied RCMs can represent the observed HWMId under the historical climate when driven by reanalysis data (i.e., perfect boundary conditions). We showed in Sect. 3.1 that the ERA-interim-driven runs generally reproduce the spatial and temporal patterns of the observed annual heat wave magnitudes over Europe. However, the RCMs clearly add their own signatures to the results leading to a larger variability in the spatial distribution of HWMId than reflected in E-OBS and ERA-Interim (Fig. 1). Even larger variability was found in the HWMId spatial distribution of Years 1994,

265 2003, 2006, and 2007 (Fig. S4–S7) than for the climatological mean, although these events, which were reported by the news headlines (Russo et al., 2015), are generally captured (Fig. 2a) by the RCMs when driven by ERA-interim. Another interesting
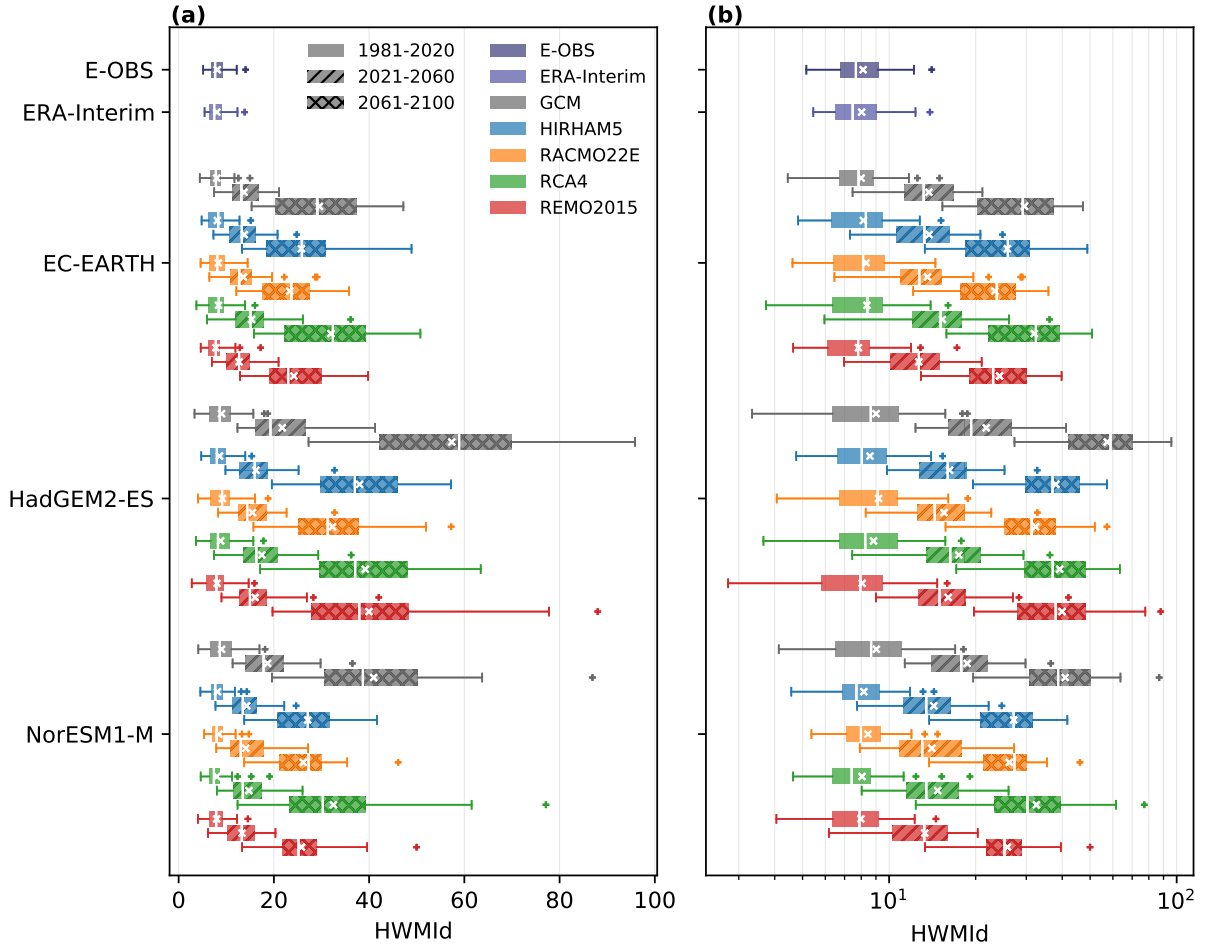
Figure 10. Box-plot for regional mean HWMId values under the defined three periods (indicated by filling patterns) in the simulation matrix of RCMs as well as the driving GCM simulations. Data are presented on different scales: (**a**) linear, and (**b**) logarithmic. Within each group (with one GCM), RCMs are identified by colors. Data from E-OBS and ERA-Interim are also shown for comparison under the recent past climate.

observation is that RCA4, which performs worse than the other RCMs (Table 2) in representing climatological mean, best reproduces the HWMId spatial distribution of Year 2003 (Fig. S5 and Table S4). A more in-depth event-based analysis would be needed to see which events are large-scale triggered from the boundaries of the RCMs and thereby picked up by the RCMs, and which events are triggered inside the RCM domain and therefore possibly missed or reproduced differently by the different RCMs. One example of a large-scale-driven event could be Year 2003, where the HWMId is picked up clearly by all the RCMs in its high intensity (Fig. 2a).

The three GCM simulations accounted for, roughly reproduce the observed patterns for the recent past climate, but miss out on details in structure and amplitude (Fig. 3). The RCMs downscaling these GCMs do well in adding more detailed

270

geographical patterns and, furthermore, pull the results closer to the observations. Again, the RCMs add their own pattern, e.g., the lower HWMId values over Eastern Europe in REMO2015. Signatures of the RCMs can also be demonstrated by the high similarity in HWMId spatial pattern between the simulations for each RCM despite a large difference between the driving GCMs (Table S1). Certainly, the RCMs also inherit signals from their driving GCMs. This is clearest for RACMO22E among the four RCMs, as the simulations with RCAMO22E show the highest r (except when driven by NorESM1-M) and the lowest RMSE against its driving GCMs (Table S2). Despite the inherited signals from their driving GCMs, the simulations of different RCMs when driven by a same GCM generally share higher similarity with each other than with the driving GCM (Table S2), implicating a coherence for RCMs in representing heat wave magnitudes when downscaling GCMs.

Some studies (e.g., Schiermeier, 2010; Kerr, 2011) show that uncertainty may increase when downscaling GCMs with RCMs as biases from the GCMs are conveyed to the RCMs and RCMs additionally add their own biases, referred to as the 'cascade of uncertainty' (e.g., Wilby and Dessai, 2010). However, many other studies (e.g., Torma et al., 2015; Di Luca et al., 2016; Rummukainen, 2016; Sørland et al., 2018; Strandberg and Lind, 2021) indicate that RCMs can also add value upon the driving GCM simulations. This study demonstrates the added value for heat wave magnitudes that have so far not been studied as extensively as for other aspects of climate and climate change, reflected in adding more detailed geographical patterns and pulling the results closer to the observations (Fig. 3; Table 4 and 5). Such added value confirms the usefulness of RCMs for downscaling coarse-scale GCM simulations, because downscaling can be considered as an act of adding useful information to that already provided by GCMs: information about how local geography influences the local climate (as in this case) and information about how local climates depend on the ambient large-scale conditions and teleconnections that the GCMs skillfully reproduce. Moreover, RCMs may also more realistically represent some atmospheric processes relative to the GCMs. Our analysis of the ensemble spread along the GCM dimension, reflecting uncertainty associated with driving data, reveals that the RCMs alter the spatial HWMId pattern from their driving GCM simulations, and that the alteration is different between the RCMs (Fig. 5 and Table S3). This, on the other hand, suggests that the uncertainties of GCMs in simulating heat wave magnitudes would be transformed by RCMs in a complex manner, rather than simply inherited, due to the nonlinear nature of model dynamics and physics, thus rejecting the concept of 'cascade of uncertainty'.

To reveal the specific factors/processes behind the added value of RCMs in simulating heat wave magnitudes, however, further analysis is required. Table 4 and Fig. 3 clearly show that RCMs capture the HWMId better than the GCMs for the recent past climate. An interesting finding of this study is that some spatial features of HWMId showing orographic traces, thereby considered to be related to orographic effects, can be seen as one aspect of the added value of RCMs as GCMs cannot represent such features due to too coarse spatial resolution. Nevertheless, the orographic effects are represented differently across the RCMs (i.e., the large ensemble spread along the RCM dimension; Fig. 4), suggesting that the representation of orographic effects, possibly related to dynamical/thermodynamical interaction with parameterizations, is one of the major sources of uncertainty. This supports the statement by Sørland et al. (2018) that increasing the spatial resolution should not be the single factor contributing to the added value of RCMs.

4.2 Possible processes behind the climate change signals

The RCMs, as well as the driving GCMs, project a rise in HWMId values at an exponential rate under RCP8.5 on the European continent. The exponential increase is patent as we see the linear shape of time series when plotting on a logarithmic scale (Fig. S8). As a result, heat waves more severe than the most severe one that has been recorded until now are projected to occur almost every year at the end of the century if we follow the high-end emission pathway (RCP8.5). According to the definition of HMWId, the approximately exponential rise can be expected because the projected warming will on one hand increase the daily magnitude (Eq. 2) and on the other hand extend the duration simultaneously. Apart from the agreement on the future severity of heat wave magnitudes under this scenario, the RCMs modify the future climate change signals projected by the driving GCMs, tending to moderate the rise in HWMId values and also deliver some different features in the spatial pattern. The underlying drivers of heat waves may be related to land-atmosphere interactions as well as atmospheric processes (Horton et al., 2016; Liu et al., 2020). For example, plant physiological CO_2 response may have a positive effect on $T_{\max}$ (Schwingshackl et al., 2019) and thereby on HWMId values. Understanding how RCMs differ from GCMs in representing processes that modify the climate change signals can have implications for how to utilize model projections in studies on climate change mitigation and adaptation.

Here, rather than directly looking into the associated processes, we investigate the corresponding climate change signals in annual mean $T_{\max}$ (Fig. S9), dry days (Fig. S10), and $P - E$ (Fig. S11). As the HWMId is directly calculated from daily $T_{\max}$, the increase in annual $T_{\max}$ provides a warming background that the surge in HWMId is associated with. An increase in the annual number of dry days may indicate a higher tendency of longer warm spells and hence a rise in HWMId. Compared to the number of dry days, $P - E$ (or effective precipitation if we do not consider the variation in runoff) is more strongly related to dry soil conditions; the correlation between effective precipitation and HWMId may reveal the effect of land-atmosphere feedbacks, as the deficit in soil moisture may reduce latent heat flux allowing temperatures to rise further (Zhang et al., 2020).

The RCMs dampen the increase, and to some extent modify the spatial pattern, in the annual mean of daily $T_{\max}$ in HadGEM2-ES and NorESM1-M (Fig. S9 and Table S5). This could potentially explain the more moderate rise in HWMId simulated by the RCMs compared to their driving GCMs. EC-EARTH and its downscaling by the RCMs show a similar level of increase as well as a similar spatial pattern in the annual mean of daily $T_{\max}$ (Fig. S9 and Table S5), and we recognize that they also project a similar rise in HWMId (Fig. 6). According to Schwingshackl et al. (2019), these results may be linked to the positive effect of plant physiological CO_2 response on $T_{\max}$, since the GCMs except for EC-EARTH consider this response but all the RCMs do not. That RCMs are better at representing near-surface processes that are important for reducing the drying feedback is also considered a reason (Sørland et al., 2018). For the annual number of dry days (Fig. S10), on regional average, the RCMs also reduce the increase from their driving GCMs, but with a few exceptions (e.g., HIRHAM5 downloading EC-EARTH and HadGEM2-ES). For annual $P - E$ (Fig. S11), on a continental basis, the RCMs, except for HIRHAM5, show a smaller decrease (or larger increase) than the driving GCMs. In general, these results are consistent with Coppola et al. (2021) who discussed similar indices.

We further examined the correlation relating to spatial characteristics between the change in HWMId and that in each of the three indices accounted for, in order to find any trace whether, and if so how, the drying processes (of either atmospheric or soil) regulate the spatial pattern of ΔHWMId. For the detailed spatial pattern of ΔHWMId, it cannot be explained by the change in the annual mean of daily T_{max} alone, even though HWMId is calculated based only on daily T_{max}. This is especially true for GCM simulations as they have a poor spatial correlation between ΔHWMId and ΔT_{max} (Fig. S9). All the GCM and RCM simulations agree on the high ΔHWMId in southern Europe (Fig. 3), which may be amplified by the projected drying trend (Fig. S10 and S11), while in northern Europe where rapid warming is projected (Fig. S9), high ΔHWMId is seen in the RCM simulations only (Fig. 3). The r values in Fig. S9–S11 read that the general warming, compared to drying, plays a small role in regulating the spatial pattern of ΔHWMId in GCM simulations, different from the case of the RCMs.

As a preliminary effort, concerning only the spatial pattern, the above analysis is however far from building clear causal links. Moreover, we are not yet clear about what is leading to the weaker drying trend in the RCM simulations. Atmospheric blocking, with adiabatic warming of sinking air and anomalous clear-sky radiative forcing, is an important driver of heat waves (e.g., Bieli et al., 2015; Schaller et al., 2018). According to Masato et al. (2013), the three CMIP5 GCMs assessed show a decrease in summertime North Atlantic blockings and an increase in blockings over eastern Europe or Russia indicating an eastward shift of the blocking activity. This implies that the underlying processes of ΔHWMId is possibly beyond the atmospheric dynamics. It is of interest whether, and if so how, the RCMs modify the climate change signals of atmospheric blocking from their driving GCMs, and whether this modification is related to the differences in ΔHWMId patterns between GCMs and RCMs as presented here, which is worth further study, though representing atmospheric blocking is considered a challenge (e.g., Masato et al., 2013; Davini and D'Andrea, 2016; Jury et al., 2019). Regarding the representation of the effect of land-atmosphere interactions, investigation detailed into different sets of parameterizations in GCMs and RCMs as well as some additional sensitivity experiments may be necessary for a better understanding.

4.3 Other matters

A full simulation matrix without gaps facilitates a fair comparison after aggregating along either the GCM or the RCM dimension. This requisite limits the size of RCM simulation matrix available for analysis. The limited size of the GCM-RCM simulation matrix could be a shortcoming influencing the robustness of the results in the study, especially for the uncertainty analysis where the uncertainty is described by the spread (maximum − minimum) across three or four ensemble members. In fact, we conducted the same analyses upon another GCM-RCM simulation matrix (with the same size but different GCM/RCM members) and derived similar results (not shown), which can alternatively support the conclusions herein.

The HWMId can also be applied to other temperature variables, but with different processes/impacts involved. For example, the HWMId applied to daily minimum temperature serve a measurement of heat wave magnitude taking into account also the nighttime cooling effect (Russo et al., 2015). As another example, Apparent Heat Wave Index (AHWI, Russo et al., 2017), the HWMId applied to daily apparent temperature, considers also the impact of air humidity on human beings. Such variants of HWMId are being considered for future studies.

5 Summary and conclusions

375 By deploying the HWMId index, the study addresses how four different RCMs downscaling i) reanalysis and ii) three different GCMs, represent European heat wave magnitudes.

Initially, the performance of the RCMs in reproducing historical heat wave magnitudes is evaluated by comparing the ERA-Interim-driven evaluation runs of the RCMs with E-OBS. It shows that the RCMs generally capture most spatial and temporal features of the observed HWMId, suggesting the RCMs as a reliable tool for simulating heat wave magnitudes. Our results

380 prove the added value of RCMs over GCMs in representing the observed heat wave magnitudes. Compared to the driving GCMs the RCMs generally have lower RMSE and higher r against the observational data for the climatological mean of HWMId values under the recent past climate. In general, the RCMs improve the spatial pattern of HWMId across the European continent compared to the driving GCMs. In addition, the RCMs reveal some small-scale features (e.g., relating to orographic effects) that the GCMs fail likely due to their coarse resolutions. The closest agreement with observations is seen for the RCM

385 ensemble mean.

A rise in HWMId at an exponential rate is projected consistently by all the GCM and RCM simulations accounted for in the study, probably because the warming boosts both the intensity and duration. However, the RCMs modify some feature of the climate change signals in the driving GCM simulations. A somewhat more moderate rise across the European continent is projected by the RCMs, as a corresponding result of the reduced warming. The RCMs also differ from the driving GCMs in

390 the spatial pattern of the climate change signals of HWMId.

We also analyzed the uncertainties of the GCM-RCM simulation matrix in simulating heat wave magnitudes. The results show that the uncertainty associated with choice of RCM is of similar importance as with driving data. The ensemble spread-/mean ratio is approximately one fifth for the present-climate HWMId and over half for the climate change signals. A major source of the uncertainty associated with the RCMs appears to be associated with the representation of orographic effects.

395 The RCMs reduce the large ensemble spread across the GCM simulations though, especially for the climate change signals in HWMId. Moreover, no consistent spatial pattern is observed in the ensemble spreads along the GCM dimension for different RCMs. Consequently, the results indicate that the uncertainties of GCMs in simulating heat wave magnitudes would not be simply inherited by RCMs but are transformed in a complex manner due to the nonlinear nature of model dynamics and physics.

400 *Code availability.* All the analyses were done using Python packages.

Data availability. The EURO-CORDEX RCM data and the CMIP5 GCM data analyzed in this work are available for download via the Earth System Grid Federation (ESGF) under the project name "CORDEX" and "CMIP5", respectively, at the NSC-LIU-SMHI (Swedish) datanode: https://esg-dn1.nsc.liu.se/projects/esgf-liu/ (last access: June 29, 2022). The E-OBS data are available for download via the ECA&D project: https://www.ecad.eu/ (last access: June 29, 2022). The ERA-Interim reanalysis data from ECMWF can be accessed using their Meteorological

405 Archival and Retrieval System (MARS): https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim (last access: June 29, 2022).

Author contributions. CL set up the analysis framework with the scientific contributions of EK and RW. CL produced the figures and tables. CL wrote the publication with important contributions from EK, RW and DC.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

410 *Acknowledgements.* This work is partly supported by Swedish FORMAS (2018-02858 and 2019-01520) and VR (2019-03954). The authors would like to thank the EURO-CORDEX network (https://www.euro-cordex.net/) and WCRP CORDEX (https://cordex.org/) for ensuring availability of CORDEX data. We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (https://www.uerra.eu/) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu/).

References

415   Åström, C., Bjelkmar, P., and Forsberg, B.: Attributing summer mortality to heat during 2018 heatwave in Sweden, Environmental Epidemi-
ology, 3, 16–17, https://doi.org/10.1097/01.EE9.0000605788.56297.b5, 2019.

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, 332, 220–224, https://doi.org/10.1126/science.1201224, 2011.

Benestad, R. E., Chen, D., and Hanssen-Bauer, I.: Empirical-statistical downscaling, World Scientific Publishing Company, 2008.

420   Benestad, R. E., van Oort, B., Justino, F., Stordal, F., Parding, K. M., Mezghani, A., Erlandsen, H. B., Sillmann, J., and Pereira-Flores, M. E.:
Downscaling probability of long heatwaves based on seasonal mean daily maximum temperatures, Advances in Statistical Climatology,
Meteorology and Oceanography, 4, 37–52, https://doi.org/10.5194/ascmo-4-37-2018, 2018.

Bieli, M., Pfahl, S., and Wernli, H.: A Lagrangian investigation of hot and cold temperature extremes in Europe, Quarterly Journal of the
Royal Meteorological Society, 141, 98–108, https://doi.org/10.1002/qj.2339, 2015.

425   Ceccherini, G., Russo, S., Ameztoy, I., Marchese, A. F., and Carmona-Moreno, C.: Heat waves in Africa 1981–2015, observations and
reanalysis, Natural Hazards and Earth System Sciences, 17, 115–125, https://doi.org/10.5194/nhess-17-115-2017, 2017.

Coppola, E., Nogherotto, R., Ciarlò, J. M., Giorgi, F., van Meijgaard, E., Kadygrov, N., Iles, C., Corre, L., Sandstad, M., Somot, S., Nabat, P.,
Vautard, R., Levavasseur, G., Schwingshackl, C., Sillmann, J., Kjellström, E., Nikulin, G., Aalbers, E., Lenderink, G., Christensen, O. B.,
Boberg, F., Sørland, S. L., Demory, M.-E., Bülow, K., Teichmann, C., Warrach-Sagi, K., and Wulfmeyer, V.: Assessment of the European
430   Climate Projections as Simulated by the Large EURO-CORDEX Regional and Global Climate Model Ensemble, Journal of Geophysical
Research: Atmospheres, 126, e2019JD032 356, https://doi.org/10.1029/2019JD032356, 2021.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and
Precipitation Data Sets, Journal of Geophysical Research: Atmospheres, 123, 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Davini, P. and D'Andrea, F.: Northern Hemisphere Atmospheric Blocking Representation in Global Climate Models: Twenty Years of
435   Improvements?, Journal of Climate, 29, 8823–8840, https://doi.org/10.1175/JCLI-D-16-0242.1, 2016.

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., Haim-
berger, 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., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis:
440   Configuration and performance of the data assimilation system, Quarterly Journal of the royal meteorological society, 137, 553–597,
https://doi.org/10.1002/qj.828, 2011.

Di Luca, A., Argüeso, D., Evans, J. P., de Elía, R., and Laprise, R.: Quantifying the overall added value of dynamical down-
scaling and the contribution from different spatial scales, Journal of Geophysical Research: Atmospheres, 121, 1575–1590,
https://doi.org/10.1002/2015JD024009, 2016.

445   Dosio, A., Mentaschi, L., Fischer, E. M., and Wyser, K.: Extreme heat waves under 1.5 °C and 2 °C global warming, Environmental Research
Letters, 13, 054 006, https://doi.org/10.1088/1748-9326/aab827, 2018.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model
Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958,
https://doi.org/10.5194/gmd-9-1937-2016, 2016.

450 Giorgi, F., Jones, C., and Asrar, G. R.: Addressing climate information needs at the regional level: the CORDEX framework, World Meteorological Organization (WMO) Bulletin, 58, 175, 2009.

Guo, Y., Gasparrini, A., Armstrong, B. G., Tawatsupa, B., Tobias, A., Lavigne, E., de Sousa Zanotti Stagliorio Coelho, M., Pan, X., Kim, H., Hashizume, M., Honda, Y., Guo, Y.-L. L., Wu, C.-F., Zanobetti, A., Schwartz, J. D., Bell, M. L., Scortichini, M., Michelozzi, P., Punnasiri, K., Li, S., Tian, L., Garcia, S. D. O., Seposo, X., Overcenco, A., Zeka, A., Goodman, P., Dang, T. N., Dung, D. V., Mayvaneh, F., Saldiva,

455 P. H. N., Williams, G., and Tong, S.: Heat wave and mortality: a multicountry, multicommunity study, Environmental Health Perspectives, 125, 087 006, https://doi.org/10.1289/EHP1026, 2017.

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J. M., Wibig, J., Casanueva, A., and Soares, P. M.: Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE, International Journal of Climatology, 39, 3846–3867, https://doi.org/10.1002/joc.5469, 2019.

460 Horton, R. M., Mankin, J. S., Lesk, C., Coffel, E., and Raymond, C.: A review of recent advances in research on extreme heat events, Current Climate Change Reports, 2, 242–259, https://doi.org/10.1007/s40641-016-0042-x, 2016.

IPCC: Part A: Global and Sectoral Aspects, in: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Field, C., Van Aalst, M., Aalst, M., Adger, W., Arent, D., Barnett, J., Betts, R., Bilir, E., Birkmann, J., Carmin, J., Chadee, D., Challinor, A., Chatterjee, M., Cramer, W., Davidson,

465 D., Estrada, Y., Gattuso, J.-P., Hijioka, Y., Guldberg, O., Huang, H.-Q., Insarov, G., Jones, R., Kovats, S., Lankao, P., Larsen, J., nigo Losada, I., Marengo, J., McLean, R., Mearns, L., Mechler, R., Morton, J., Niang, I., Oki, T., Olwoch, J., Opondo, M., Poloczanska, E., Pörtner, H.-O., Redsteer, M., Reisinger, A., Revi, A., Schmidt, D., Shaw, R., Solecki, W., Stone, D., Stone, J., Strzepek, K., Suarez, A., Tschakert, P., Valentini, R., Vicuna, S., Villamizar, A., Vincent, K., Warren, R., White, L., Wilbanks, T., Wong, P., and Yoh, G., p. 1132, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

470 IPCC: Summary for Policymakers, in: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, edited by Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J., Chen, Y., Zhou, X., Gomis, M., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T., p. 32, World Meteorological Organization, Geneva, Switzerland,

475 2018.

IPCC: Summary for Policymakers, in: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Pörtner, H.-O., Roberts, D., Poloczanska, E., Mintenbeck, K., Tignor, M., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., and Okem, A., p. 37, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, in Press, 2022.

480 Jacob, D., Bärring, L., Christensen, O. B., Christensen, J. H., de Castro, M., Deque, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellström, E., Lenderink, G., Rockel, B., Sánchez, E., Schär, C., Seneviratne, S. I., Somot, S., van Ulden, A., and van den Hurk, B.: An inter-comparison of regional climate models for Europe: model performance in present-day climate, Climatic Change, 81, 31–52, https://doi.org/10.1007/s10584-006-9213-4, 2007.

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., Benestad, R., Boberg, F., Buonomo, E., Cardoso, R. M.,

485 Casanueva, A., Christensen, O. B., Christensen, J. H., Coppola, E., Cruz, L. D., Davin, E. L., Dobler, A., Domínguez, M., Fealy, R., Fernandez, J., Gaertner, M. A., García-Díez, M., Giorgi, F., Gobiet, A., Goergen, K., Gómez-Navarro, J. J., Alemán, J. J. G., Gutiérrez, C., Gutiérrez, J. M., Güttler, I., Haensler, A., Halenka, T., Jerez, S., Jiménez-Guerrero, P., Jones, R. G., Keuler, K., Kjellström, E., Knist, S.,

Kotlarski, S., Maraun, D., van Meijgaard, E., Mercogliano, P., Montávez, J. P., Navarra, A., Nikulin, G., de Noblet-Ducoudré, N., Panitz, H.-J., Pfeifer, S., Piazza, M., Pichelli, E., Pietikäinen, J.-P., Prein, A. F., Preuschmann, S., Rechid, D., Rockel, B., Romera, R., Sánchez, E., Sieck, K., Soares, P. M. M., Somot, S., Srnec, L., Sørland, S. L., Termonia, P., Truhetz, H., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community, Regional Environmental Change, 20, 1–20, https://doi.org/10.1007/s10113-020-01606-9, 2020.

Jerez, S., Palacios-Peña, L., Gutiérrez, C., Jiménez-Guerrero, P., López-Romero, J. M., Pravia-Sarabia, E., and Montávez, J. P.: Sensitivity of surface solar radiation to aerosol–radiation and aerosol–cloud interactions over Europe in WRFv3. 6.1 climatic runs with fully interactive aerosols, Geoscientific Model Development, 14, 1533–1551, https://doi.org/10.5194/gmd-14-1533-2021, 2021.

Jury, M. W., Herrera, S., Gutiérrez, J., and Barriopedro, D.: Blocking representation in the ERA-Interim driven EURO-CORDEX RCMs, Climate Dynamics, 52, 3291–3306, https://doi.org/10.1007/s00382-018-4335-8, 2019.

Kerr, R. A.: Vital Details of Global Warming Are Eluding Forecasters, Science, 334, 173–174, https://doi.org/10.1126/science.334.6053.173, 2011.

Liu, X., He, B., Guo, L., Huang, L., and Chen, D.: Similarities and differences in the mechanisms causing the European summer heatwaves in 2003, 2010, and 2018, Earth's Future, 8, e2019EF001 386, https://doi.org/10.1029/2019EF001386, 2020.

Luo, Z., Yang, J., Gao, M., and Chen, D.: Extreme hot days over three global mega-regions: Historical fidelity and future projection, Atmospheric Science Letters, 21, e1003, https://doi.org/10.1002/asl.1003, 2020.

Masato, G., Hoskins, B. J., and Woollings, T.: Winter and Summer Northern Hemisphere Blocking in CMIP5 Models, Journal of Climate, 26, 7044 – 7059, https://doi.org/10.1175/JCLI-D-12-00466.1, 2013.

Molina, M., Sánchez, E., and Gutiérrez, C.: Future heat waves over the Mediterranean from an Euro-CORDEX regional climate model ensemble, Scientific Reports, 10, 1–10, https://doi.org/10.1038/s41598-020-65663-0, 2020.

Perkins, S. E. and Alexander, L. V.: On the measurement of heat waves, Journal of Climate, 26, 4500–4517, https://doi.org/10.1175/JCLI-D-12-00383.1, 2013.

Rennie, S., Goergen, K., Wohner, C., Apweiler, S., Peterseil, J., and Watkins, J.: A climate service for ecologists: sharing pre-processed EURO-CORDEX regional climate scenario data using the eLTER Information System, Earth System Science Data, 13, 631–644, https://doi.org/10.5194/essd-13-631-2021, 2021.

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, 331, 171–178, https://doi.org/10.1016/j.crvi.2007.12.001, 2008.

Rummukainen, M.: Added value in regional climate modeling, Wiley Interdisciplinary Reviews: Climate Change, 7, 145–159, https://doi.org/10.1002/wcc.378, 2016.

Russo, S., Sillmann, J., and Fischer, E. M.: Top ten European heatwaves since 1950 and their occurrence in the coming decades, Environmental Research Letters, 10, 124 003, https://doi.org/10.1088/1748-9326, 2015.

Russo, S., Marchese, A. F., Sillmann, J., and Immé, G.: When will unusual heat waves become normal in a warming Africa?, Environmental Research Letters, 11, 054 016, https://doi.org/10.1088/1748-9326/11/5/054016, 2016.

Russo, S., Sillmann, J., and Sterl, A.: Humid heat waves at different warming levels, Scientific Reports, 7, 1–7, https://doi.org/10.1038/s41598-017-07536-7, 2017.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., and Russo, S.: Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles, Environmental Research Letters, 13, 054 015, https://doi.org/10.1088/1748-9326/aaba55, 2018.

Schiermeier, Q.: The real holes in climate science: like any other field, research on climate change has some fundamental gaps, although not the ones typically claimed by sceptics. Quirin Schiermeier takes a hard look at some of the biggest problem areas, Nature, 463, 284–288, https://doi.org/10.1038/463284a, 2010.

Schwingshackl, C., Davin, E. L., Hirschi, M., Sørland, S. L., Wartenburger, R., and Seneviratne, S. I.: Regional climate model projections underestimate future warming due to missing plant physiological CO_2 response, Environmental Research Letters, 14, 114 019, https://doi.org/10.1088/1748-9326/ab4949, 2019.

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and climate extreme events in a changing climate, in: Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekç, O., Yu, R., and Zhou, B., p. 345, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.

Soares, P. M. M., Maraun, D., Brands, S., Jury, M. W., Gutiérrez, J. M., San-Martín, D., Hertig, E., Huth, R., Belušić Vozila, A., Cardoso, R. M., Kotlarski, S., Drobinski, P., and Obermann-Hellhund, A.: Process-based evaluation of the VALUE perfect predictor experiment of statistical downscaling methods, International Journal of Climatology, 39, 3868–3893, 2019.

Sørland, S. L., Schär, C., Lüthi, D., and Kjellström, E.: Bias patterns and climate change signals in GCM-RCM model chains, Environmental Research Letters, 13, 074 017, https://doi.org/10.1088/1748-9326/aacc77, 2018.

Sørland, S. L., Fischer, A. M., Kotlarski, S., Künsch, H. R., Liniger, M. A., Rajczak, J., Schär, C., Spirig, C., Strassmann, K., and Knutti, R.: CH2018–National climate scenarios for Switzerland: How to construct consistent multi-model projections from ensembles of opportunity, Climate Services, 20, 100 196, https://doi.org/10.1016/j.cliser.2020.100196, 2020.

Strandberg, G. and Lind, P.: The importance of horizontal model resolution on simulated precipitation in Europe–from global to regional models, Weather and Climate Dynamics, 2, 181–204, https://doi.org/10.5194/wcd-2-181-2021, 2021.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Torma, C., Giorgi, F., and Coppola, E.: Added value of regional climate modeling over areas characterized by complex terrain—Precipitation over the Alps, Journal of Geophysical Research: Atmospheres, 120, 3957–3972, https://doi.org/10.1002/2014JD022781, 2015.

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., Coppola, E., Corre, L., van Meijgaard, E., Nogherotto, R., Sandstad, M., Schwingshackl, C., Somot, S., Aalbers, E., Christensen, O. B., Ciarlò, J. M., Demory, M.-E., Giorgi, F., Jacob, D., Jones, R. G., Keuler, K., Kjellström, E., Lenderink, G., Levavasseur, G., Nikulin, G., Sillmann, J., Solidoro, C., Sørland, S. L., Steger, C., Teichmann, C., Warrach-Sagi, K., and Wulfmeyer, V.: Evaluation of the large EURO-CORDEX regional climate model ensemble, Journal of Geophysical Research: Atmospheres, n/a, e2019JD032 344, https://doi.org/10.1029/2019JD032344, 2020.

Wilby, R. L. and Dessai, S.: Robust adaptation to climate change, Weather, 65, 180–185, https://doi.org/10.1002/wea.543, 2010.

Wilcke, R. A. I., Kjellström, E., Lin, C., Matei, D., Moberg, A., and Tyrlis, E.: The extremely warm summer of 2018 in Sweden–set in a historical context, Earth System Dynamics, 11, 1107–1121, https://doi.org/10.5194/esd-11-1107-2020, 2020.

Zampieri, M., Russo, S., di Sabatino, S., Michetti, M., Scoccimarro, E., and Gualdi, S.: Global assessment of heat wave magnitudes from 1901 to 2010 and implications for the river discharge of the Alps, Science of the Total Environment, 571, 1330–1339, https://doi.org/10.1016/j.scitotenv.2016.07.008, 2016.

Zhang, P., Jeong, J.-H., Yoon, J.-H., Kim, H., Wang, S.-Y. S., Linderholm, H. W., Fang, K., Wu, X., and Chen, D.: Abrupt shift to hotter and drier climate over inner East Asia beyond the tipping point, Science, 370, 1095–1099, https://doi.org/10.1126/science.abb3368, 2020.