Are dependencies of extreme rainfall on humidity more reliable in convection-permitting climate models?

Geert Lenderink[1], Nikolina Ban[2], Erwan Brisson[3], Ségolène Berthou[4], Virginia Edith Cortés-Hernández[3], Elizabeth Kendon[4,5], Hayley J. Fowler[6], Hylke de Vries[1]

5 [1]R&D Weather and Climate, Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[2]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Austria
[3]Centre National de Recherches Météorologiques (CNRM), Université de Toulouse, Météo-France, CNRS, Toulouse, France
[4]Met Office Hadley Centre, Exeter, UK
[5]Bristol University, Bristol, UK
10 [6]School of Engineering, Newcastle University, Newcastle upon Tyne, UK

Correspondence to: Geert Lenderink (geert.lenderink@knmi.nl)

Abstract. Convection-permitting climate models (CPMs) are becoming increasingly used in climate change studies. These models show greatly improved convective rainfall statistics compared to parameterized-convection regional climate models (RCMs), but are they also more reliable in a climate change setting? Increases of rainfall extremes are generally considered

15 to be caused by increases in *absolute* humidity, primarily following from the Clausius-Clapeyron relation, while the influence of *relative* humidity changes is uncertain and not systematically explored. Quantifying these humidity dependencies in the present-day climate may help the interpretation of future changes, which are driven by increases in absolute humidity, but also decreases in relative humidity in most continental areas in summer. Here, we systematically analyse hourly rainfall extremes and their dependencies on 2m dew point temperature (absolute humidity) and dew point depression (relative humidity) in 7

20 RCM and 5 CPM simulations for the present-day climate. We compare these to observations from The Netherlands (a moderate moist climate) and Southern France (a warmer and drier climate). We find that the RCMs display a large spread in outcomes, in particular in their relative humidity dependence, with a strong suppression of hourly rainfall extremes in low relative humidity conditions. CPMs produce better overall rainfall statistics, show less inter-model spread, and have absolute and relative humidity dependencies more consistent with the observations. In summary, our results provide evidence that future

25 changes in convective rainfall extremes in CPMs are more reliable compared to RCMs, whereas the discussed dependencies also provide a metric to evaluate and further improve model performance as well as improving convection schemes.

1 Introduction

In recent years, climate modelling has rapidly progressed towards higher resolution with so-called *convection-permitting* climate models that explicitly resolve the largest convective motions in the atmosphere (Stevens et al., 2019; Prein et al., 2020;

30 Schär et al., 2020; Ban et al., 2021; Kendon et al., 2021b; Pichelli et al., 2021; Lucas-Picher et al., 2021). These models typically run at <4-5 km grid spacing, with non-hydrostatic dynamics, which allows them to explicitly resolve deep convection

without the need for a convection parameterization. As computational demands are high, CPMs can only be applied for relatively short time periods; on regional domains typically 10-30 years are simulated. In recent years, large efforts have been made to produce a number of these CPM runs, for instance the international coordinated action CORDEX flagship pilot study

35   over the European Alpine region (Coppola et al., 2020; Ban et al., 2021; Pichelli et al., 2021) and a 12-member ensemble for 100 year projections from 1980 to 2080 produced as part of the UK national climate scenarios (Kendon et al., 2020, 2021a, 2023).

The primary reason to run CPMs at km-scales is their improved representation of convective processes, which give rise to

40   locally-intense convective rainfall on sub-daily time scales that e.g. can lead to flash floods and debris flows. In recent years, knowledge on these type of extremes, and how they could be affected by climate change, has advanced considerably based on observational evidence, process understanding, and high-resolution modelling such as with CPMs (Fowler et al., 2021a, b). Convection parameterization schemes have been identified as responsible for substantial biases in convective rainfall statistics in coarser-resolution climate models, causing rainfall intensities that are too low and too frequent, and a too early triggering

45   during the day (Kendon et al., 2012; Ban et al., 2021). CPMs improve the frequency and diurnal cycle of rainfall, as well as the distribution of the extremes and rainfall spatial patterns (Kendon et al., 2012; Ban et al., 2020, 2021; Berthou et al., 2020; Fumière et al., 2020). CPMs have also been evaluated for process-based metrics such as intensity-duration characteristics and cell-size distribution (Kendon et al., 2012), assessing the reliability of the underlying rainfall processes. Finally, rain cell tracking has been used to investigate the lifecycle of convective rainfall events, providing additional information on how

50   organized convective systems behave in CPMs versus radar data (Purr et al., 2019, 2021; Müller et al., 2023).

Limited evidence exists that the improved representation of convective cloud processes in CPMs, as compared to RCMs, leads to more certain estimates of future extreme rainfall changes (Fosser et al., 2020, 2024). Indeed, when future changes are dominated by simple thermodynamic factors (the moisture increase) then CPMs often project similar changes to RCMs (Ban

55   et al., 2020). Despite their improved behaviour, CPMs are not without faults: at 2-3 km resolution important parts of the physics and dynamics of convective clouds are still not (well) resolved, e.g. convective plumes, entrainment processes at the cloud interface, detailed boundary layer cold pool dynamics and cloud microphysical processes (Prein et al., 2021; Lochbihler et al., 2021). Therefore, it is important to establish the trustworthiness of these models using measures that relate to climate change. Here, we study the dependency of extreme rainfall on measures of absolute and relative humidity – key drivers of

60   future changes, as explained below.

The intensification of rainfall extremes due to global warming is to first order explained by increases in absolute humidity following from the Clausius-Clapeyron (CC) relation (Pall et al., 2006; Trenberth, 2011; Fischer and Knutti, 2016). The Clausius-Clapeyron relation governs the increase in the water vapor *holding* capacity of the air as a function of temperature,

65   and near to the surface gives a rate of increase of 6-7% per degree warming. Evidence from observations suggests that the rate

of change of precipitation extremes with absolute humidity could be much stronger. In observations, so-called scaling rates of precipitation extremes on near surface point temperatures up to 12-14% per degree (2xCC) have been found (Lenderink and van Meijgaard, 2008, 2010; Lenderink et al., 2011; Fowler et al., 2021a; Ali et al., 2022) as well as strong rainfall intensification beyond a critical value of the integrated water vapor path (Neelin et al., 2022). Scaling rates beyond the CC relationship are

70   commonly denoted as super-CC scaling. But scaling rates also differ for different areas of the globe, and depend on the temperature measure used (Lenderink et al., 2018; Bui et al., 2019). It is not always straightforward to connect the temperature measure used to the humidity of the air in which a rain shower develops. In this paper, we use dew point temperature which directly measures absolute humidity.

75   Several explanations have been proposed to explain observed super-CC scaling. These range from systematic statistical shift in the rainfall type with temperature (Haerter and Berg, 2009; Fowler et al., 2021a; Molnar et al., 2015), to positive feedback mechanisms from physical processes. Latent heat release in the cloud could invigorate updraft motions, leading to more condensation of moist updraft air, and stronger precipitation rates (Loriaux et al., 2013). Stronger sub-cloud cold pool dynamics with warming could lead to stronger cloud organisation and larger moisture availability at the cloud condensation level (Haerter

80   and Schlemmer, 2018; Lochbihler et al., 2021). Deeper warm cloud levels below the freezing level could lead to more efficient warm rain processes (Prein and Heymsfield, 2020). Finally, large-scale dynamical adjustments to latent heating could also lead to enhanced moisture convergence and destabilization of the atmosphere (Lenderink et al., 2017; Nie et al., 2018).

We do not aim here to further study the causes of super-CC behaviour. Instead, we conjecture that reproducing the observed

85   near-surface dew point temperature dependence in a model is an essential prerequisite for confidence in its future projections. Nonetheless, we acknowledge that the relation between present-day derived scaling rates and future projection of changes in rainfall extremes is rather complex (Lenderink and Attema, 2015; Bao et al., 2017; Zhang et al., 2017; Lenderink et al., 2019, 2021; Fowler et al., 2021a). One major complication is that co-varying atmospheric variables with (dew point) temperature – such as for example, large-scale circulation or atmospheric stability – may have a different correlation in day-to-day variability

90   as compared to long time climate change (Lenderink et al., 2017; Fowler et al., 2021a). Thus, although it is important to reproduce observed scaling rates, it is not a sufficient requirement.

In contrast to the anticipated increase of rainfall extremes due to absolute humidity change, the sign of the extreme rainfall response due to relative humidity change is very uncertain (Fowler et al., 2021a, b). This is important as the summer season,

95   in which most of the convective extremes occur in mid-latitudes, is expected to see substantial reductions in relative humidity in a warmer future climate (Lenderink and Attema, 2015; Byrne and O'Gorman, 2018; Williams and O'Gorman, 2022; Zhou et al., 2023). Links to relative humidity can be very complex. Reductions of relative humidity may lead to a smaller number of showers (Dai et al., 2020), for instance because (near surface) air needs to rise to higher levels to reach saturation reducing the cases where convection is triggered. In addition, low relative humidity in the cloud layer leads to a strong dilution of

100   convective updrafts by entrainment processes (Derbyshire et al., 2004), possibly leading to weaker convection and reductions in rainfall extremes. However, lower RH is also connected to larger values of convective inhibition, leading to a longer build-up of convective instability and therefore stronger updraft motions (Rasmussen et al., 2020). Enhanced warming of the surface layer before convection is triggered could compensate the stabilization (in the dry lapse rate) of the atmosphere due to enhanced warming of the upper air; this stabilization occurs as a consequence of a (partly) moist adiabatic adjustment of the atmosphere

105   (Loriaux et al., 2013; Attema et al., 2014). In addition, the stronger evaporation of rain in a drier and deeper atmospheric boundary layer may lead to stronger cold pool dynamics, promoting larger and more organized convective cloud systems (Lochbihler et al., 2021). Finally, deeper and drier boundary layers could lead to wider updrafts at the cloud base, eventually leading to stronger convection (Mulholland et al., 2021). It is therefore important that the dependency of precipitation extremes on relative humidity is well understood and modelled.

110

Within the context of absolute and relative humidity changes, this paper aims to answer the following questions:

- What is the relationship between absolute and relative humidity and precipitation extremes? And how can we establish these relationships with simple measures?
- How do these relationships in parameterized-convection and convection-permitting models compare to observations?

115   Are convection-permitting models substantially better?

- How can we use this information to enhance our (confidence in) future projections of extreme rainfall?

We use data from The Netherlands – a relatively mild and humid climate – and the southern part of France – warmer with lower relative humidity – to study these questions. We also specifically examine commonalities and differences between these

120   two regions. This is potentially interesting since the future climate of the Netherlands is projected to be more like the present-day climate of southern France, with higher temperatures, lower relative humidity and likely decreases in rainfall frequency (Lenderink et al., 2014; Aalbers et al., 2022). We perform a series of analyses to answer the questions above, using scaling on dew point temperature and then further stratifying the data on dew point temperature depression. To complement these analyses, we also examine the distributions of absolute and relative humidity conditional on different rainfall intensity

125   classifications (Lenderink and van Meijgaard, 2010; Lenderink et al., 2011).

2 Methods

2.1 Observations and models

Two observational data sets are used, both with hourly accumulated rainfall and hourly temperature and humidity measurements. For the Netherlands (NL), we use 33 automatic weather stations distributed rather evenly over the country and

130   operated by the Royal Netherlands Meteorological Institute from 1991 to 2020. Most stations have almost complete time series (25 stations with >90% data). For southern France (SFR), 59 stations are used that have at least 10 years data coverage

in the period 1991-2020. These stations are also selected based on their altitude below 400 meters to limit orographic effects. In earlier work we found that scaling relations of hourly extremes derived from these lower altitude stations are quite robust, and not dependent on either the time period or region analysed within western Europe (Lenderink et al., 2021), which is

135 reaffirmed here by considering the differences between central France and The Netherlands. On average the stations in SFR contain 16 years of data, with a maximum of 25 years. Since we pool station data together (and do not examine stations individually) the dataset is large, containing over 900 years of data.
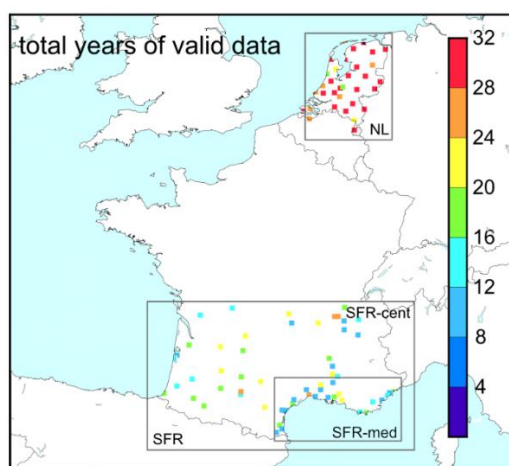
140 **Figure 1: Map of the stations used in the study. The colour scale indicates the number of total years of valid data for each of the stations.**

We use outputs from 7 RCMs: CLMcom, HadRM3, RACMO, RCA4, HIRHAM5, REMO and ALADIN. References for these models can be found in the Supplementary Information. All models have been run at 12 km grid spacing, and we analyse the

145 period from 1991-2010. All models use a convection parameterization, both for deep as well as shallow clouds.

We use 4 different CPMs: HCLIM, AROME, COSMO and UKMO-UM. These simulations are performed on a domain extending sufficiently north to cover The Netherlands. Typically, the runs are 10-12 years, with at least the period 2000-2009 covered in all models (see Supplementary Information for details). For HCLIM two runs are analysed: HCLIM-ALP covers a

150 large domain around the Alps (as prescribed in the CORDEX-FPS study) for the period 1999-2010, and HCLIM-NWE covers a large domain in north-west Europe for the period 2008-2018. All models run without a deep convection parameterization, but some keep a convection parameterization for shallow convective clouds.

All model runs are driven by ERA-Interim reanalysis data (Dee et al., 2011), providing realistic atmospheric forcing boundary

155 conditions. For the CPMs, UKMO is directly nested into ERA-Interim, but the other models use an RCM as an intermediate

nest, providing boundary conditions at high spatial and temporal resolution.

2.2 Analysis

All analysis uses data pooled over many stations (or grid points for model data). Data is pooled for the Netherlands over all

stations (NL). For southern France (SFR) we separate the data set into 25 stations close to the Mediterranean coast (lat <

160 43.8°N, lon > 3.8°E, labelled SFR-med), and 34 stations more inland, labelled SFR-cent (Figure 1).

We mainly show results for the summer period from May to September (MJJAS) in which most of the convective extremes

occur in the Netherlands and SFR-cent. But we also briefly look at the spring to mid-summer period, April to July (AMJJ) and

the late-summer to autumn period, August to November (ASON). Results for SFR-cent and NL are very robust, with very

165 distinct scaling behaviour only (very) weakly dependent on the season analysed. In contrast, results from SFR-med reveal how

a more complex meteorology, with influences from the Mediterranean Sea as well as the nearby orography, affects the results.

For the CPMs, we first re-grid the hourly data based on blocks of 5x5 grid points. For these 5x5 grid blocks, we compute the

mean ($Pmean_{5x5}$) and take the value from the point centred in the middle of the 5x5 box ($Psample_{5x5}$, called *sample* as it samples

170 the 25 grid points in the block). By comparing $Pmean_{5x5}$ with $Psample_{5x5}$ we can establish the dependency on spatial scale, and

the scale of $Pmean_{5x5}$ is approximately comparable to RCM grid box scale. Here, we use $Psample_{5x5}$, which is most

representative for the local station observation. However, we note that very similar results are obtained using $Pmean_{5x5}$, with

dependencies not strongly dependent on spatial scale.

175 The closest grid point to the station location is used, without spatial interpolation. For the model results we take the complete

time series, and we do not correct for missing observational data by neglecting the same hours from the model results.

Uncertainties in the observations are computed using a bootstrap procedure. We take 10 random years from the 1991-2020

period and calculate all analysis statistics on the 10 year sampled data. We note that we take 10 years as comparison to CPM

180 model simulations with a ~10-year length. We take 100 bootstrap samples and plot the 5-95% range of samples. Since the

data in France contains more missing data, this implies that uncertainties in the observations for this area are somewhat larger

than in the models, yet experimenting with 15-year bootstrap samples only gave small differences.

In this study we use the atmospheric temperature T, dew point temperature TD, and dew point depression, DPD = T – TD, all

185 near the surface at 2m. The dew point is defined as the temperature at which an air parcel reaches saturation (100% relative

humidity) when cooled adiabatically and at constant pressure; TD is therefore a measure of absolute humidity, and each degree

rise in TD reflects an increase of 6-7% in the humidity of the air following the CC relation. The dew depression is a good measure of *relative* humidity, which follows from the fact that the saturation specific humidity varies approximately exponentially with temperature (according to the CC relation). Each degree of DPD increase represents an approximately 4-

190   6% drop in relative humidity.

Dependencies on TD (absolute humidity) and DPD (relative humidity) are computed from two procedures (following Lenderink and van Meijgaard 2010). For both procedures we first pair the instantaneous values of dew point temperature to the precipitation 4 hours later (or closest to that). We note that some models only provide 3-hourly (dew point) temperature;

195   hence, using a temperature 4-hours prior guarantees that the (dew point) temperature is taken before the precipitation, so that it is not likely strongly affected by the rainfall event itself. Using a lag of 4 hours also provided the most robust results in earlier scaling analyses based on observations and, in physical terms, limits the influence of showers themselves on the surface temperature by cold (and dry) downdrafts (Lenderink *et al* 2011). All further analysis is done on this paired data set (temperature on a wet hour paired with temperature at 4 hours preceding the hour with rainfall).

200

We first compute TD and DPD statistics conditional on rainfall intensity. We sub-select pairs based on rainfall intensity, taking all hours including those without rain (all), all hours with rain exceeding 0.1 mm hour^{-1} (wet), and hours with rain exceeding the 50, 90, 95, 99, 99.5 and 99.9th percentiles of rain (with the percentile computed from wet hours only). Subsequently, we compute various statistics from the selection of TD and DPD, for instance, the median TD.

205

For the scaling analysis, we compute rainfall intensity conditional on TD. First, the pooled data is divided into two-degree wide TD bins. As in previous studies we use overlapping bins with steps of one degree. Conditional percentiles (90, 99, 99.9th) are computed for each bin, taking only hours with rainfall exceeding a threshold of 0.1 mm hour^{-1}. Besides binning on TD only, we also distinguish between different relative humidity classes based on DPD, with DPD < 3 °C marking high relative

210   humidity, DPD from 3 to 6 °C marking medium relative humidity, and DPD > 6 °C marking low relative humidity. This procedure divides the data into three approximately equal parts, although obviously the drier climate in SFR leads to more data in the low relative humidity class.

Finally, in some figures with many lines from different models, results are slightly smoothed using a LOESS filter (Chan et

215   al., 2016) in order to make the figures less cluttered and easier to understand (this is mentioned in the figure caption). Bootstrap uncertainties are only given when the statistics can be derived in more than 90% of the bootstrap samples. Percentiles are only computed when there are more than 20, 200, and 2000 data points for the 90, 99 and 99.9th percentiles, respectively. The distributions of dew point and dew point depression based on rain intensity class are only computed when there are more than 100 paired data points. Some statistics are plotted outside the range for which the bootstrap uncertainty could be established

220 (note that the bootstrap uses only 10 years instead of the 30 years in the full data set), but obviously the values in this case are very uncertain.

3 Results

3.1 Extreme statisics

Before analysing temperature and humidity dependencies, we first examine the rainfall distribution derived from the pooled

225 dataset for both regions in southern France, SFRA-cent and SFR-med, and the Netherlands, NL. Plotted in Figure 2 is the *probability* of exceedance derived from the pooled dataset over all stations and all time-steps. The most striking result is that the ensemble of RCMs has a much larger spread than the CPM ensemble. All of the RCMs underestimate hourly rainfall for NL up to moderate intense events (10 mm hour^{-1}). In the extreme tail about half of the models still substantially underestimate, whereas the others catch-up; even leading to an overestimation for HIRHAM5. We note that these high extremes in the RCMs

230 could be related to unphysical grid-point storms (Chan et al., 2014). In RACMO, however, we also observe that the *resolved* dynamics produce a convective-system-like up and downdraft when the atmosphere is very moist, producing a convective rain systems that is too large scale and too persistent (see Fig. S8). For SFRA-cent, the situation is similar, yet one model (HadRM3) appears to be surprisingly good except for a very few high intensity events (>50 mm hour^{-1}).

235 CPMs generally show much more consistent results, with the majority of the models close to the observations for intensities below 20 mm hour^{-1}. However, except for UKMO-UM, they appear to underestimate the extreme tail of the rainfall distribution in SFRA-cent and SFR-med; a behaviour that is also apparent for NL, with models not exceeding 50 mm hour^{-1}, with the exception again of UKMO-UM which is now too extreme. Summarizing, in line with previous studies, CPM results show more consistent behaviour across the multi-model ensemble, with extreme rainfall distributions closer to the observations than

240 for the RCMs.
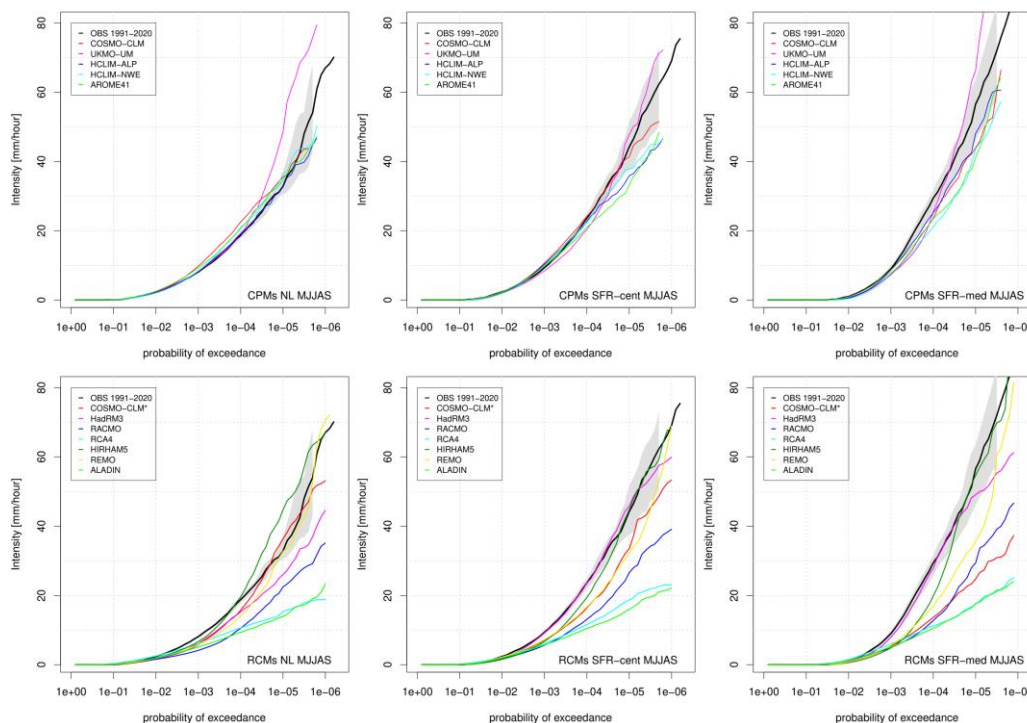
Figure 2: Probability of exceedance derived from the pooled data of all stations for MJJAS. Grey bands around the observations for 1991-2020 are 90% uncertainty ranges estimated from bootstrapping.

3.2 Dependencies conditional on rain intensity

Before discussing dependencies on rainfall intensity, we first look at the climatological distribution of near-surface dew point temperature, TD and DPD. Results for NL and SFR-cent are shown here in Figure 3 (left point labelled with "all" in graph). Since CPMs and RCMs are, except for higher resolution, not fundamentally different in many aspects of the parameterized processes in the soil and in clouds, we do not necessarily expect better behaviour from the CPMs here. In addition, the larger intermittency and higher rainfall intensity in CPMs may lead to more runoff and drier soils (Berthou et al., 2020). Indeed, looking at the median TD for all hours and the median TD for wet events in SFR-cent, CPMs do not show better behaviour than RCMs, with both HCLIM as well as UKMO-UM showing a dry bias corresponding to the low-end of the RCM ensemble. This dry bias is fairly constant across the full TD distribution in most, but not all, models indicating that the variability in TD is approximately correct (Fig. S2, panels on the left).

As expected, higher intensity rainfall events occur on average with higher TD, gradually increasing with the extremity of the rainfall event; see Figure 3 (left panel) showing the median TD for different rainfall intensity classes. CPMs and RCMs display rather similar behaviour in this respect. Typically, the slope of the line is parallel to the observed slope showing that a model with a dry mean bias typically also produces extreme rainfall at too low absolute humidity, and vice versa. This general behaviour is also obtained by investigating the full distributions of TD, instead of only the median shown in Figure 3 (see Fig.

S2). There appear to be no substantial systematic shifts in the anomaly of TD in the two model ensembles with rainfall intensity class (all, wet and 99th percentile). Thus, the differences in the extreme rainfall statistics (Figure 2) between the RCMs and CPMs cannot be easily explained by differences in absolute humidity alone.

265

The observed dewpoint depression, DPD, shows a moderate increase with rainfall intensity class for NL and a more pronounced increase for SFR-cent, revealing that more intense rain occurs on average at lower values of relative humidity (Figure 3). For SFR-med (Fig. S1) there is almost no dependency, which may be related to the more complex meteorology connected to rainfall extremes in this region (Duffourg and Ducrocq, 2013). The CPM ensemble represents this dependency

270 rather well. But, particularly for SFR-cent, the spread is considerable and outside of the observed uncertainty ranges, with most CPMs simulating too high DPD (too low relative humidity). The RCMs are clearly biased. Except for HadRM3, they all require too low values of DPD (too high relative humidity) to produce extreme events. This is the case for all three regions studied (Figure 3, and Fig. S1). Potential causes of this behaviour – related to how these models represent convection – will be elaborated on in the discussion.

275

Hydrology and
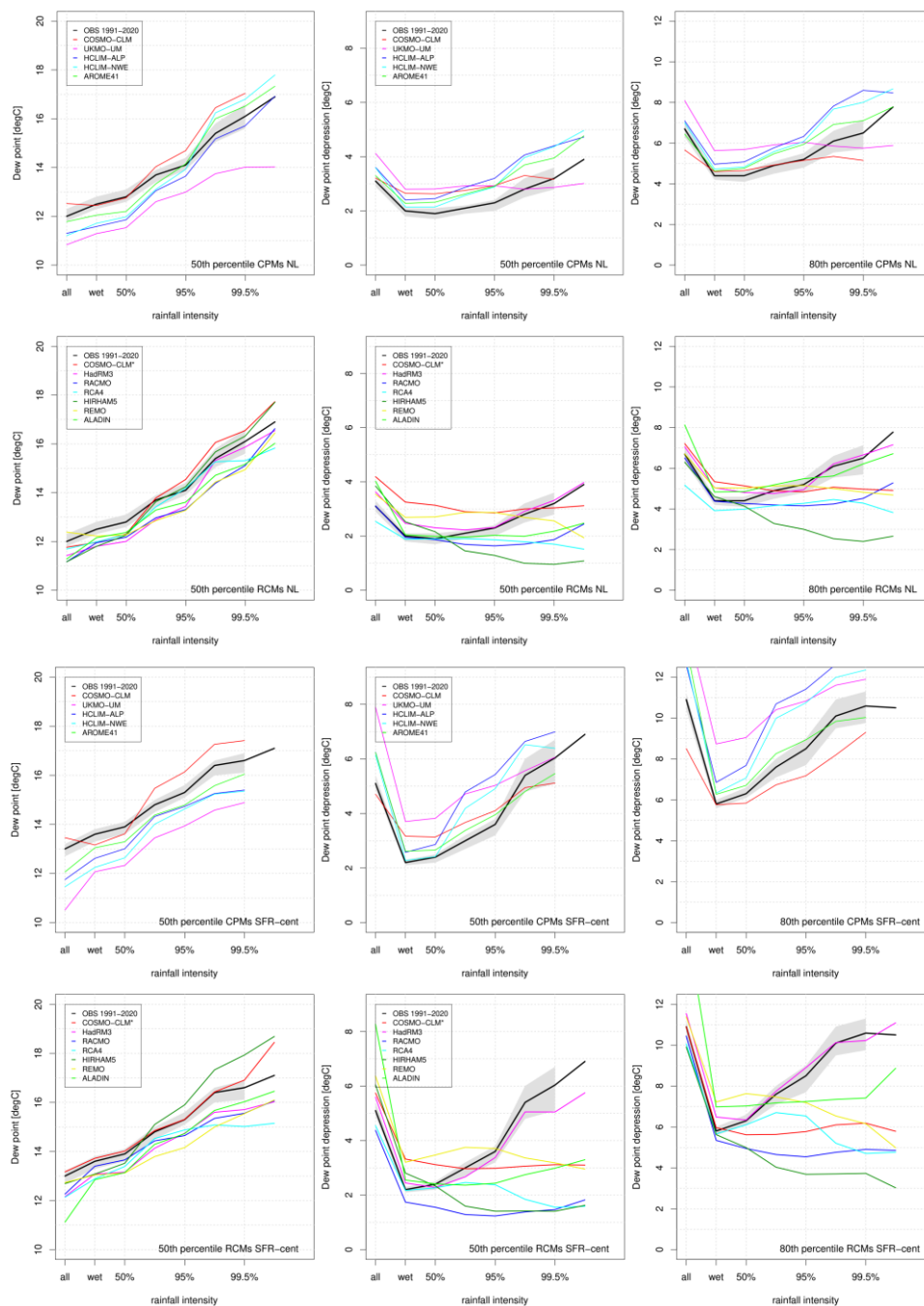Earth System
Sciences

Discussions

Figure 3: Median dew point temperature, TD, conditional on hourly rainfall intensity (left), and median (middle) and 80th percentile (right) of the dew point depression, DPD; for all hours (that is, climatology, including dry hours), wet hours, and those with hourly rainfall exceeding the conditional 50, 90, 95, 99, 99.5, and 99.9th percentiles. First two rows show results for NL, and last two rows show results for SFR-cent; results of the CPMs in 1st and 3rd row, and RCMs in 2nd and 4th row. We use paired precipitation and TD time series (TD are 4 hours prior to rainfall). (Statistics are only computed with more than 100 paired measurements, causing the shorter 10-year bootstrap samples to give no data for the 99.9th percentile rain intensity class).

285 To further investigate the robustness of this result, Figure 4 shows the full distribution of the DPD anomaly compared to the observations for rainfall events exceeding the 99th percentile. The difference between the RCMs and CPMs is again obvious; the RCMs are almost all substantially below the zero line, reflecting the occurrence of rainfall extremes at predominantly too high relative humidity. The CPMs, however, do not show this behaviour, and many of them even seem to display a reversed dependency; they tend to produce heavy rainfall at too low relative humidity.

290

Understanding the underlying causes of the differences between the model derived dependencies and observations is difficult with such a statistical analysis. A model bias in the DPD distribution for extreme events could be related to a bias in the DPD climatology of the model but could also reflect differences in how convective rainfall responds to relative humidity. As an example, it could be that convective rainfall in a model is only triggered in case the boundary layer is close to saturation due 295 to convection parameterization (Hohenegger et al., 2009). In this case, a model could have mean bias with too low relative humidity, but still display too high relative humidities for extreme rainfall events.

Thus, we argue that by considering how the anomaly in dew point and dew point depression depends on rainfall intensity class we can get a hint towards the causal mechanisms. Our conjecture is that when these anomalies depend strongly on the intensity 300 class, the role of parameterization and convection-related mesoscale processes is important. In contrast, when it is mainly independent then other mechanisms – for example, related to soil memory and runoff processes – could be more relevant. But because a large fraction of the total rainfall falls as "extreme" and because runoff is a function of rainfall intensity these two effects are intrinsically intertwined.

305 Acknowledging the subtleties in the interpretation described above, the RCMs show a very clear shift to high relative humidity (low values of the dew point depression) for high intensity rainfall events (Figure 3, Figs. S1-3). This behaviour is most strongly present in SFR-cent, but also visible for NL and SFR-med (Figure 4, and Figs. S1-3). There is one remarkable exception to these results: the Met Office model (HadRM3) is almost bias free in this respect. For the CPMs, however, we do not observe such a dependence on rainfall intensity, suggesting that general model climatological biases, for instance due to 310 dry soils, are more important.
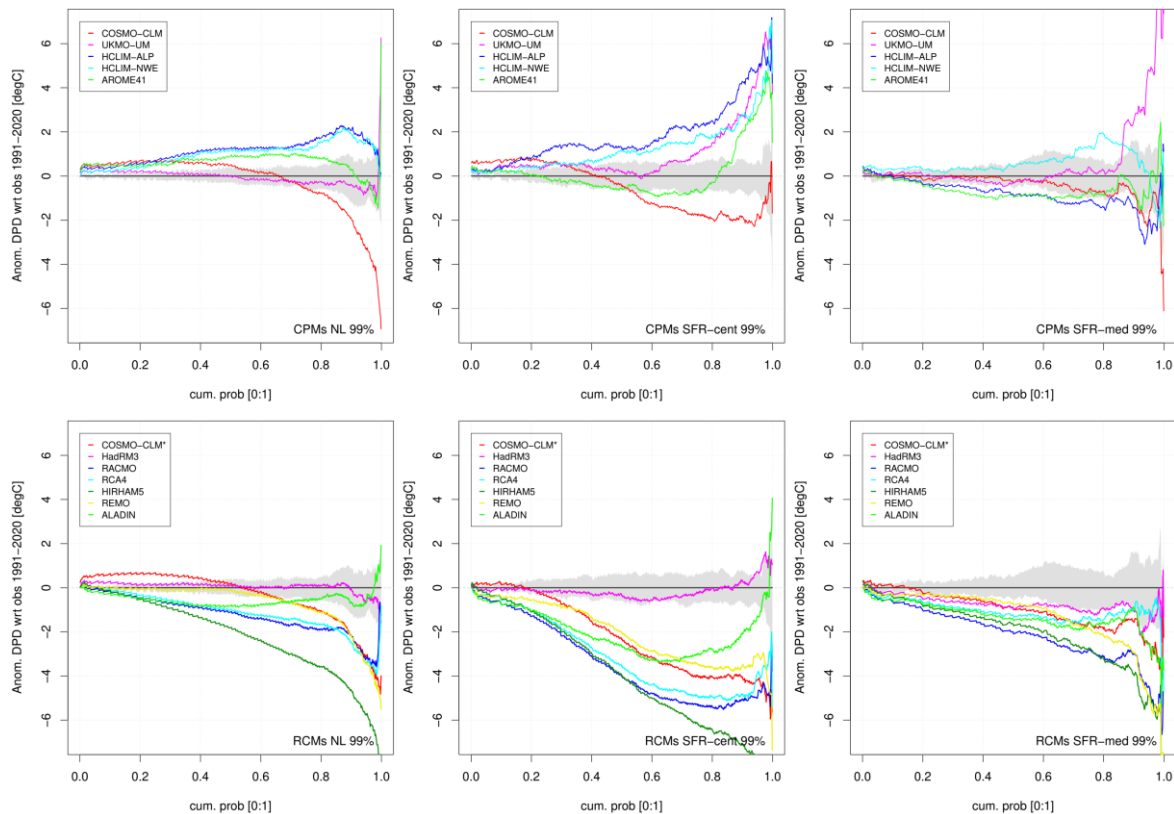
Figure 4: Anomalies with respect to the observations (1991-2020) of the distribution of dew point depression, DPD, for events exceeding the 99th percentile of rainfall intensity. Plotted is the delta between model and observations as function of the cumulative probability, similar to quantile difference plot.

3.3 Dependencies conditional on absolute and relative humidity

Humidity dependencies are studied in more detail using a scaling analysis. Figure 6 shows that the highest percentiles, the 99 and 99.9th, clearly reveal close to 2CC dependencies (13% per degree) on TD for the observations in NL and SFR-cent. This behaviour is surprisingly robust for both areas, with almost the same behaviour obtained by analysing the spring/early summer period (AMJJ) and late summer/autumn period (ASON) (Fig. S4). Following earlier studies, we use conditional percentiles – investigating rainfall intensity preconditioned on the occurrence of rain – as these show the most consistent behaviour; e.g. results show larger bootstrap uncertainty bands and less consistent scaling rates across the TD range using absolute percentiles. This is consistent with rainfall intensity rather than rainfall frequency being related to the moisture availability in the atmosphere. But we note that the results using absolute percentiles are quite similar, with only somewhat worse behaviour for the lowest percentiles (Fig. S5; and see results in the discussion).

For the SFR-med, positive dependencies on dew point temperature are obtained, but these dependencies are less robust and there are quite strongly varying scaling rates with dew point temperature (Figure 5) as well as seasonality effects (Fig. S4). The indicates local near surface humidity is not the main source of moisture (or not correlated well with it) for rainfall (e.g. moisture at higher levels may be more important) or other factors (e.g. specific dynamical circulation patterns) are important

335 for controlling rainfall intensity. This may reflect the more complex topography and associated meteorology of the region. Given the less reliable scaling results for SFR-med, we hereafter focus our analysis on results from NL and SFR-cent.
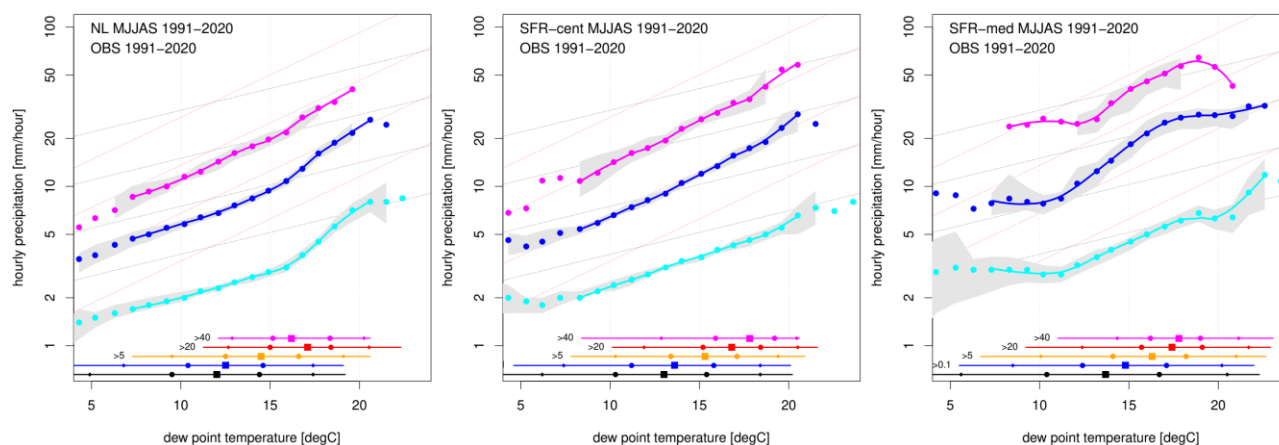
340

Figure 5: Scaling of hourly observations of precipitation extremes on dew point temperature, TD, with cyan, blue, and magenta lines showing the 90, 99, and 99.9th percentiles based on wet hours only. Dotted black and red lines show a dependency of 1 and 2 times the CC relation. Horizontal lines at the bottom show the dew point temperature range for all hours (black), hours with rain (blue) and hours with rain exceeding 5 (orange), 20 (red) and 40 (magenta) mm, with the line showing the 1 to 99th percentile range and
345 **markers indicating the 5, 25, 50 (square), 75, 95th percentiles. Grey bands are uncertainty estimates from the bootstrapping procedure. Loess smoothed lines are plotted for a relevant TD range (orange line, showing the 1-99th percentile range of TD for hourly rainfalls exceeding 5 mm), and uncertainty bands are plotted where more than 90% of the bootstrap samples contain sufficient data (plots for individual models are given for NL and SFR-cent in the supplementary information).**

For further investigation we use the 99th percentile, which is a compromise between analysing sufficiently extreme events and
350 still containing sufficient data for a robust estimate. For this percentile we further classify on DPD as a measure of relative humidity. In general, scaling rates – the rate of increase in intensity per degree TD – are rather similar for the different DPD classes. Results for SFR-cent (Figure 6) and NL (Fig. S6) all show a dependency close to the 2CC rate, with only small variations with dew point temperature.

355 The CPMs generally reproduce the observed dependencies rather well (Figure 6, upper panels, and Fig. S6). Most CPMs tend to overestimate the rainfall intensities for SFR-cent and NL; this is most prominent for high relative humidities. The RCMs reveal much more diversity in scaling behaviour (Figure 6, lower panels). Generally, their hourly rainfall amounts are substantially too low. But some model results appear to be reasonable; for example, HadRM3 overall shows good comparison

with observations. Yet, other RCMs show very divergent behaviour. HIRHAM5 is an (extreme) example. The model has
360   reasonable dew point temperature scaling for high relative humidity, but is far off for low relative humidity (consistent with
the erroneous relative humidity dependency of HIRHAM5 shown in Figure 4). This inability to reproduce the correct TD
scaling for low relative humidity is apparent in most RCMs; for instance, for SFR-cent most RCMs show hardly any
dependency between 10 and 20 °C (RACMO, RCA, REMO, HIRHAM5). As relative humidity is expected to decrease in the
future climate, this could imply that expected increases in rainfall extremes are underestimated in these models.

365

To summarize these findings, we plotted rainfall amounts at 15 °C against the scaling coefficient for the 99th percentile in
observations (black, with estimates of uncertainty from the bootstrap), RCMs (open triangles) and CPMs (solid squares)
(Figure 7). These results are derived from fitting a linear dependency to the logarithmic of precipitation between 10- and 20-
degrees dew point temperature and plotting the fit coefficient against the value at 15-degrees. For the observations of NL and
370   SFR-cent this is rather accurate (and resulting errors are small) since the scaling lines in Figure 5 are almost linear in that dew
point temperature range. For all relative humidity classes – all, low and high relative humidity – and both regions, scaling rates
are close to 2CC in the observations. In addition, rainfall intensities are higher for low relative humidity as compared to high
relative humidity, typically increasing from 10 mm to 15-17 mm.

375   The CPMs tend to slightly overestimate the scaling rates for high relative humidity and the moister NL area. Conversely, they
tend to underestimate scaling rates for low relative humidity, as visible for SFR-cent. Overall, however, scaling rates are rather
accurate in the CPMs. There is a small overestimation of the intensity of rain; this could be partly related to an underestimation
of rain frequency, as will be discussed below. But generally CPMs also have updrafts forced to occur at too large a scale (the
km grid scale) due to insufficient turbulent mixing, causing also an overestimation of the rainfall intensity.

380

Compared to the CPMs, the results derived from the RCMs, are clearly worse. Scaling rates are on average accurate for high
humidity, as visible for NL. But, as soon as the relative humidity drops, scaling rates start to become too low; this is most
prominent for SFR-cent, where several RCMs display almost no dependency on dew point temperature. We note that the
CPMs tend to struggle here too, with average dependencies close to the CC rate. But overall, the CPMs are clearly in better
385   agreement with observations and have lower inter-model spread.

As an example of where the CPMs improve on the RCMs, we plotted the difference in rainfall intensity between low and high
relative humidity as a function of TD (Figure 8). In the observations, we obtained for NL and SFR-cent substantially more
rainfall for low relative humidity in the TD range between 10 and 20 °C; peaking at around 15 °C, where hourly rainfall is 70-
390   100% more intense for low relative humidity. Most of the CPMs capture this intensification of precipitation for lower relative
humidity, but they also tend to reach peak values at too low TD. In contrast, the RCMs do not consistently capture an
intensification with lower relative humidity (with the exception again of HadRM3).
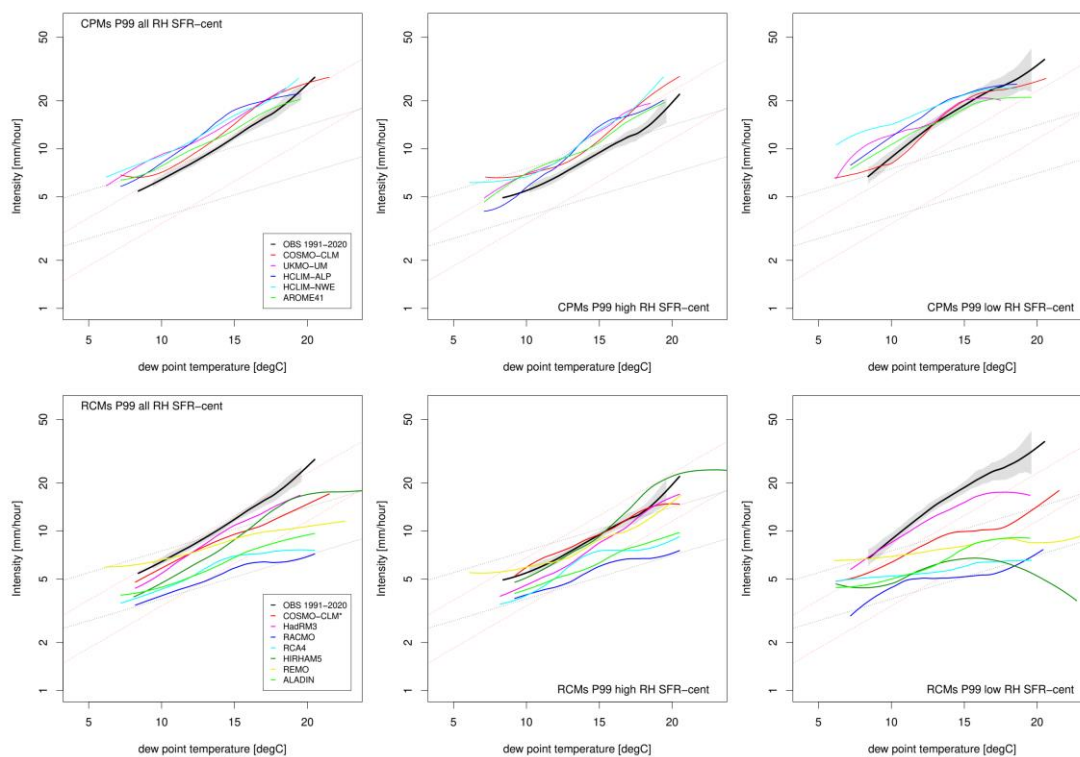
15

Figure 6: Scaling of hourly precitation based on, from left to right, dew point temperature, dew point temperature for hours with low relative humidity, dew point temperature for hours with high relative humidity. Results are for the CPMs, and upper panels are for NL, and lower for SFR-cent. Lines filtered with LOESS filter (span = 0.5).
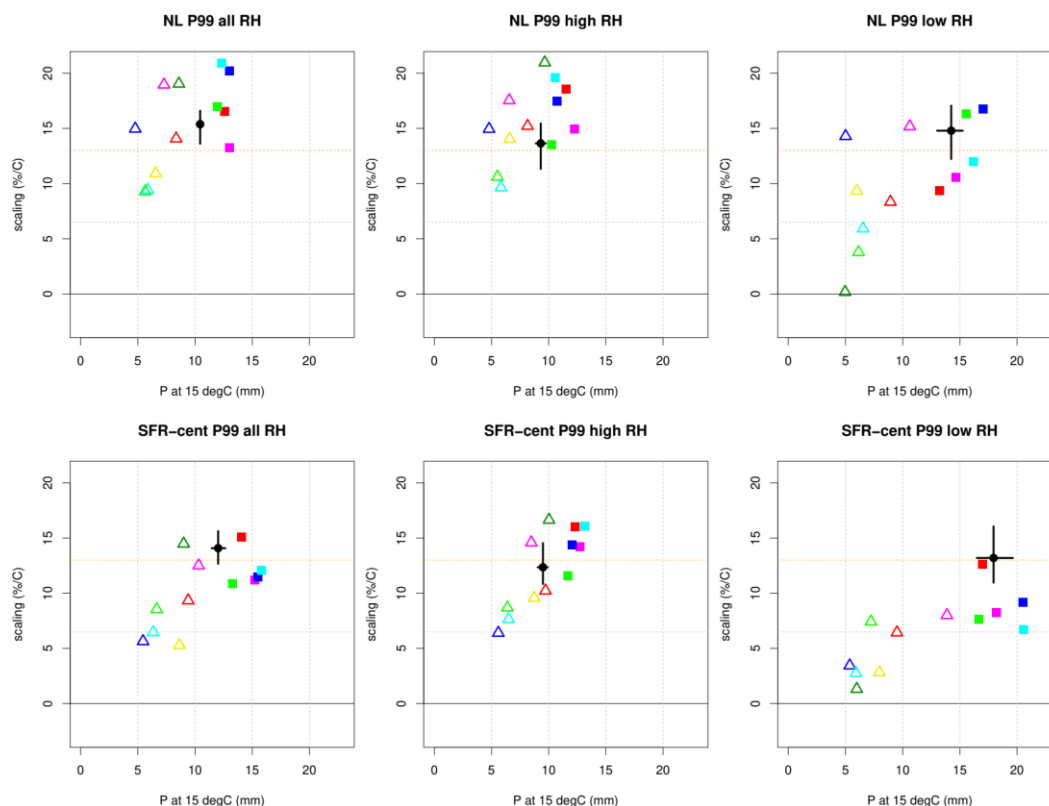
Figure 7: Intensity at 15 °C versus scaling coefficient for the 99th percentile of hourly precipitation. Both are derived from a linear fit to log(P), fitted between 10 and 20 °C dew point. From left to right, NL, SFR-cent and SFR-med, and from top to bottom, all data, and data selected on low RH and high RH. CPM results are shown by solid squares, whereas RCM results are open triangles, and colouring is the same as Fig. 6.

Besides rainfall intensity, which is conditional on the occurrence of rainfall, the frequency of rain is important as well. As expected, the frequency of rainfall is much higher for high relative humidities (of around 10-15%) than for low relative humidities (1-4%) (Figure 9, note the difference in scale). The wet hour frequency for SFR-cent and NL are surprisingly similar for high relative humidity, with values between 0.1 and 0.14 (Figure 9 and Fig. S12). This is not the case for low relative humidity, where the frequency of rain events is lower in SFR-cent than NL, possibly related to the lower relative humidity conditions in SFR-cent.

Hydrology and
Earth System
Sciences
Discussions
EGU
Open Access
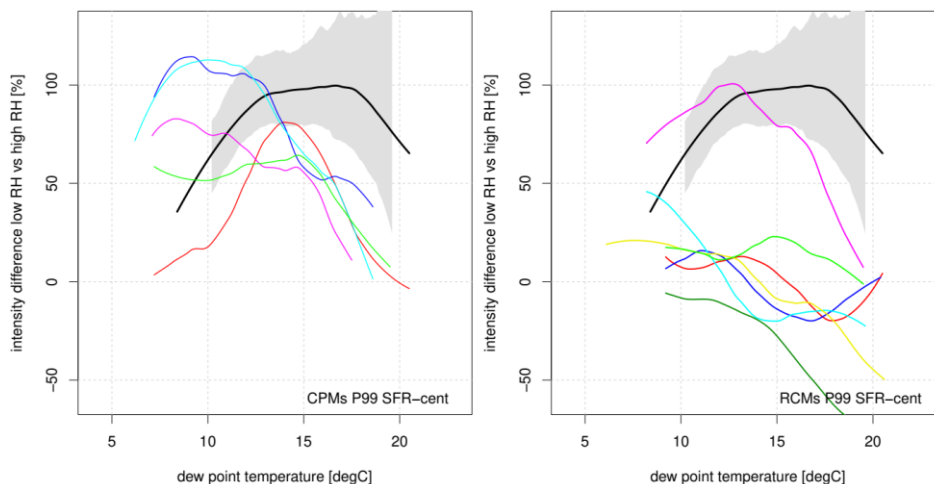
415

Figure 8: Difference in intensity between low and high relative humidity (%) as a function of dew point temperature for SFR-cent (results for NL in the supplementary information). Colors are the same as Fig. 6.

The CPMs generally reproduce the observed rainfall frequencies rather well (Figure 9). The CPMs tend to underestimate the

420 frequency of rainfall for high relative humidity, particularly for the high TD range (except for AROME). For low relative humidity they are approximately correct (except for COSMO in SFR-cent, which appears to underestimate). The RCMs clearly show more divergent behavior. In general, they strongly overestimate the frequency of rainfall (Figure 9)— a common deficiency of this type of model (Berthou et al., 2020; Ban et al., 2021; Lucas-Picher et al., 2021).

425 The dependence of the frequency of rainfall on relative humidity could have several causes, partly related to relative humidity directly affecting rain processes but also to confounding large-scale atmospheric conditions. Higher lifting condensation levels and higher values of convective inhibition associated with lower relative humidity could prevent an updraft from reaching its level of free convection and therefore suppress convective showers. However, by selecting on low relative humidity, large-scale atmospheric conditions are favoured that are more hostile to the development of convective showers. For instance, high

430 pressure systems cause low relative humidity as well as conditions in which showers are unlikely to develop, due to a lack of large-scale moisture convergence or stable tropospheric temperature profiles. The biases in model behaviour therefore cannot be easily attributed to the local cloud processes or to the confounding large-scale factors. Nevertheless, the overall negative bias in the simulation of the frequency of rainfall in the CPMs for high relative humidity, in particular for high TD, suggests that CPMs have difficulties in triggering convection in cases where the surface is relatively cold – noting that high relative

435 humidity implies a cold surface temperature for a given TD – and therefore for cases where the surface forcing is relatively weak.
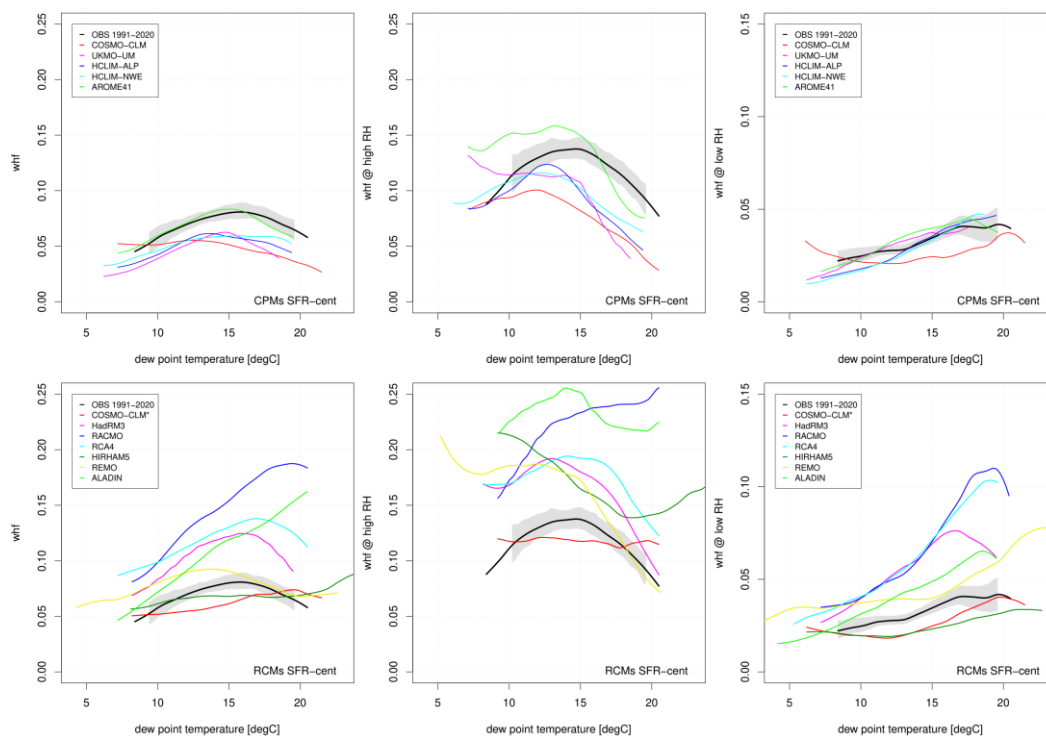
Figure 9: Fraction of wet events as a function of dew point temperature, TD, for all hours (left), high relative humidity(middle) and low relative humidity RH (right) (note the difference in plotted range between low and high relative humidity). Lines filtered with LOESS filter (span = 0.3).

The biases in the frequency of occurrence of rainfall may also affect results from the scaling analysis (Schär et al., 2016). We investigate this further in the discussion but note here that the main conclusions remain the same.

3.4 Area differences

Finally, we inspect differences between the extreme rainfall statistics in SFR-cent and NL. In addition to correctly representing humidity dependencies, a good representation of the systematic differences between these two contrasting regions adds confidence to the trustworthiness of these models.
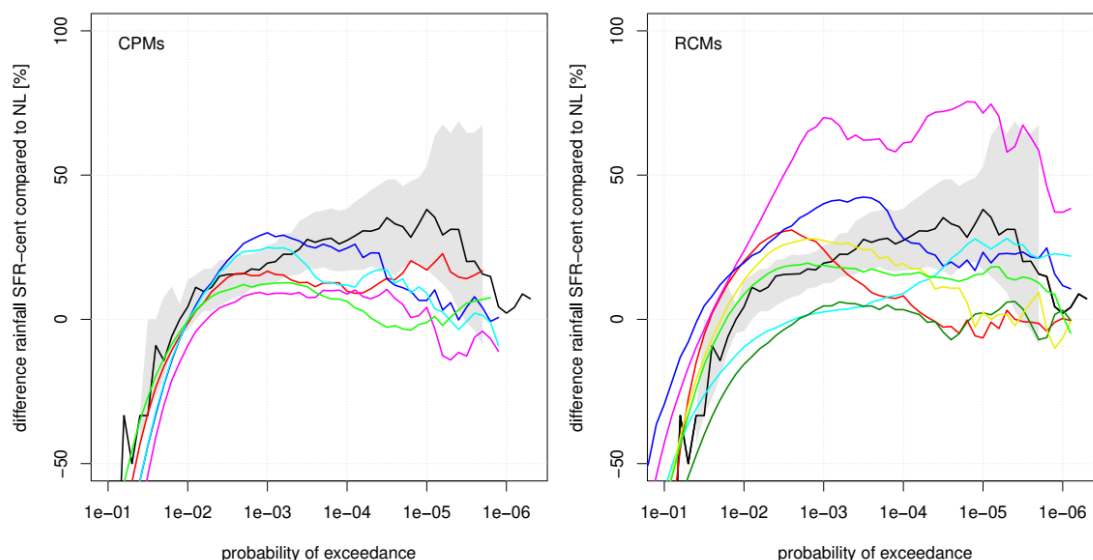
Figure 10: Difference in hourly rainfall between SFR-cent and NI. The difference is the percentage difference (SFR-med minus NL) as function of the probability of exceedance estimated from pooling all stations within the region and for the summer half year.

455   Before looking at the rainfall distribution differences between NL and SFR-cent, it is useful to inspect the differences in absolute and relative humidity in more detail. In general, dew point temperatures are slightly higher in SFR-cent as compared to NL; the climatological value is approximately 1 °C higher, and this difference decreases to almost zero for the most intense rain events (Figure 3, noting the difference in slope, and also Fig. S13). The differences in relative humidity are more pronounced. The median dew point depression is approximately 3-4 °C for the most extreme rainfall events in NL, whereas it
460   is 5-6 °C for SFR-cent with the largest difference for the most extreme rainfall. Thus, rainfall extremes in SFR-cent occur at slightly higher humidity and substantially lower relative humidity as compared to those in NL.

The difference in (extreme) rainfall is measured by $D_{SFR\text{-}cent-NL}$, which is defined as the fractional difference (in %) in hourly extreme rainfall (as a function of the probability of exceedance) with respect to the value for the NL. We note that results
465   beyond a probability of exceedance of ~10^{-4} (~ 20 mm hour^{-1}) are less certain, as shown by the grey shaded areas based on the bootstrap procedure. Values for $D_{SFR\text{-}cent-NL}$ are positive for a probability of exceedance of less than 10^{-2}, showing that rainfall extremes in central southern France produce 10-30% more rain in an hour than in the Netherlands. Based on the differences in humidity it appears likely that this is partly explained by the higher humidity values, and partly by the lower relative humidity. In particular, for the most extreme rainfall hours, the latter could be more important.

470

The CPMs capture the observed behaviour in $D_{SFR\text{-}cent-NL}$ well but are slightly too low for intermediate values of the probability of exceedance between 10^{-2} and 10^{-4}, yet often within the uncertainty margins of the observations. For rarer events, with a

20

Hydrology and
Earth System
Sciences
Discussions

probability of exceedance less than ~10^{-4}, modelled $D_{SFR\text{-}cent-NL}$ values are (generally) 10-20% too low and are outside the uncertainty margins of the observations. We believe this to be primarily caused by the lower relative humidity in SFR-cent,

475 and the general tendency (small, but consistent) of the CPMs to underestimate rainfall extremes for low relative humidity.

The RCMs, on the other hand, display more diverse behaviour, and are generally further outside the uncertainty margins of the observations. However, the average across the RCM model ensemble does not appear to be worse than the CPM ensemble. This is at odds with the expectation that the too strong reduction of rainfall intensity to decreases in relative humidity should

480 lead to a more pronounced underestimation of $D_{SFR\text{-}cent-NL}$ as compared to the CPMs. We could not find a clear reason to explain this finding. It appears that $D_{SFR\text{-}cent-NL}$ is determined by a combination of different factors: the rain frequency and how it depends on relative humidity and other factors, the humidity dependencies of rain intensity, and the differences in underlying climatology. The interplay between these factors is complex, and most RCMs appear to suffer from compensating errors which are hard to disentangle.

485

As one example of this complexity, we discuss results from HIRHAM5. By all measures, HIRHAM5 has a very strong erroneous relative humidity dependency, only producing extreme rainfall for high relative humidity. Given the lower relative humidity of SFR-cent, one would therefore expect $D_{SFR\text{-}cent-NL}$ to be negative, but in fact values close to neutral are obtained. Part of this could be explained by the finding that, despite a mean negative bias, dew point temperatures are overestimated in

490 the upper range of the dew point temperature distribution by up to 2-4 °C (Fig. S2, and also visible in Figure 6 with scaling extending to much higher dew point temperatures). Similarly, the dew point depression is underestimated (Fig. S14). Thus it appears that compensating effects play a role here, and speculate that these may related to soil processes and soil memory, and how these interact with rainfall.

495 **4 Discussion**

Different measures have shown that the sensitivity of hourly extremes to relative humidity is essentially different in RCMs compared to CPMs, and that CPMs behave (much) more realistically in this respect. Lower relative humidity tends to lead to stronger rainfall extremes compared to high relative humidity: extremes co-occur on average with lower relative humidity (Figure 3) and at a given dew point temperature (given absolute humidity) rainfall intensities are higher for low relative

500 humidity (Figure 8).

It could be that these results are (partly) caused by a statistical effect caused by taking conditional percentiles. One may argue that low relative humidity suppresses showers when the large-scale conditions – for example, convergence at large scales – are not very favorable, taking light showers out of the rainfall distribution, but hardly affecting the occurrence of heavy

505   showers. In that case, the statistics of the extremes could still be the same, but the conditional percentiles would give an artificial increase (Schär et al., 2016). Also, this effect could play a role in explaining the differences between the models and the observations as the RCMs have large errors in the frequency of rain.

For this reason, we also considered the unconditional percentiles. The disadvantage of unconditional percentiles is that scaling
510   results become less robust, more dependent on the percentile, and more variable over the dew point temperature range (Figs. S5 and S15). But the difference is not very large, and is most pronounced for the lowest percentiles where the bootstrap sampling gives wider uncertainty estimates. The latter suggests that uncertainty in large-scale circulation conditions, which are captured by the resampling procedure, mostly affects the rain frequency leaving the conditional intensity distribution unaffected.

515

Taking the 99.9th percentile (roughly corresponding to the 99th percentile based on wet hours only) we still find that low relative humidity leads to stronger precipitation extremes. Yet, the degree of the effect is substantially reduced. In general, we now find that the extremes are 10-20% more intense in the case of low relative humidity (Figs. S15-17) – far below the 50-100% from estimated from the conditional percentiles. However, one should note that we now likely suffer from a reverse
520   compounding effect; low relative humidity is likely to co-occur with atmospheric high pressure systems. Thus, by comparing low with high relative humidity samples, we also sample the influence of large-scale circulation, for low relative humidity likely leading to a more hostile environment for convection to occur with less moisture convergence at larger scales and more stable lapse rates. Thus, we think that that this estimate is a lower bound and that in reality – comparing only systematic differences in near surface relative humidity, but otherwise similar other atmospheric drivers like stability and large-scale
525   convergence (Lenderink et al., 2017) – the sensitivity is in between the high estimate derived from the conditional rainfall distribution and the low estimate presented here. This collaborates also with the finding that the most extreme events co-occur with lower relative humidity (Figure 3).

Physically, several processes could lead to an enhancement of rainfall extremes at lower relative humidity. At a fixed value of
530   the surface dew point temperature, lower relative humidity implies higher surface temperatures, so that low relative humidity is likely associated with atmospheric conditions with stronger heat forcing from the surface and higher values of convection inhibition (CIN). Likewise, lower relative humidity could be associated with mesoscale circulations transporting moisture to drier regions (Klein and Taylor, 2020; Hohenegger and Stevens, 2018). Lower relative humidity could lead to stronger cold pool dynamics because the deeper and drier (in terms of relative humidity) boundary layer causes more evaporation of rainfall
535   (Lochbihler et al., 2021). The growth and collision of cold pools have been shown to play an important role in converging moisture and triggering new precipitation cells and this process could lead to bigger and more vigorous cloud systems

22

(Lochbihler *et al* 2021, Haerter and Schlemmer 2018). Also, it has been argued that deeper boundary layers lead to wider updrafts at the cloud base promoting stronger convection (Mulholland et al., 2021)

540 Conversely, low relative humidity could lead to weaker extreme rainfall as entrainment of dry air into the updraft motions could lead to strong cloud erosion and deep and dry boundary layers could lead to substantial evaporation of rain before it reaches the ground (Fowler *et al* 2021a, Derbyshire *et al* 2004). Also, the atmosphere may become too dry even to trigger moist convection if the lifting condensation level is not reached or values of CIN are too high. Since, in the limit of relative humidity approaching zero, there will be no moisture and therefore no rainfall, these processes that act to weaken extreme
545 rainfall must start to dominate at a certain (low) relative humidity value.

The causes of the relatively poor performance of the RCMs in representing the relative humidity dependencies are complex, but likely the consequence of the way the convection scheme interacts with its environment and lack of ability to capture local scale dynamics – with the parameterizations essentially treating this as a 1D problem whereas, since there is limited moisture
550 available in an atmospheric column, it is a 3D problem where moisture convergence is essential to sustain high rainfall rates. A number of processes leading to the enhancement of rain extremes with lower relative humidities are not parameterized or are poorly resolved in most RCMs, such as for instance the interactions between cold pools. Many mass flux based convection parameterizations use strong assumptions on the mass of air entering at cloud base, the lateral entrainment and detrainment of air into the cloud, lag memory from one time step to the next, and often suffer from numerical noise (Lenderink et al., 2004;
555 Hohenegger et al., 2009; de Rooy and Pier Siebesma, 2010; Yano et al., 2013). Although we cannot pinpoint a specific process being misrepresented in parameterizations or under-resolved at low resolution, it may be worthwhile using these relations to guide and evaluate further convection parameterization development, including upcoming schemes developed using machine learning techniques (Dwyer and O'Gorman, 2017; Gentine et al., 2018)

560 Further, we question whether differences in the humidity dependencies are reflected in the spatial differences between NL and SFR-cent. The climate in SFR-cent is most comparable to NL, with extreme rain occurring primarily in the summer months, but it is characterized by slightly higher absolute humidity (<1 °C higher TD), and considerably lower relative humidity (2-4 °C higher DPD). It is clear that CPMs show much more realistic humidity dependencies (by all metrics discussed here) with much smaller spread as compared to the RCMs; they are also more realistic and show less spread in simulating the difference
565 in extreme rain statistics between France and The Netherlands.

All CPMs underestimate the difference in extremes between SFR-cent compared to NL. This is likely related to the finding that they underestimate the sensitivity to humidity when the relative humidity is low. The dry bias in climatology – with too low relative humidity in most CPMs – could also play a role here. It may well be that low relative humidity is still enhancing

570   rainfall extremes for the Netherlands, while the lower relative humidity in SFR starts to limit the simulation of rainfall extremes in the CPMs.

Finally, we note that several of the dependencies shown in this paper appear to be comparable in the two regions studied, the Netherlands (NL) and central southern France (SFR-cent). Scaling based on dew point temperature shows ~2CC behaviour for

575   a large dew point temperature range. Dependencies on relative humidity are also broadly consistent with a general increase in rainfall amount with decreasing relative humidity. The frequency of wet events is also similar between the regions, particularly for high levels of relative humidity.

5 Conclusions

We evaluated 5 CPM and 7 RCM simulations – all driven by reanalysis data – using hourly rainfall observations over The

580   Netherlands and in southern France. We investigated extreme rainfall statistics as well as their sensitivities to absolute and relative humidity.

We have shown that, on average for a model ensemble, CPMs have much better extreme hourly rainfall distributions in comparison to RCMs, in agreement with earlier results. Importantly, CPMs more faithfully reproduce dependencies of hourly

585   rainfall extremes on absolute humidity (as measured by dew point temperature) and relative humidity (as measured by dew point depression) than RCMs. RCMs show much greater spread, generally too low rainfall intensities, often (but not always) underestimating the dependency on dew point temperature and are generally (except for HadRM3) unable to reproduce the observed dependency of rainfall extremes on relative humidity.

590   The most striking result is the sensitivity of hourly rainfall to relative humidity, with all measures showing that lower relative humidity leads to more intense rainfall for the climate conditions in NL and the inland part of southern France (SFR-cent). This is particularly relevant in a climate change context as widespread decreases in relative humidity are expected in future summers. The RCMs (with exception of one) cannot reproduce the observed extreme rainfall increase with decreasing relative humidity and are therefore unreliable for extreme rainfall projections in a climate change setting. The CPMs qualitatively

595   reproduce the correct sensitivity to relative humidity, but still appear to underestimate the observed enhancement of rainfall extremes with lower relative humidity. They are therefore more reliable, but may still underestimate climate change effects on rainfall extremes.

In short, our results support the greater trustworthiness of CPM results as compared to RCMs in a climate change setting.

600   However, we also note deficiencies obtained in most CPMs: in general too low relative humidity, a negative bias in rain frequency for high relative humidity with likely weak forcing from the surface, and biases in differences for the most extreme

hourly rainfall between The Netherlands and southern France. By comparing models with the observations using the metrics describing absolute and relative humidity dependencies as proposed in this paper, we can systematically improve climate models and obtain better future predictions of rainfall extremes.

605 **References**

Aalbers, E. E., van Meijgaard, E., Lenderink, G., de Vries, H., and van den Hurk, B. J. J. M.: The 2018 west-central European drought projected in a warmer climate: how much drier can it get?, EGUsphere, 2022, 1–33, https://doi.org/10.5194/egusphere-2022-954, 2022.

Ali, H., Fowler, H. J., Pritchard, D., Lenderink, G., Blenkinsop, S., and Lewis, E.: Towards Quantifying the Uncertainty in
610 Estimating Observed Scaling Rates, Geophysical Research Letters, 49, https://doi.org/10.1029/2022GL099138, 2022.

Attema, J. J., Loriaux, J. M., and Lenderink, G.: Extreme precipitation response to climate perturbations in an atmospheric mesoscale model, Environmental Research Letters, 9, 014003, https://doi.org/10.1088/1748-9326/9/1/014003, 2014.

Ban, N., Rajczak, J., Schmidli, J., and Schär, C.: Analysis of Alpine precipitation extremes using generalized extreme value theory in convection-resolving climate simulations, Climate Dynamics, 55, 61–75, https://doi.org/10.1007/s00382-018-4339-
615 4, 2020.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., Belušić, D., Berthou, S., Brisson, E., Cardoso, R. M., Chan, S. C., Christensen, O. B., Fernández, J., Fita, L., Frisius, T., Gašparac, G., Giorgi, F., Goergen, K., Haugen, J. E., Hodnebrog, Ø., Kartsios, S., Katragkou, E., Kendon, E. J., Keuler, K., Lavin-Gullon, A., Lenderink, G., Leutwyler, D., Lorenz, T., Maraun, D., Mercogliano, P., Milovac, J., Panitz, H. J., Raffa, M.,
620 Remedio, A. R., Schär, C., Soares, P. M. M., Srnec, L., Steensen, B. M., Stocchi, P., Tölle, M. H., Truhetz, H., Vergara-Temprado, J., de Vries, H., Warrach-Sagi, K., Wulfmeyer, V., and Zander, M. J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation, Climate Dynamics, 57, 275–302, https://doi.org/10.1007/s00382-021-05708-w, 2021.

Bao, J., Sherwood, S. C., Alexander, L. V., and Evans, J. P.: Future increases in extreme precipitation exceed observed scaling
625 rates, Nature Clim Change, 7, 128–132, https://doi.org/10.1038/nclimate3201, 2017.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., and Fosser, G.: Pan-European climate at convection-permitting scale: a model intercomparison study, Climate Dynamics, 55, 35–59, https://doi.org/10.1007/s00382-018-4114-6, 2020.

Bui, A., Johnson, F., and Wasko, C.: The relationship of atmospheric air temperature and dew point temperature to extreme
630 rainfall, Environmental Research Letters, 14, https://doi.org/10.1088/1748-9326/ab2a26, 2019.

Byrne, M. P. and O'Gorman, P. A.: Trends in continental temperature and humidity directly linked to ocean warming, Proceedings of the National Academy of Sciences of the United States of America, 115, 4863–4868, https://doi.org/10.1073/pnas.1722312115, 2018.

Chan, S. C., Kendon, E. J., Fowler, H. J., Blenkinsop, S., Roberts, N. M., and Ferro, C. a. T.: The Value of High-Resolution
635 Met Office Regional Climate Models in the Simulation of Multihourly Precipitation Extremes, Journal of Climate, 27, 6155–6174, https://doi.org/10.1175/JCLI-D-13-00723.1, 2014.

Chan, S. C., Kendon, E. J., Roberts, N. M., Fowler, H. J., and Blenkinsop, S.: Downturn in scaling of UK extreme rainfall with temperature for future hottest days, Nature Geosci, 9, 24–28, https://doi.org/10.1038/ngeo2596, 2016.

640 Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., Ban, N., Bastin, S., Belda, M., Belusic, D., Caldas-Alvarez, A., Cardoso, R. M., Davolio, S., Dobler, A., Fernandez, J., Fita, L., Fumiere, Q., Giorgi, F., Goergen, K., Güttler, I., Halenka, T., Heinzeller, D., Hodnebrog, Ø., Jacob, D., Kartsios, S., Katragkou, E., Kendon, E., Khodayar, S., Kunstmann, H., Knist, S., Lavín-Gullón, A., Lind, P., Lorenz, T., Maraun, D., Marelle, L., van Meijgaard, E., Milovac, J., Myhre, G., Panitz, H.-J., Piazza, M., Raffa, M., Raub, T., Rockel, B., Schär, C., Sieck, K., Soares, P. M. M., Somot, S., Srnec, L., Stocchi, P., Tölle, M. H., Truhetz, H., Vautard, R., de Vries, H., and Warrach-Sagi, K.: A first-of-its-kind multi-model
645 convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean, Clim Dyn, 55, 3–34, https://doi.org/10.1007/s00382-018-4521-8, 2020.

Dai, A., Rasmussen, R. M., Liu, C., Ikeda, K., and Prein, A. F.: A new mechanism for warm-season precipitation response to global warming based on convection-permitting simulations, Clim Dyn, 55, 343–368, https://doi.org/10.1007/s00382-017-3787-6, 2020.

650 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., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of
655 the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Derbyshire, S. H., Beau, I., Bechtold, P., Grandpeix, J.-Y., Piriou, J.-M., Redelsperger, J.-L., and Soares, P. M. M.: Sensitivity of moist convection to environmental humidity, Quarterly Journal of the Royal Meteorological Society, 130, 3055–3079, https://doi.org/10.1256/qj.03.130, 2004.

Duffourg, F. and Ducrocq, V.: Assessment of the water supply to Mediterranean heavy precipitation: a method based on finely
660 designed water budgets, Atmospheric Science Letters, 14, 133–138, https://doi.org/10.1002/asl2.429, 2013.

Dwyer, J. G. and O'Gorman, P. A.: Changing duration and spatial extent of midlatitude precipitation extremes across different climates, Geophysical Research Letters, 44, 5863–5871, https://doi.org/10.1002/2017GL072855, 2017.

Fischer, E. M. and Knutti, R.: Observed heavy precipitation increase confirms theory and early models, Nature Clim Change, 6, 986–991, https://doi.org/10.1038/nclimate3110, 2016.

665 Fosser, G., Kendon, E. J., Stephenson, D., and Tucker, S.: Convection-Permitting Models Offer Promise of More Certain Extreme Rainfall Projections, Geophysical Research Letters, 47, 1–9, https://doi.org/10.1029/2020GL088151, 2020.

Fosser, G., Gaetani, M., Kendon, E. J., Adinolfi, M., Ban, N., Belušić, D., Caillaud, C., Careto, J. A. M., Coppola, E., Demory, M.-E., De Vries, H., Dobler, A., Feldmann, H., Goergen, K., Lenderink, G., Pichelli, E., Schär, C., Soares, P. M. M., Somot, S., and Tölle, M. H.: Convection-permitting climate models offer more certain extreme rainfall projections, npj Clim Atmos
670 Sci, 7, 51, https://doi.org/10.1038/s41612-024-00600-w, 2024.

Fowler, H. J., Lenderink, G., Prein, A. F., Westra, S., Allan, R. P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Do, H. X., Guerreiro, S., Haerter, J. O., Kendon, E. J., Lewis, E., Schaer, C., Sharma, A., Villarini, G., Wasko, C., and Zhang, X.: Anthropogenic intensification of short-duration rainfall extremes, Nat Rev Earth Environ, 2, 107–122, https://doi.org/10.1038/s43017-020-00128-6, 2021a.

675   Fowler, H. J., Ali, H., Allan, R. P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Cabi, N. S., Chan, S., Dale, M., Dunn, R. J. H., Ekström, M., Evans, J. P., Fosser, G., Golding, B., Guerreiro, S. B., Hegerl, G. C., Kahraman, A., Kendon, E. J., Lenderink, G., Lewis, E., Li, X., O'Gorman, P. A., Orr, H. G., Peat, K. L., Prein, A. F., Pritchard, D., Schär, C., Sharma, A., Stott, P. A., Villalobos-Herrera, R., Villarini, G., Wasko, C., Wehner, M. F., Westra, S., and Whitford, A.: Towards advancing scientific knowledge of climate change impacts on short-duration rainfall extremes, Phil. Trans. R. Soc. A., 379, 20190542, 680   https://doi.org/10.1098/rsta.2019.0542, 2021b.

Fumière, Q., Déqué, M., Nuissier, O., Somot, S., Alias, A., Caillaud, C., Laurantin, O., and Seity, Y.: Extreme rainfall in Mediterranean France during the fall: added value of the CNRM-AROME Convection-Permitting Regional Climate Model, Climate Dynamics, 55, 77–91, https://doi.org/10.1007/s00382-019-04898-8, 2020.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., and Yacalis, G.: Could Machine Learning Break the Convection 685   Parameterization Deadlock?, Geophys. Res. Lett., 45, 5742–5751, https://doi.org/10.1029/2018GL078202, 2018.

Haerter, J. O. and Berg, P.: Unexpected rise in extreme precipitation caused by a shift in rain type?, Nature Geoscience, 2, 372–373, https://doi.org/10.1038/ngeo523, 2009.

Haerter, J. O. and Schlemmer, L.: Intensified Cold Pool Dynamics Under Stronger Surface Heating, Geophysical Research Letters, 45, 6299–6310, https://doi.org/10.1029/2017GL076874, 2018.

690   Hohenegger, C. and Stevens, B.: The role of the permanent wilting point in controlling the spatial distribution of precipitation, 115, 5692–5697, https://doi.org/10.1073/pnas.1718842115, 2018.

Hohenegger, C., Brockhaus, P., Bretherton, C. S., and Schär, C.: The Soil Moisture–Precipitation Feedback in Simulations with Explicit and Parameterized Convection, Journal of Climate, 22, 5003–5020, https://doi.org/10.1175/2009JCLI2604.1, 2009.

695   Kendon, E., Short, C., Pope, J., Chan, S., Wilkinson, J., Tucker, S., Bett, P., Harris, G., and Murphy, J.: UKCP Local Projections Update: Technical Note, 2021a.

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J.: Realism of Rainfall in a Very High-Resolution Regional Climate Model, Journal of Climate, 25, 5791–5806, https://doi.org/10.1175/JCLI-D-11-00562.1, 2012.

Kendon, E. J., Roberts, N. M., Fosser, G., Martin, G. M., Lock, A. P., Murphy, J. M., Senior, C. A., and Tucker, S. O.: Greater 700   future U.K. winter precipitation increase in new convection-permitting scenarios, Journal of Climate, 33, 7303–7318, https://doi.org/10.1175/JCLI-D-20-0089.1, 2020.

Kendon, E. J., Prein, A. F., Senior, C. A., and Stirling, A.: Challenges and outlook for convection-permitting climate modelling, Phil. Trans. R. Soc. A., 379, 20190547, https://doi.org/10.1098/rsta.2019.0547, 2021b.

Kendon, E. J., Fischer, E. M., and Short, C. J.: Variability conceals emerging trend in 100yr projections of UK local hourly 705   rainfall extremes, Nat Commun, 14, 1133, https://doi.org/10.1038/s41467-023-36499-9, 2023.

Klein, C. and Taylor, C. M.: Dry soils can intensify mesoscale convective systems, Proc. Natl. Acad. Sci. U.S.A., 117, 21132–21137, https://doi.org/10.1073/pnas.2007998117, 2020.

Lenderink, G. and Attema, J.: A simple scaling approach to produce climate scenarios of local precipitation extremes for the Netherlands, Environmental Research Letters, 10, 085001, https://doi.org/10.1088/1748-9326/10/8/085001, 2015.

Lenderink, G. and van Meijgaard, E.: Increase in hourly precipitation extremes beyond expectations from temperature changes, Nature Geoscience, 1, 511–514, https://doi.org/10.1038/ngeo262, 2008.

Lenderink, G. and van Meijgaard, E.: Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes, Environmental Research Letters, 5, 025208, https://doi.org/10.1088/1748-9326/5/2/025208, 2010.

Lenderink, G., Siebesma, a. P., Cheinet, S., Irons, S., Jones, C. G., Marquet, P., Müller, F., Olmeda, D., Calvo, J., Sánchez, E., and Soares, P. M. M.: The diurnal cycle of shallow cumulus clouds over land: A single-column model intercomparison study, Quarterly Journal of the Royal Meteorological Society, 130, 3339–3364, https://doi.org/10.1256/qj.03.122, 2004.

Lenderink, G., Mok, H. Y., Lee, T. C., and van Oldenborgh, G. J.: Scaling and trends of hourly precipitation extremes in two different climate zones – Hong Kong and the Netherlands, Hydrology and Earth System Sciences, 15, 3033–3041, https://doi.org/10.5194/hess-15-3033-2011, 2011.

Lenderink, G., van den Hurk, B. J. J. M., Klein Tank, a M. G., van Oldenborgh, G. J., van Meijgaard, E., de Vries, H., and Beersma, J. J.: Preparing local climate change scenarios for the Netherlands using resampling of climate model output, Environmental Research Letters, 9, 115008, https://doi.org/10.1088/1748-9326/9/11/115008, 2014.

Lenderink, G., Barbero, R., Loriaux, J. M., and Fowler, H. J.: Super-Clausius–Clapeyron Scaling of Extreme Hourly Convective Precipitation and Its Relation to Large-Scale Atmospheric Conditions, Journal of Climate, 30, 6037–6052, https://doi.org/10.1175/JCLI-D-16-0808.1, 2017.

Lenderink, G., Barbero, R., Westra, S., and Fowler, H. J.: Reply to comments on "Temperature-extreme precipitation scaling: a two-way causality?," International Journal of Climatology, 8–10, https://doi.org/10.1002/joc.5799, 2018.

Lenderink, G., Belušić, D., Fowler, H. J., Kjellström, E., Lind, P., van Meijgaard, E., van Ulft, B., and de Vries, H.: Systematic increases in the thermodynamic response of hourly precipitation extremes in an idealized warming experiment with a convection-permitting climate model, Environmental Research Letters, 14, 074012, https://doi.org/10.1088/1748-9326/ab214a, 2019.

Lenderink, G., de Vries, H., Fowler, H. J., Barbero, R., van Ulft, B., and van Meijgaard, E.: Scaling and responses of extreme hourly precipitation in three climate experiments with a convection-permitting model, Phil. Trans. R. Soc. A., 379, 20190544, https://doi.org/10.1098/rsta.2019.0544, 2021.

Lochbihler, K., Lenderink, G., and Siebesma, A. P.: Cold Pool Dynamics Shape the Response of Extreme Rainfall Events to Climate Change, Journal of Advances in Modeling Earth Systems, 13, 1–16, https://doi.org/10.1029/2020MS002306, 2021.

Loriaux, J. M., Lenderink, G., De Roode, S. R., and Siebesma, a. P.: Understanding Convective Extreme Precipitation Scaling Using Observations and an Entraining Plume Model, Journal of the Atmospheric Sciences, 70, 3641–3655, https://doi.org/10.1175/JAS-D-12-0317.1, 2013.

Lucas-Picher, P., Argüeso, D., Brisson, E., Tramblay, Y., Berg, P., Lemonsu, A., Kotlarski, S., and Caillaud, C.: Convection-permitting modeling with regional climate models: Latest developments and next steps, WIREs Climate Change, 1–59, https://doi.org/10.1002/wcc.731, 2021.

Molnar, P., Fatichi, S., Gaál, L., Szolgay, J., and Burlando, P.: Storm type effects on super Clausius–Clapeyron scaling of intense rainstorm properties with air temperature, Hydrol. Earth Syst. Sci., 19, 1753–1766, https://doi.org/10.5194/hess-19-1753-2015, 2015.

Mulholland, J. P., Peters, J. M., and Morrison, H.: How Does LCL Height Influence Deep Convective Updraft Width?, Geophysical Research Letters, 48, e2021GL093316, https://doi.org/10.1029/2021GL093316, 2021.

Müller, S. K., Caillaud, C., Chan, S., De Vries, H., Bastin, S., Berthou, S., Brisson, E., Demory, M.-E., Feldmann, H., Goergen, K., Kartsios, S., Lind, P., Keuler, K., Pichelli, E., Raffa, M., Tölle, M. H., and Warrach-Sagi, K.: Evaluation of Alpine-
750 Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms, Clim Dyn, 61, 939–957, https://doi.org/10.1007/s00382-022-06555-z, 2023.

Neelin, J. D., Martinez-Villalobos, C., Stechmann, S. N., Ahmed, F., Chen, G., Norris, J. M., Kuo, Y.-H., and Lenderink, G.: Precipitation Extremes and Water Vapor: Relationships in Current Climate and Implications for Climate Change, Curr Clim Change Rep, 8, 17–33, https://doi.org/10.1007/s40641-021-00177-z, 2022.

755 Nie, J., Sobel, A. H., Shaevitz, D. A., and Wang, S.: Dynamic amplification of extreme precipitation sensitivity, Proc Natl Acad Sci USA, 115, 9467–9472, https://doi.org/10.1073/pnas.1800357115, 2018.

Pall, P., Allen, M. R., and Stone, D. A.: Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO_2 warming, Climate Dynamics, 28, 351–363, https://doi.org/10.1007/s00382-006-0180-2, 2006.

Pichelli, E., Coppola, E., Sobolowski, S., Ban, N., Giorgi, F., Stocchi, P., Alias, A., Belušić, D., Berthou, S., Caillaud, C.,
760 Cardoso, R. M., Chan, S., Christensen, O. B., Dobler, A., de Vries, H., Goergen, K., Kendon, E. J., Keuler, K., Lenderink, G., Lorenz, T., Mishra, A. N., Panitz, H.-J., Schär, C., Soares, P. M. M., Truhetz, H., and Vergara-Temprado, J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation, Climate Dynamics, 56, 3581–3602, https://doi.org/10.1007/s00382-021-05657-4, 2021.

Prein, A. F. and Heymsfield, A. J.: Increased melting level height impacts surface precipitation phase and intensity, Nat. Clim.
765 Chang., 10, 771–776, https://doi.org/10.1038/s41558-020-0825-x, 2020.

Prein, A. F., Rasmussen, R., Castro, C. L., Dai, A., and Minder, J.: Special issue: Advances in convection-permitting climate modeling, Climate Dynamics, 55, 1–2, https://doi.org/10.1007/s00382-020-05240-3, 2020.

Prein, A. F., Rasmussen, R. M., Wang, D., and Giangrande, S. E.: Sensitivity of organized convective storms to model grid spacing in current and future climates, Philosophical Transactions of the Royal Society A: Mathematical, Physical and
770 Engineering Sciences, 379, 20190546, https://doi.org/10.1098/rsta.2019.0546, 2021.

Purr, C., Brisson, E., and Ahrens, B.: Convective shower characteristics simulated with the convection-permitting climate model COSMO-CLM, Atmosphere, 10, https://doi.org/10.3390/ATMOS10120810, 2019.

Purr, C., Brisson, E., and Ahrens, B.: Convective rain cell characteristics and scaling in climate projections for Germany, International Journal of Climatology, 41, 3174–3185, https://doi.org/10.1002/joc.7012, 2021.

775 Rasmussen, K. L., Prein, A. F., Rasmussen, R. M., Ikeda, K., and Liu, C.: Changes in the convective population and thermodynamic environments in convection-permitting regional climate simulations over the United States, Clim Dyn, 55, 383–408, https://doi.org/10.1007/s00382-017-4000-7, 2020.

de Rooy, W. C. and Pier Siebesma, A.: Analytical expressions for entrainment and detrainment in cumulus convection, Quarterly Journal of the Royal Meteorological Society, n/a-n/a, https://doi.org/10.1002/qj.640, 2010.

780 Schär, C., Ban, N., Fischer, E. M., Rajczak, J., Schmidli, J., Frei, C., Giorgi, F., Karl, T. R., Kendon, E. J., Tank, A. M. G. K., O'Gorman, P. A., Sillmann, J., Zhang, X., and Zwiers, F. W.: Percentile indices for assessing changes in heavy precipitation events, Climatic Change, 137, 201–216, https://doi.org/10.1007/s10584-016-1669-2, 2016.

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Di Girolamo, S., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T. C., Sprenger, M., Ubbiali, S., and Wernli, H.: Kilometer-Scale Climate Models: Prospects and Challenges, Bulletin of the American Meteorological Society, 101, E567–E587, https://doi.org/10.1175/BAMS-D-18-0167.1, 2020.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S. J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., Wedi, N., and Zhou, L.: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains, Progress in Earth and Planetary Science, 6, https://doi.org/10.1186/s40645-019-0304-z, 2019.

Trenberth, K.: Changes in precipitation with climate change, Climate Research, 47, 123–138, https://doi.org/10.3354/cr00953, 2011.

Williams, A. I. L. and O'Gorman, P. A.: Summer-Winter Contrast in the Response of Precipitation Extremes to Climate Change Over Northern Hemisphere Land, Geophysical Research Letters, 49, e2021GL096531, https://doi.org/10.1029/2021GL096531, 2022.

Yano, J.-I., Bister, M., Fuchs, Ž., Gerard, L., Phillips, V. T. J., Barkidija, S., and Piriou, J.-M.: Phenomenology of convection-parameterization closure, Atmos. Chem. Phys., 13, 4111–4131, https://doi.org/10.5194/acp-13-4111-2013, 2013.

Zhang, X., Zwiers, F. W., Li, G., Wan, H., and Cannon, A. J.: Complexity in estimating past and future extreme short-duration rainfall, Nature Geosci, 10, 255–259, https://doi.org/10.1038/ngeo2911, 2017.

Zhou, W., Leung, L. R., and Lu, J.: The Role of Interactive Soil Moisture in Land Drying Under Anthropogenic Warming, Geophysical Research Letters, 50, e2023GL105308, https://doi.org/10.1029/2023GL105308, 2023.

Data availability: https://doi.org/10.5281/zenodo.11066905. For France only the postprocessed statistics used in this paper are available though this link. For the raw station data please contact Meteo France.

820   *Code availability*: https://github.com/mister-superCC/CCscaling-Evaluation

825