Meteorologische Zeitschrift, Vol. **25**, No. 6, 753–766 (published online July 6, 2016)

MiKlip

Validation of Convective Parameters in MPI-ESM Decadal Hindcasts (1971–2012) against ERA-Interim Reanalyses

Georg Pistotnik[1,3*], Pieter Groenemeijer[1] and Robert Sausen[2]

[1]European Severe Storms Laboratory, Wessling, Germany
[2]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[3]Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria

(Manuscript received September 5, 2014; in revised form August 13, 2015; accepted September 9, 2015)

Abstract

Decadal forecasts of the Max Planck Institute – Earth System Model (MPI-ESM) are validated against ERA-Interim reanalysis data for the period 1979 to 2012 over Europe with respect to (a) the medians of 500 hPa temperature, 925 hPa temperature and 925 hPa mixing ratio, and (b) the 90th percentiles of parameters that are relevant for convective storms. These parameters are the vertical temperature gradient between 850 and 500 hPa, the low-level mixing ratio (25 hPa above the local model topography), and Convective Available Potential Energy (CAPE). The MPI-ESM evaluation reveals very little bias for the median 925 hPa temperature but a pronounced negative bias for the median 500 hPa temperature. Besides, a positive bias of the median 925 hPa moisture is found. The 90th percentiles of the convective parameters exhibit positive biases of the vertical temperature gradient and the low-level mixing ratio, which results in an overestimation of CAPE. It is hypothesized that MPI-ESM's convective parameterization scheme is reluctant in dissipating CAPE across Europe, often leaving the atmosphere in a state that appears too favorable for convective storms. Biases for convective parameters are slightly larger for an enhanced resolution of MPI-ESM. A categorization with respect to forecast years indicates that these biases mostly result from a systematically deviant model climatology. However, a considerable part of the 500 hPa temperature bias is also introduced by the initial conditions. ERA-Interim shows an increase of low-level moisture over most parts of Europe in the past three decades, in particular after 1990. Steep vertical temperature gradients became less frequent in northwestern Europe and more frequent in southeastern Europe. The 90th percentile of CAPE exhibits little evolution in northwestern Europe, where these two changes largely compensate each other, whereas it increases in southeastern Europe, where they add up. MPI-ESM seems able to reproduce some of these tendencies, but not regional structures.

Keywords: convective storms, ERA-Interim reanalyses, decadal climate predictions

1 Introduction

Convective storms can produce severe weather events such as hail, strong wind gusts, tornadoes, or flash floods. These phenomena are estimated to cause an average material damage of 5–8 billion Euro per year in Europe (Sander, 2011), comparable to that inflicted by large-scale winter storms (MunichRe, 2006). It is therefore of high interest to predict and monitor the occurrence frequency of these extremes on a multi-annual and decadal basis.

Due to the rare occurrence of severe thunderstorms at any given location, observational data from routine weather station networks give an incomplete picture of their true frequency. An alternative approach is the collection of event-based information on severe weather instances in databases like Storm Data (Schaefer and Edwards, 1999) or the European Severe Weather Database (Dotzek et al., 2009). However, the varying reporting rate across different times and countries still imposes high challenges on the creation of transnational severe storm climatologies. Artifacts by changes in reporting procedures obscure meteorological signals and trends, as shown, e.g., for tornadoes in the United States (Verbout et al., 2006) and particularly in Europe (Dotzek, 2003; Groenemeijer and Kühne, 2014). For example, an improving report procedure could give the false impression that the frequency of severe weather would increase.

A way to minimize the presence of non-meteorological artifacts is to identify covariates: meteorological quantities that are observed, analyzed or predicted consistently in space and time, and that can be used to estimate the probability that the event of interest occurs (Brooks et al., 2003). Such covariates may be parameters computed from radiosonde measurements (Kunz et al., 2009; Mohr and Kunz, 2013), or objectively classified weather patterns from model data (e.g., Kapsch et al., 2012). More often, the covariates are grid-scale variables that serve as proxies for the occurrence of

*Corresponding author: Georg Pistotnik, Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria, Hohe Warte 38, 1190 Vienna, Austria, e-mail: georg.pistotnik@zamg.ac.at

DOI 10.1127/metz/2016/0649
Gebrüder Borntraeger Science Publishers, Stuttgart, www.borntraeger-cramer.com

subgrid-scale phenomena. Naturally, the success of this approach depends on whether such proxies can be found and whether a numerical model can simulate them accurately.

Various authors found high Convective Available Potential Energy (CAPE) and strong 0–6 km vertical wind shear (deep-layer shear) to characterize environments conducive to severe convective storms (e.g. Rasmussen and Blanchard, 1998; Craven and Brooks, 2004). Brooks et al. (2003) first computed these two proxies from the global NCAR/NCEP reanalysis dataset (Kalnay et al., 1996) to assess the global climatology of conditions favorable to severe storms. Brooks and Dotzek (2008) performed a temporal analysis of the frequency of such conditions across the central and eastern USA and identified a weak decrease till the early 1970s and a weak increase thereafter. Marsh et al. (2007) compared a similar severe storm proxy for a 20-year dataset of a CCSM3 climate hindcasts (Collins et al., 2006) and NCAR/NCEP reanalyses across the contiguous USA and found a good qualitative agreement with the climatology derived by Brooks et al. (2003).

Several authors also applied proxies based on CAPE and deep-layer shear to climate models (e.g. Trapp et al. 2007; Trapp et al. 2009; Van Klooster and Robber, 2009). Their area of study was the contiguous USA. These studies largely agree on an expected increase in the frequency of favorable severe thunderstorm conditions until the end of the 21st century, which is caused by a predicted moistening of the lower troposphere and resulting higher CAPE, whose effect overcompensates a predicted reduction of deep-layer shear. Diffenbaugh et al. (2013) recently demonstrated that high-CAPE situations would not be affected by this decrease of deep-layer shear. A more frequent combined occurrence of high CAPE and strong deep-layer shear, despite a general decrease of the latter, was also brought forward by Del Genio et al. (2007).

Similar studies for Europe are comparably sparse. In their worldwide comparison of the frequency of severe storm conditions based on NCAR/NCEP reanalyses, Brooks et al. (2003) found the highest severe storm risk in Europe for Spain and for the western Balkan states. Investigations with the CCSM3 model and NCAR/NCEP reanalyses exhibited weak changes in the overall distributions of either CAPE or deep-layer shear over much of Europe from the 20th to the 21st century (Marsh et al., 2009). A study by Sander (2011) with the COSMO-CLM regional climate model (Böhm et al., 2006) showed an expected increase of mean CAPE from the late 20th to the late 21st century. Eccel et al. (2012) considered trends of convective parameters at selected grid points in the ERA-40 reanalysis.

In the present study, we are concerned with climate forecasts on a shorter, decadal timescale. Decadal predictions are still a relatively new field of climate forecasts. They are challenging because inter-decadal variations are mainly driven by high-frequency variability whose more or less stochastic nature often defies predictability (Latif, 2011; Liu, 2012). It obscures low-frequency variability (e.g., multi-decadal trends) or effects resulting from external forcing (e.g., an anthropogenic influence on greenhouse gas concentrations) from which predictability of the climate system mainly arises (Meehl et al., 2009; Latif and Keenlyside, 2011). Nonetheless, some aspects of observed decadal variability can already be reproduced by state-of-the-art climate models, which raises hopes that decadal variations may be partly predictable (Smith et al., 2007; Keenlyside and Ba, 2010; Pohlmann et al., 2013).

This study is a first step in developing proxies for the occurrence of severe thunderstorm hazards using the decadal predictions based on the Earth System Model of the Max Planck Institute in Hamburg (MPI-ESM; Giorgetta et al., 2013; Stevens et al., 2013). The aim of the present study is to perform a validation of quantities upon which such a proxy will depend, i.e. temperature and moisture. These quantities and the derived proxies are validated against a reanalysis dataset that was developed with the intention to represent the atmosphere as consistently as possible with all available observations using state-of-the-art data assimilation techniques. Similar comparisons of the model climate of a GCM with reanalysis datasets have recently been undertaken by Brands et al. (2013), Devis et al. (2014) and Jury et al. (2015). We favor the validation with respect to reanalysis data over one that uses actual observations, since the comparison of two gridded data sets is more straightforward than the comparison of point measurements to grid-box averages (Sander, 2011; Graf et al., 2011). This is true in particular for horizontally strongly varying fields and for lower model resolutions.

2 Data and methods

2.1 Data

Our study is based on decadal hindcasts provided by MPI-ESM (Pohlmann et al., 2013). The "baseline 1" version of MPI-ESM used for this study is initialized with temperature and salinity anomalies from the ORAS4 ocean reanalysis system (Balmaseda et al., 2012) and with atmospheric anomalies from ERA-40 (Uppala et al., 2005) and ERA-Interim reanalyses (Dee et al., 2011). The initialization of MPI-ESM results in improved forecast skill in the Northern Atlantic and European region compared to uninitialized reference runs (Müller et al., 2012).

Ten "low resolution" (LR) runs and five "mixed resolution" (MR) runs of MPI-ESM initialized each year were analyzed for this study. The MR and LR members have different resolutions of the ocean ($0.4\,° \times 0.4\,°$ L40 and $1.5\,° \times 1.5\,°$ L40, respectively) and the atmosphere (T63L95 and T63L47, respectively). The finer resolution improves the forecast of the quasi-biennial oscillation and of the surface air temperature over the tropical Pacific Ocean (Pohlmann et al., 2013). Ensemble

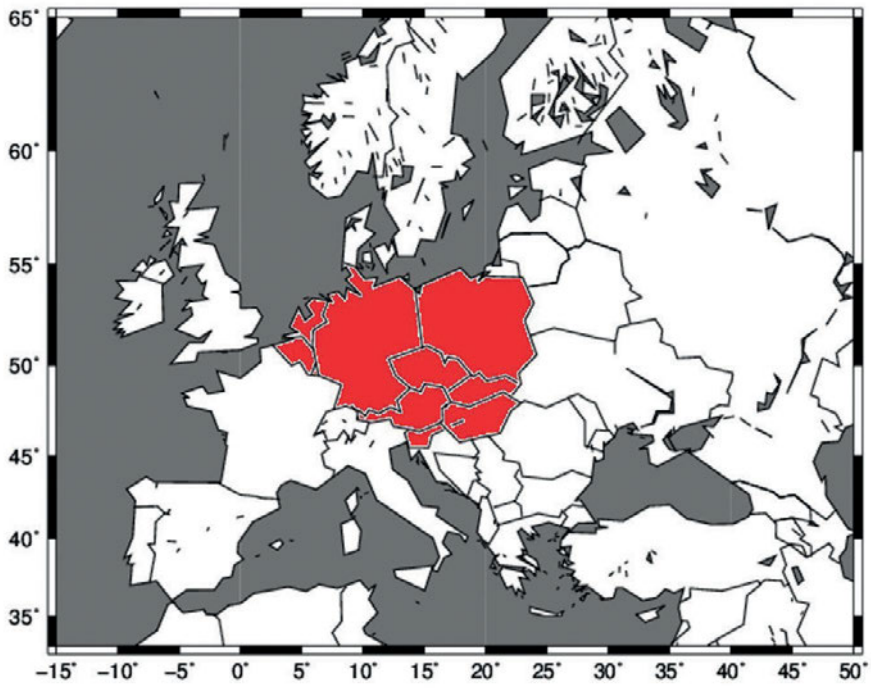

Figure 1: Subdomain of ten central European countries (Belgium, the Netherlands, Luxemburg, Germany, Poland, the Czech Republic, Slovakia, Hungary, Austria and Slovenia) to which the areal averages relate which are illustrated in Figs. 3 and 5 and described in the text.

spread in the MPI-ESM decadal hindcasts is created by initializing the model on consecutive days around 1 January each year. With ten LR and five MR runs for each starting year and a forecast horizon of ten years, a total of 150 realizations are available for each target year. We evaluate the years from 1971 to 2012.

The 6-hourly output of MPI-ESM is validated against ERA-Interim reanalyses (Dee et al., 2011), which are available since 1979 at a temporal resolution of 6 hours, a spatial resolution of 0.75 ° × 0.75 ° and a vertical resolution of 37 pressure levels. The data assimilation procedure of ERA-Interim ensures that the final reanalysis fields closely reproduce the observational data as long as these fulfill the prescribed error statistics. The resulting root mean square distance between the temperature at levels around 500 hPa measured by radiosondes and reanalyzed in ERA-Interim was shown to be generally below 1.5 K (Dee et al., 2011). We are not aware of any direct comparison between ERA-Interim and radiosonde data with respect to convective parameters and take as a working hypothesis that ERA-Interim is the best possible approximation of the truth. We will therefore refer to any deviations of MPI-ESM from ERA-Interim as biases.

2.2 Methods

Our comparison of MPI-ESM runs with ERA-Interim is divided into two parts. First, we investigate three basic quantities: the 500 hPa temperature (T500), the 925 hPa temperature (T925), and the 925 hPa mixing ratio (MR925). For these parameters, we compute annual medians (p50).

Second, we compare quantities that directly influence the probability of convective storms. The most common such parameter is the Convective Available Potential Energy (CAPE):

$$CAPE = g \int_{LFC}^{EL} \frac{T_v - \overline{T_v}}{\overline{T_v}} dz$$

LFC is the level of free convection (where a theoretically lifted air parcel first becomes positively buoyant), *EL* the equilibrium level (where it becomes cooler than its surroundings again), and g the gravitational acceleration. Two conditions have to be met so that the virtual temperature T_v of a rising parcel becomes higher than the environmental virtual temperature $\overline{T_v}$ and positive CAPE results. First, the vertical gradient of $\overline{T_v}$ needs to be at least steeper than the moist-adiabatic lapse rate. Second, sufficient moisture at low levels needs to be available to allow condensation and latent heat release in the rising parcel. These prerequisites, or "ingredients" can evolve mostly independently from each other and be brought together by synoptic processes (Doswell et al., 1996; Markowski and Richardson, 2010). We compute the "low-level mixing ratio" (LLMR) by the mixing ratio value at 25 hPa below the local surface pressure and the "lapse rate" (LARA) by the vertical temperature gradient between 850 and 500 hPa.

Since thunderstorms typically occur when these parameters have values well above their respective climatological median values (Brooks et al., 2003; Marsh et al., 2009; Brooks, 2013), we study the 90th percentiles of their annual distributions (p90). For a smaller

area of ten central European countries (Fig. 1), we also analyze MPI-ESM LR and MR hindcasts separately to investigate the effect of model resolution, and we evaluate the 1st, 2nd to 5th, and 6th to 10th forecast years separately to assess a potential benefit from the initialization of the model with observational data, as suggested by Goddard et al. (2013). Across this sub-domain, the coverage of severe weather observations in ESWD may be sufficiently high since 2006 to eventually enable a validation of severe storm frequency predictions on decadal timescales as follow-up work to the present study.

CAPE is computed from vertical temperature and humidity profiles interpolated to 1 hPa resolution, and a computed virtual temperature curve of lifted parcels using a numerical integration with steps of 1 hPa. The CAPE version used in our analysis is the maximum value of CAPE for a number of parcels that originate from different source layers ("most unstable CAPE"). The lowest layer is set at 25 hPa below the model's surface pressure, i.e. approximately 250 m above the earth's surface. This shift ensures that very shallow near-surface layers, which are often not representative for a storm's inflow (Craven et al., 2002), are discarded. Additional source layers are located every 25 hPa up to 700 hPa in order to consider layers of potentially buoyant air at higher altitudes, which can still act as a source for elevated convection. Besides our study, the same set of convection parameters was also used to find a proxy for thunderstorm occurrence from EUCLID lightning data and ERA-Interim reanalyses (Westermayer et al., 2016).

3 Results

3.1 Comparison of temperature and moisture

For p50_T500 (Fig. 2.a and 2.b), MPI-ESM exhibits a pronounced negative bias on the order of 1.5 to 2 K. Apart from that, it is able to represent the averaged structures superimposed onto the zonal gradient, namely a weak thermal trough from southeastern Europe into western Russia and an opposed weak thermal ridge over western Europe, though the axis of the latter is shifted a few hundred kilometers to the east. p50_T925 (Fig. 2.c and 2.d) matches fairly well, since only minor biases, with a magnitude below 1 K, are confined to some regions over southeastern Europe (positive) and over France and the Iberian Peninsula (negative). Smaller-scale structures like weak thermal ridges over the North Sea and the Baltic Sea and a thermal trough in-between over Norway and Sweden can be found both in ERA-Interim and MPI-ESM. For p50_MR925 (Fig. 2.e and 2.f), MPI-ESM exhibits a positive bias around 0.2 g/kg over much of Europe. A land-sea-contrast, with higher MR925 values over land, is more pronounced in ERA-Interim than in MPI-ESM in the Mediterranean region.

Averaged over the central European sub-domain and categorized with respect to model resolution and forecast time (Fig. 3), MR runs perform slightly worse than LR runs for p50_T500 (MR bias: −1.92 K; LR bias: −1.81 K) and slightly better for p50_T925 (MR bias: +0.07 K; LR bias: +0.13 K). There is hardly any difference for p50_MR925 (MR bias: +0.37 g/kg; LR bias: +0.36 g/kg).

Interestingly, MPI-ESM does not exhibit larger biases with longer forecast times for p50_T500 (year 1 bias: −1.86 K; year 2–5 bias: −1.86 K; year 6–10 bias: −1.81 K), which indicates that the model initialization has no positive impact in this case. In contrast, the bias initially rises with increasing lead times for p50_T925 (year 1 bias: +0.06 K; year 2–5 bias: +0.17 K; year 6–10 bias: +0.12 K) and for p50_MR925 (year 1 bias: +0.31 g/kg, year 2–5 bias: +0.37 g/kg; year 6–10 bias: +0.36 g/kg). This suggests that the model initialization adds skill to the 925 hPa forecast fields, since the rising bias indicates a drift from the model's initial state towards its preferred internal climatology (Goddard et al., 2013). The magnitude of the biases is largest in the 1980s for p50_T500 and in the 1990s for p50_T925 and p50_MR925, while it is smallest in the 2000s for each of these quantities (Fig. 3.b, 3.d and 3.f, respectively).

3.2 Comparison of convective parameters

Fig. 4 compares p90_LARA, p90_LLMR and p90_CAPE computed from ERA-Interim reanalyses and MPI-ESM hindcasts. A positive p90_LARA bias of MPI-ESM amounts to 0.2 K/km across most of Europe and is qualitatively consistent with the negative bias that was found for p50_T500. For p90_LLMR, MPI-ESM matches well with ERA-Interim over much of the continent but is affected by a negative bias over the Mediterranean region and the Black Sea. Interestingly, this is contrary to the positive bias for p50_MR925 identified in the previous sub-section; though LLMR and MR925 are not exactly the same quantities, the contrasting behavior of these different percentiles suggests that the annual cycle of near-surface moisture is smaller in MPI-ESM than in ERA-Interim. The combination of a positive p90_LARA bias with hardly any p90_LLMR bias results in a positive bias of p90_CAPE across most areas. Relative to the CAPE magnitude, its differences are generally larger for higher geographical latitudes, which experience a less frequent occurrence of high-CAPE situations. Along some coastal areas of the Mediterranean and Black Sea as well as in western Norway, however, the p90_CAPE bias vanishes or is negative.

Fig. 5 categorizes the spatially averaged MPI-ESM biases of convective parameters over the central European sub-domain with respect to model resolution and different lead times. The MR runs reveal higher (i.e. "more convective") values, and therefore larger positive biases, than the LR runs for all three convective parameters. The MR and LR biases amount to +0.20 and +0.17 K/km for p90_LARA, to +0.26 and +0.19 g/kg for p90_LLMR, and to +283 and +250 J/kg for p90_CAPE, respectively. This inferior performance of the MR members contrasts with the commonly expected increase of

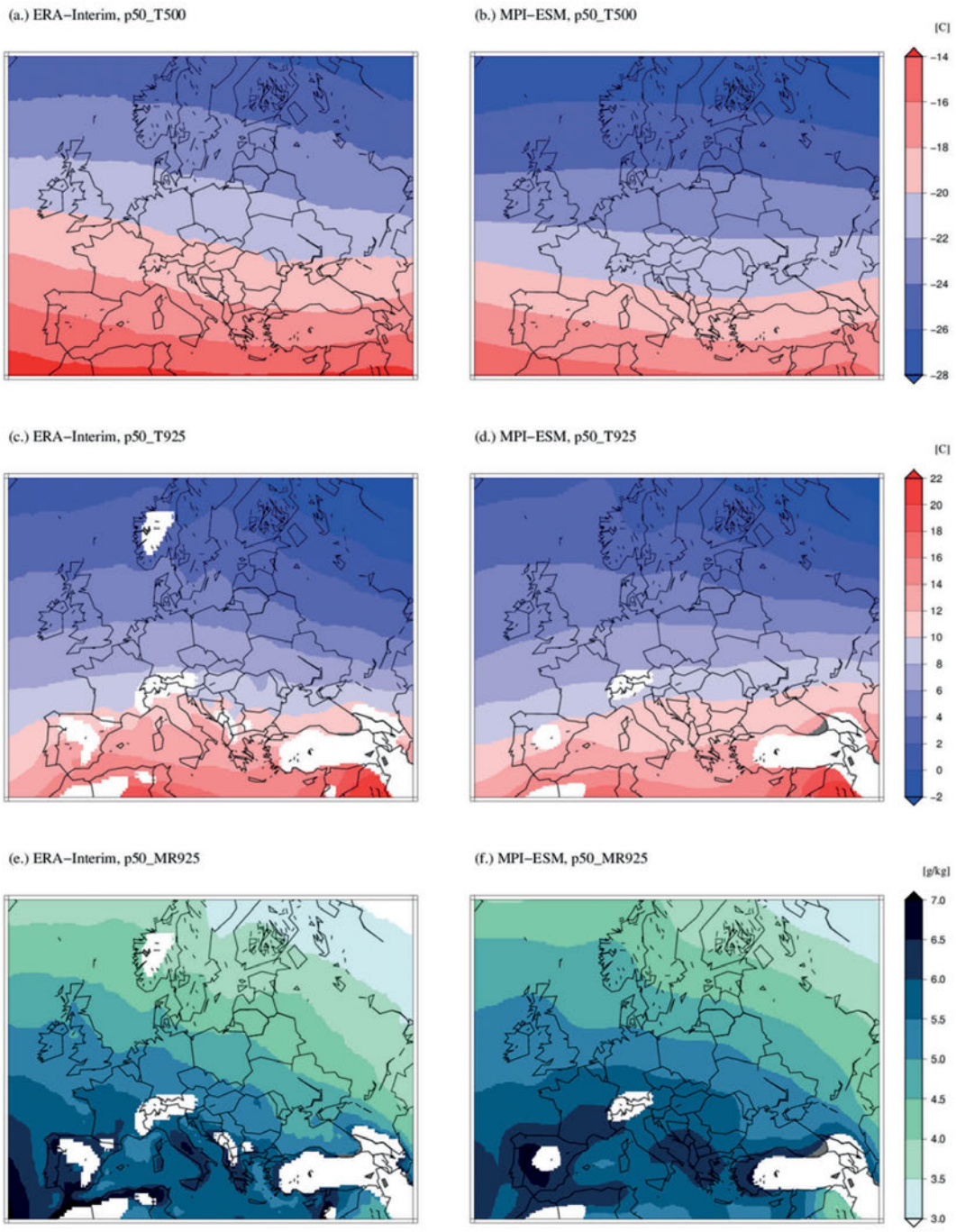

Figure 2: Medians for the period 1979–2012: (a.) and (b.) 500 hPa temperature [°C] in ERA-Interim and MPI-ESM. (c.) and (d.) 925 hPa temperature [°C] in ERA-Interim and MPI-ESM. (e.) and (f.) 925 hPa mixing ratio [g/kg] in ERA-Interim and MPI-ESM. Areas where the 925 hPa level is below the respective model topography are left white.

forecast skill by a higher resolution of MPI-ESM or similar climate models (e.g. Pohlmann et al., 2013, and references therein).

With respect to forecast time, biases of p90_LARA are almost constant (year 1 bias: +0.19 K/km; year 2–5 bias: +0.18 K/km; year 6–10 bias: +0.18 K/km), while an increase with longer lead times can be noted for p90_LLMR (year 1 bias: +0.16 g/kg; year 2–5 bias: +0.22 g/kg; year 6–10 bias: +0.23 g/kg) and consequently also for p90_CAPE (year 1 bias: +253 J/kg; year 2–5 bias: +265 J/kg; year 6–10 bias: +261 J/kg). The initialization of the model apparently adds forecast skill for p90_LLMR, while the pronounced positive bias of p90_LARA seems to be imported from the initialization fields (similar to the too low values of p50_T500). Categorized by decades, the bias is largest in the 1980s for p90_LARA and in the 1990s for p90_LLMR and p90_CAPE, while again it is smallest in the 2000s for all three convective parameters (Fig. 5.b, 5.d and 5.f, respectively).

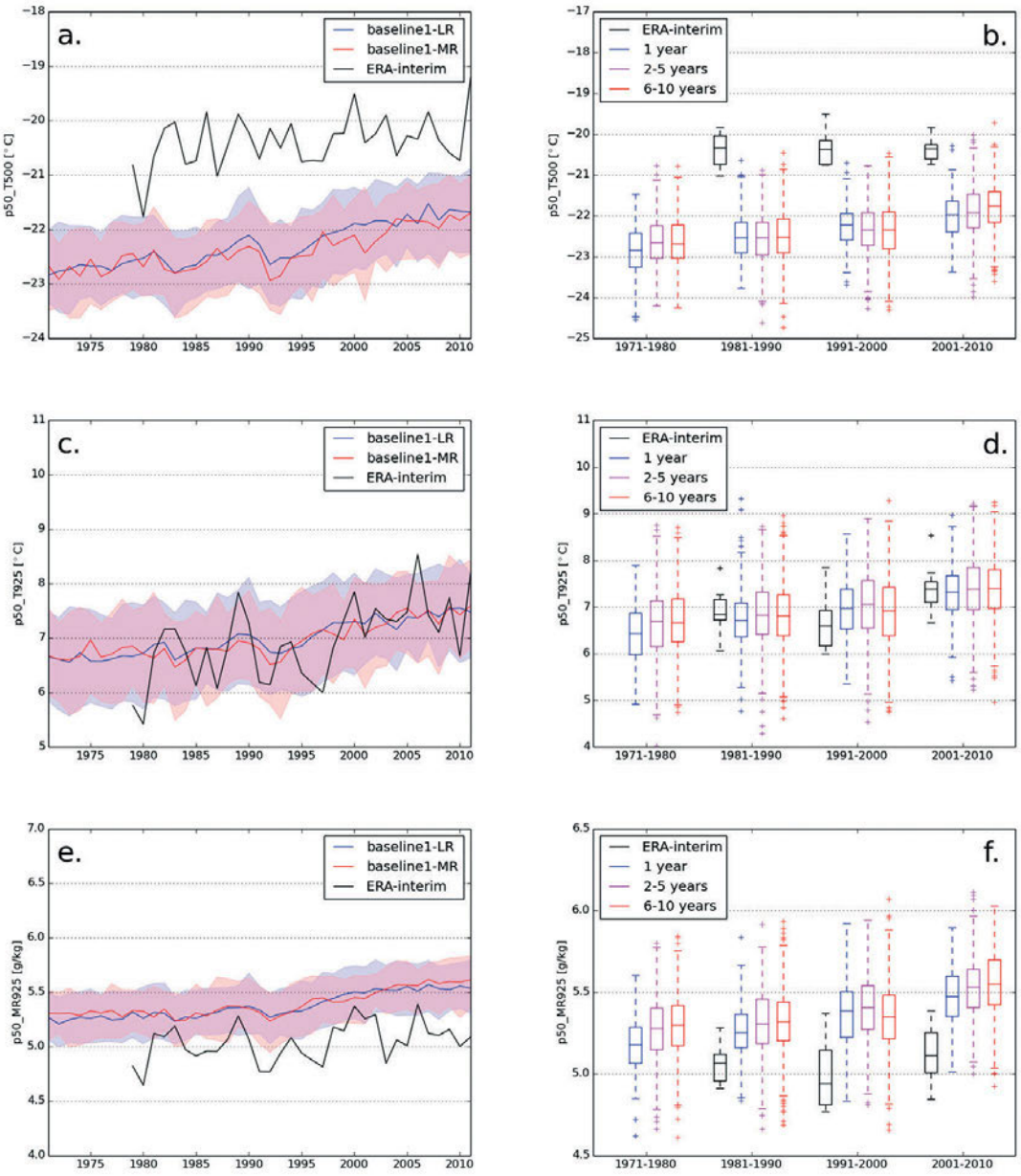

Figure 3: Annual medians averaged over the central European sub-domain from Fig. 1. (a.) Annual medians of 500 hPa temperature [°C] in ERA-Interim (black), MPI-ESM LR (blue for ensemble median and shaded blue for ensemble spread) and MPI-ESM MR (red for ensemble median and shaded red for ensemble spread). (b.) Decadal box-and-whisker plot of annual medians of 500 hPa temperature [°C] in ERA-Interim (black) and MPI-ESM for the forecast years 1 (blue), 2–5 (pink) and 6–10 (red); the boxes are delimited by the respective 25th, 50th and 75th percentiles, 10th and 90th percentiles are indicated by the whiskers, outliers are shown as crosses. (c.) as (a.), but for 925 hPa temperature [°C]. (d.) as (b.), but for 925 hPa temperature [°C]. (e.) as (a.), but for 925 hPa mixing ratio [g/kg]. (f.) as (b.), but for 925 hPa mixing ratio [g/kg].

3.3 Interdecadal changes of convective parameters

Changes of the 90th percentiles of the convective parameters from the period 1981–1990 to the period 1991–2000 are illustrated in Fig. 6. According to ERA-Interim, a rise of p90_CAPE by up to 200 J/kg occurred over the eastern Mediterranean Sea, Turkey, the Aegean Sea, Greece, Bulgaria, parts of Romania and the Black Sea (Fig. 6.e), where both p90_LARA (Fig. 6.a) and p90_LLMR (Fig. 6.c) increased. In contrast, the western and central Mediterranean region as well as western Russia experienced a negative interdecadal change of p90_CAPE by up to 300 J/kg, which mostly resulted from a decrease of p90_LLMR under neutral or even positive trends of p90_LARA. In southern France, negative trends of p90_LLMR and p90_LARA also resulted in a negative trend of p90_CAPE. Otherwise, analyzed changes were of minor importance.

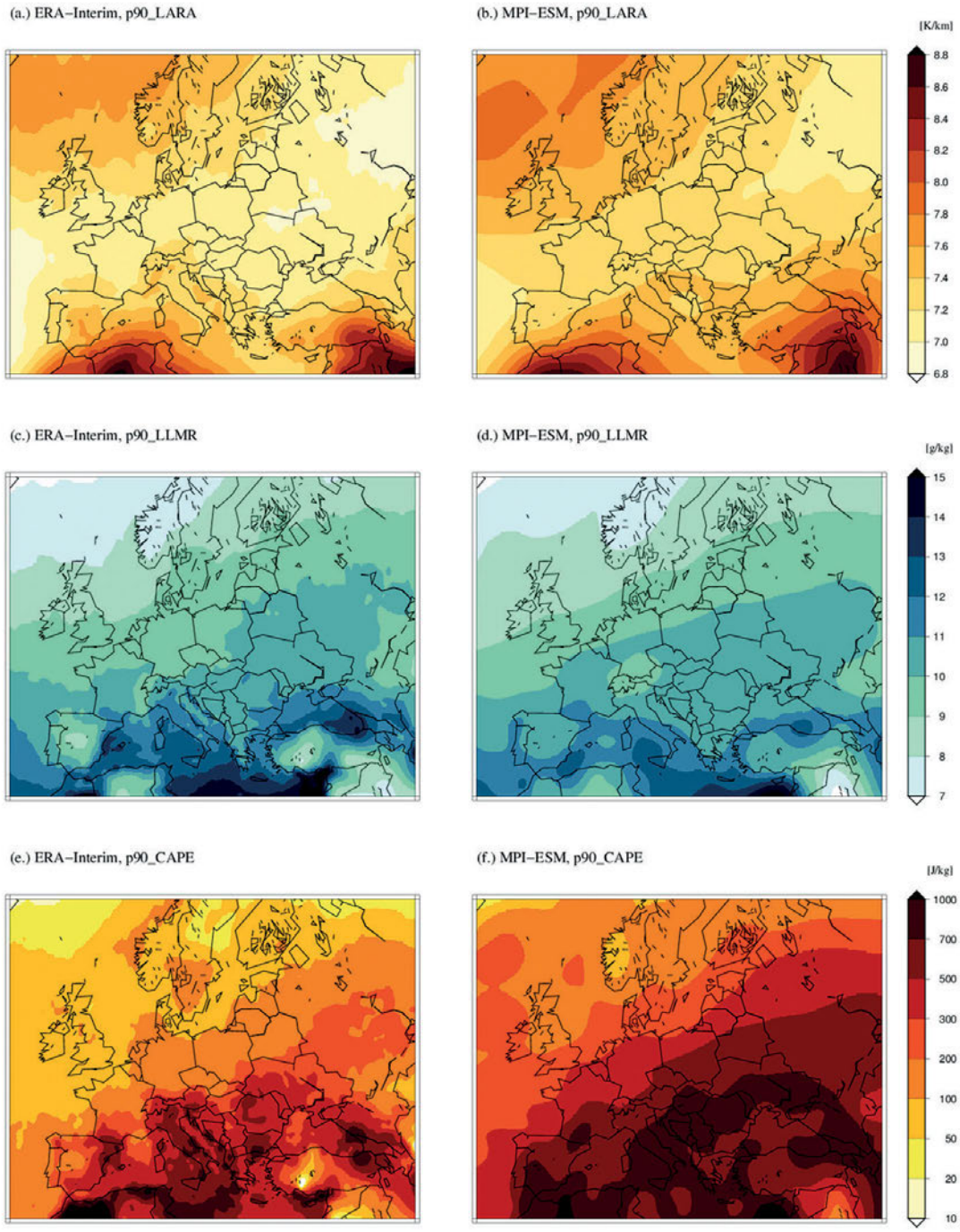

Figure 4: 90th percentiles for the period 1979–2012: (a.) and (b.) 850 to 500 hPa lapse rates [K/km] in ERA-Interim and MPI-ESM. (c.) and (d.) mixing ratio [g/kg] 25 hPa below the surface pressure in ERA-Interim and MPI-ESM. (e.) and (f.) CAPE [J/kg] in ERA-Interim and MPI-ESM.

In the MPI-ESM simulations, interdecadal changes of p90_LARA were neutral to slightly negative (Fig. 6.b) and those of p90_LLMR slightly positive (Fig. 6.d). As a result, slightly higher p90_CAPE was simulated in the 1990s over the Iberian Peninsula as well as over Belarus, the Ukraine and Russia, while it was similar to the 1980s elsewhere (Fig. 6.f).

The interdecadal changes from the period 1991–2000 to the period 2001–2010 show a remarkably different behavior (Fig. 7). According to ERA-Interim, p90_CAPE increased over almost the entire continent (Fig. 7.e). This increase often exceeded 100 J/kg from the Alpine region south- and eastward, provoked by positive changes of both p90_LARA (Fig. 7.a) and p90_LLMR (Fig. 7.c). A similar, albeit weaker, behavior can be seen over the British Isles. Across the northern half of Europe, changes of p90_LARA were smaller or even regionally negative, but partic-

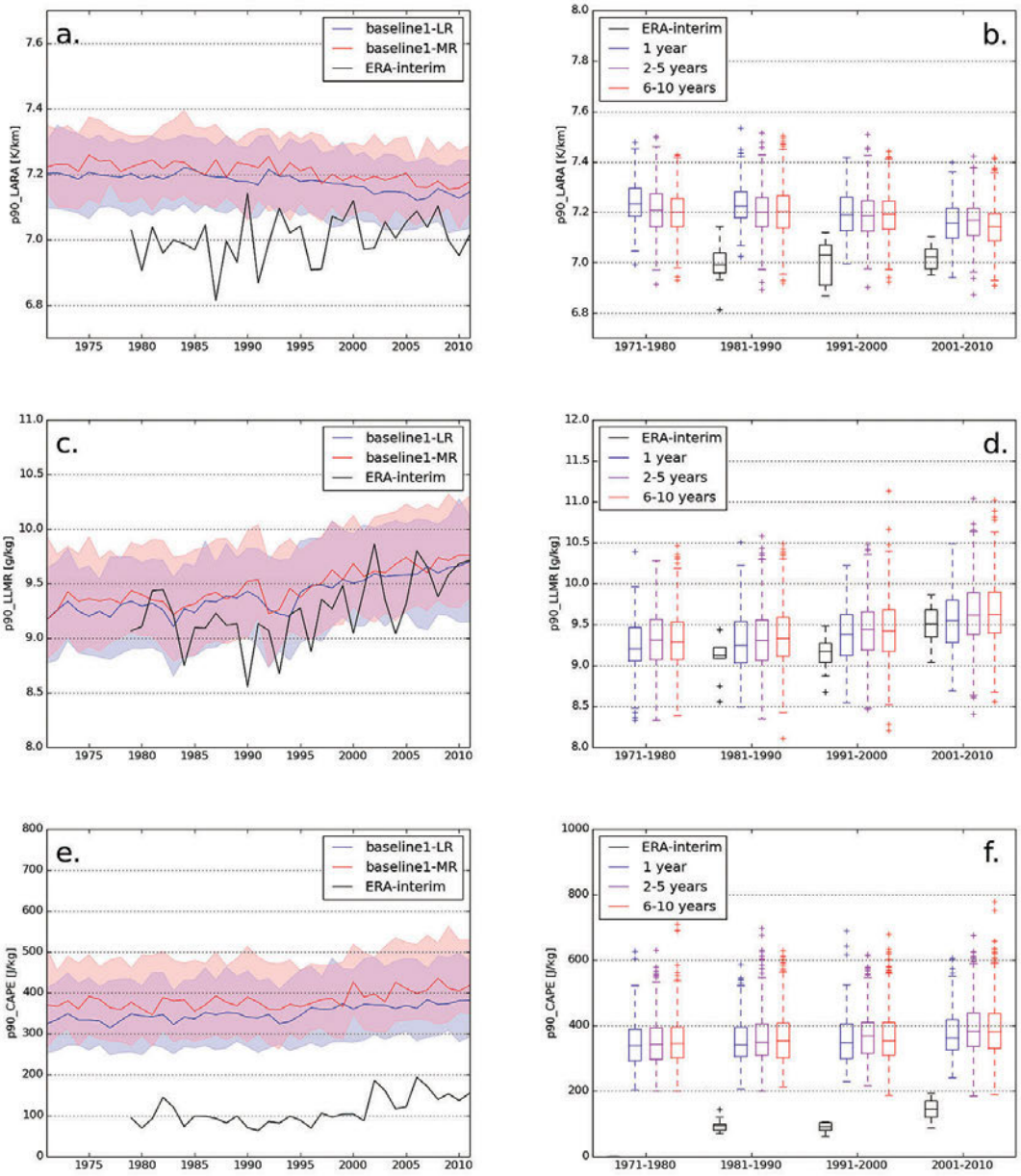

Figure 5: Annual 90th percentiles averaged over the central European sub-domain from Fig. 1. (a.) Annual 90th percentiles of 850 to 500 hPa lapse rates [K/km] in ERA-Interim (black), MPI-ESM LR (blue for ensemble median and shaded blue for ensemble spread) and MPI-ESM MR (red for ensemble median and shaded red for ensemble spread). (b.) Decadal box-and-whisker plot of annual 90th percentiles of 850 to 500 hPa lapse rates [K/km] in ERA-Interim (black) and MPI-ESM for the forecast years 1 (blue), 2–5 (pink) and 6–10 (red); the boxes are delimited by the respective 25th, 50th and 75th percentiles, outliers are shown as crosses. (c.) same as (a.), but for the mixing ratio [g/kg] 25 hPa below the surface pressure. (d.) same as (b.), but for the mixing ratio [g/kg] 25 hPa below the surface pressure. (e.) same as (a.), but for CAPE [J/kg]. (f.) same as (b.), but for CAPE [J/kg].

ularly strong increases of p90_LLMR still resulted in rising p90_CAPE. Negative interdecadal changes of p90_CAPE are confined to northeastern Spain (Catalonia), Malta and surroundings, and parts of Turkey, mostly driven by a decrease of p90_LLMR.

In the MPI-ESM simulations, p90_LARA decreased over much of Europe from the 1990s to the 2000s (Fig. 7.b) while p90_LLMR increased everywhere, most strongly in the Mediterranean region and southwestern Europe with values around 0.3 g/kg (Fig. 7.d). Increases of p90_CAPE mostly amounted to values between 20 and 50 J/kg (Fig. 7.f), indicating that the positive changes of p90_LLMR overcompensated the negative changes of p90_LARA.

Jointly, Figs. 6 and 7 reveal some interesting results. The interdecadal changes simulated by MPI-ESM are weaker and exhibit coarser spatial patterns than those derived from ERA-Interim. The changes in ERA-Interim from the 1980s to the 1990s are not reproduced at all by MPI-ESM for any of the convective parameters.

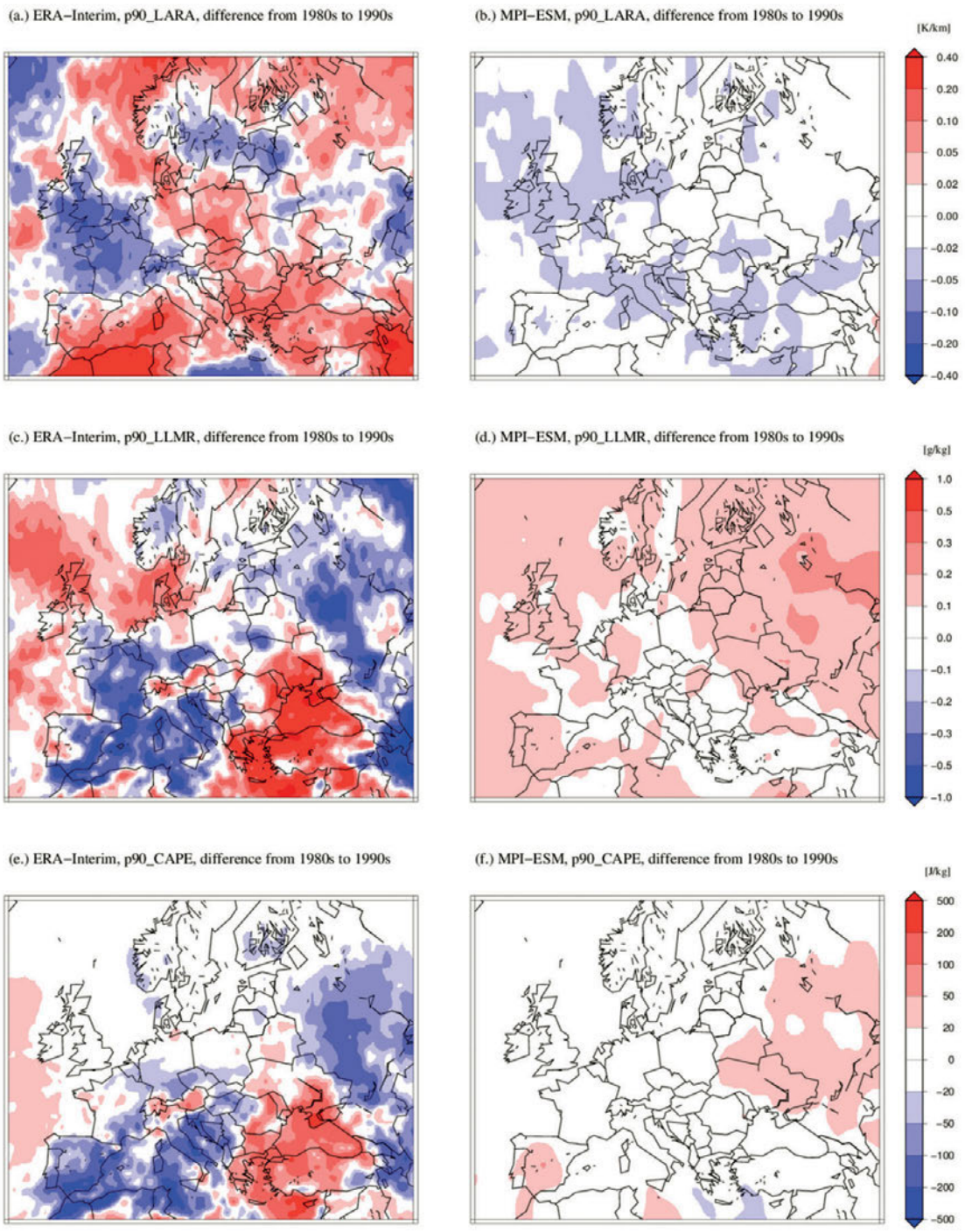

Figure 6: As Fig. 4, but for the changes of the 90th percentiles from the decade 1981–1990 to the decade 1991–2000.

The changes from the 1990s to the 2000s are obviously better predictable, and the ERA-Interim and MPI-ESM signals match at least qualitatively well. ERA-Interim and MPI-ESM show strong indications that p90_CAPE increased after the year 2000 despite the fact that lapse rates became less steep, because the effect of increasing low-level moisture dominated.

An important question is if any of these trends in ERA-Interim reflect what happened in the real atmosphere, or if these are attributable to changes of the measurement equipment used to collect assimilated data. For instance, changes in radiosonde equipment have taken place in Germany between 2003 and 2005, when the Vaisala RS80 was replaced by the Vaisala RS92 sensor. Although this change has had only minor effects on temperature measurements (Pattantyús-Ábrahám and Steinbrecht, 2015), an increase in lower troposphere moisture observed in German radiosonde data around that time (Mohr and Kunz, 2013) may, wholly or in part, be caused by these equipment changes. One

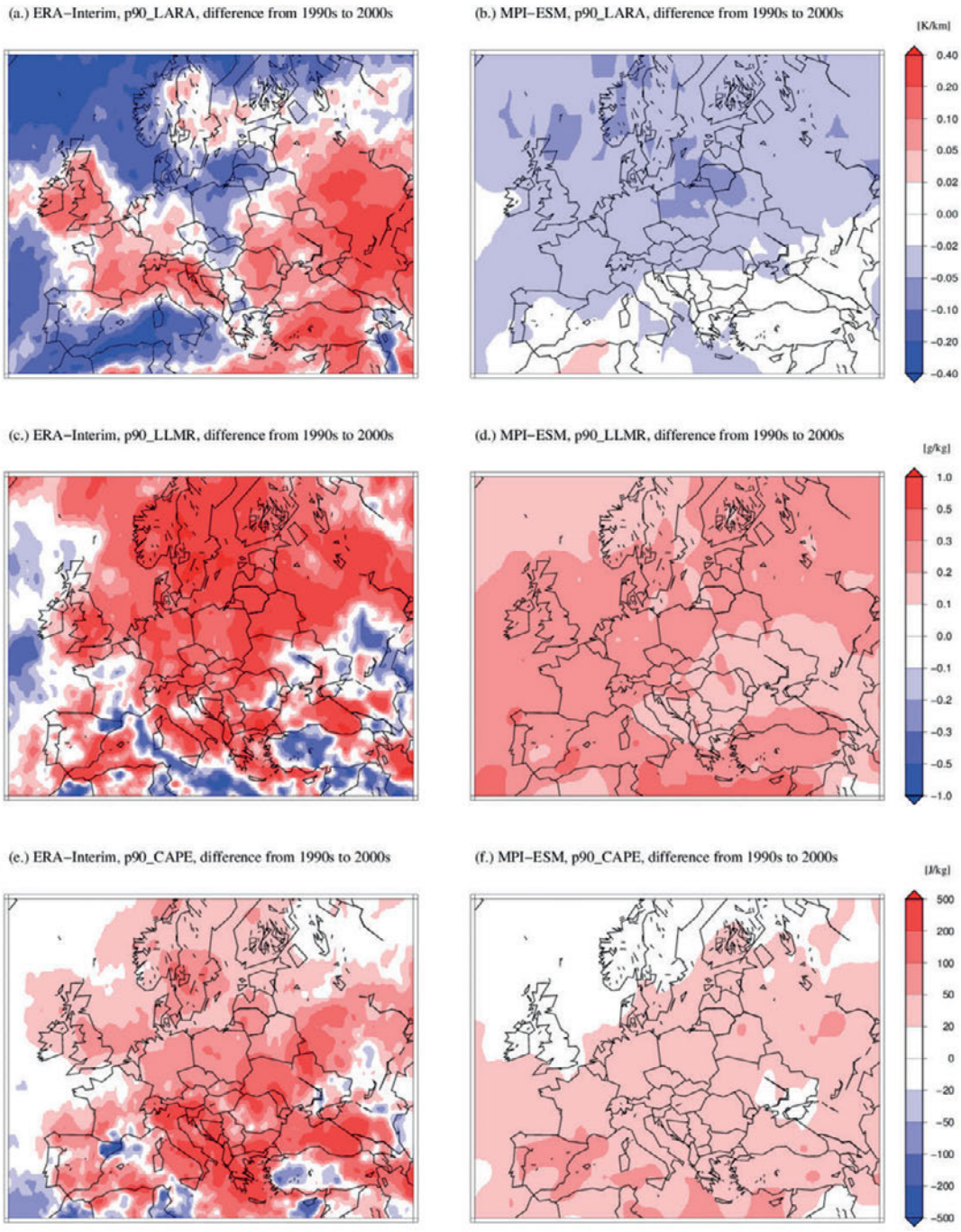

Figure 7: As Fig. 6, but for the changes of the 90th percentiles from the decade 1991–2000 to the decade 2001–2010.

may expect that reanalysis data such as ERA-interim are to some extent impacted by such changes and it could be that the increases of p90_LLMR and p90_CAPE from the 1990s to 2000s can be partly attributed to this.

Unlike in the MPI-ESM simulations, the sign of both interdecadal changes is identical only over a limited number of regions in ERA-Interim. For example, p90_CAPE decreased twice only over northeastern Spain and to a lesser extent over Tunisia and Malta, and it increased twice only over the Black Sea and surroundings and, more locally, in the Alpine region. Since the changes from the 1990s to the 2000s were remarkably larger in magnitude and occur over wide areas, they likely displayed a stronger and better predictable climatological signal. The previous changes from the 1980s to the 1990s may have reflected transient oscillations of limited extent and poorer predictability rather than longer-lasting trends.

Having checked that the 90th percentiles of all convective parameters are roughly normally distributed, we performed a t-test to examine the significance of the observed and modeled interdecadal changes. Due to rather

large interannual variations, none of the interdecadal changes were significant at the 95 % level. The maximum magnitude of t-values was a little bit above one for the interdecadal changes of the convective parameters derived from ERA-Interim, corresponding to a maximum significance roughly at the 70 % level over limited areas. However, significance was considerably lower for most gridpoints of ERA-Interim, and in particular for the entire MPI-ESM fields.

As the analyzed changes of p90_CAPE (Figs. 6.e and 7.e) depended more heavily on the variations of p90_LLMR (Figs. 6.c and 7.c) than on those of p90_LARA (Figs. 6.a and 7.a), it is interesting to link the analyzed and modeled moisture changes to concurrent temperature changes to find out to which degree they were temperature-driven. The maximum capacity of air to hold water vapor is governed by its temperature via the Clausius-Clapeyron equation: the relative change of mixing ratio with temperature at which the relative humidity remains constant is approximately +7.6 %/K at a temperature of 0 °C, subject to a slight negative temperature dependence (e.g., it is +9 %/K at −25 °C and +6 %/K at +40 °C).

Investigating the relative decadal p50_MR925 changes as a function of decadal p50_T925 changes for all ERA-Interim and MPI-ESM gridpoints on the European continent (Fig. 8), each of the linear regression slopes was below the slope following constant relative humidity. From the 1980s to the 1990s, the slopes of the regression lines were +2.8 %/K for ERA-Interim and +3.0 %/K for MPI-ESM. From the 1990s to the 2000s, they were +3.4 %/K and +1.7 %/K, respectively. A t-test of regression slopes revealed that each of them was less steep than the constant relative humidity slope at a 95 % level of significance. Warmer periods in the past few decades in Europe were therefore associated with significantly lower relative humidity, and vice versa. However, particularly for the lower panel in Fig. 8 (MPI-ESM) the variance of the residuals is of the size or even larger than the variance of the regression slope, and thus the usefulness of a linear model questionable.

This characteristic was qualitatively well captured by the MPI-ESM simulations. Its implication on thunderstorms is not entirely clear yet. However, it may be speculated that the associated higher clouds bases and enhanced subcloud evaporation (Doswell et al., 1996) could at least partly counteract a potential increase of heavy convective precipitation, which could otherwise be expected under a higher water vapor capacity of the atmosphere (Trapp et al., 2007; Trapp et al., 2009; Diffenbaugh et al., 2013).

4 Conclusions

A three decades' joint dataset of ERA-Interim reanalyses and MPI-ESM decadal climate hindcasts (1981–2010) was evaluated to study the behavior of conditions favorable for thunderstorm formation, expressed by the 90th percentiles of CAPE (p90_CAPE) and of its individual components, namely the low-level mixing ratio (p90_LLMR; chosen at 25 hPa below the surface pressure) and the vertical temperature gradient between 850 and 500 hPa (p90_LARA). This allows assessing the strengths and weaknesses of MPI-ESM decadal predictions as well as finding interdecadal changes in the frequency of situations conducive to thunderstorm development.

The spatial structures of temperature and humidity fields, and derived convective parameters, showed a good agreement between MPI-ESM and ERA-Interim. The found spatial distribution of the 90th percentile of CAPE is in line with previous results of Marsh et al. (2009), who found the highest frequency of conditions for thunderstorms over Spain and from northern Italy into the western Balkan states.

However, the most remarkable result was a strong negative bias of the median 500 hPa temperature (p50_T500) in MPI-ESM. This bias is on the order of −1.5 to −2 K across most of Europe. In contrast, no large bias was found for the median 925 hPa temperature (p50_T925) and a positive bias on the order of 0.2 g/kg for the median 925 hPa mixing ratio (p50_MR925). These biases suggest a systematic departure of the "preferred" internal model climate from the true atmospheric state; however, for p50_T500, the bias magnitude was found to be largest for the first forecast year, which indicates that a considerable part of this particular bias is likely "imported" from the initial conditions.

The simulation of a too moist lower troposphere and a too cold mid-troposphere exaggerates the atmosphere's readiness to produce convective storms, which was also reflected by mostly positive biases of p90_LARA, p90_LLMR and p90_CAPE. An important sink of CAPE in a model and the real atmosphere is the occurrence of deep convection. The high CAPE bias may be a result of a relative inactivity of MPI-ESM's convective parameterization scheme across Europe. The unrealistically large reservoirs of CAPE that persist in the absence of model convection will cause any proxy that is based on CAPE to overestimate the atmosphere's readiness to produce convective storms.

"Low resolution" (LR) and "mixed resolution" (MR) runs of MPI-ESM, the latter computed at higher resolutions of the oceanic and atmospheric model components, showed a similar behavior for p50_T500, p50_T925 and p50_MR925. However, for convective situations represented by p90_LARA, p90_LLMR and p90_CAPE, the found biases were even slightly larger in the MR runs, contrary to the assumption that an increasing resolution would improve the model's performance. Although an increased skill of the higher MR resolution was shown for the surface air temperature over the tropical Pacific Ocean (Pohlmann et al., 2013), we find decreased skill for the mid-tropospheric temperature and derived convective parameters across Europe. Apparently, the accuracy of the model physics is a greater source of error for these parameters than that which would be caused

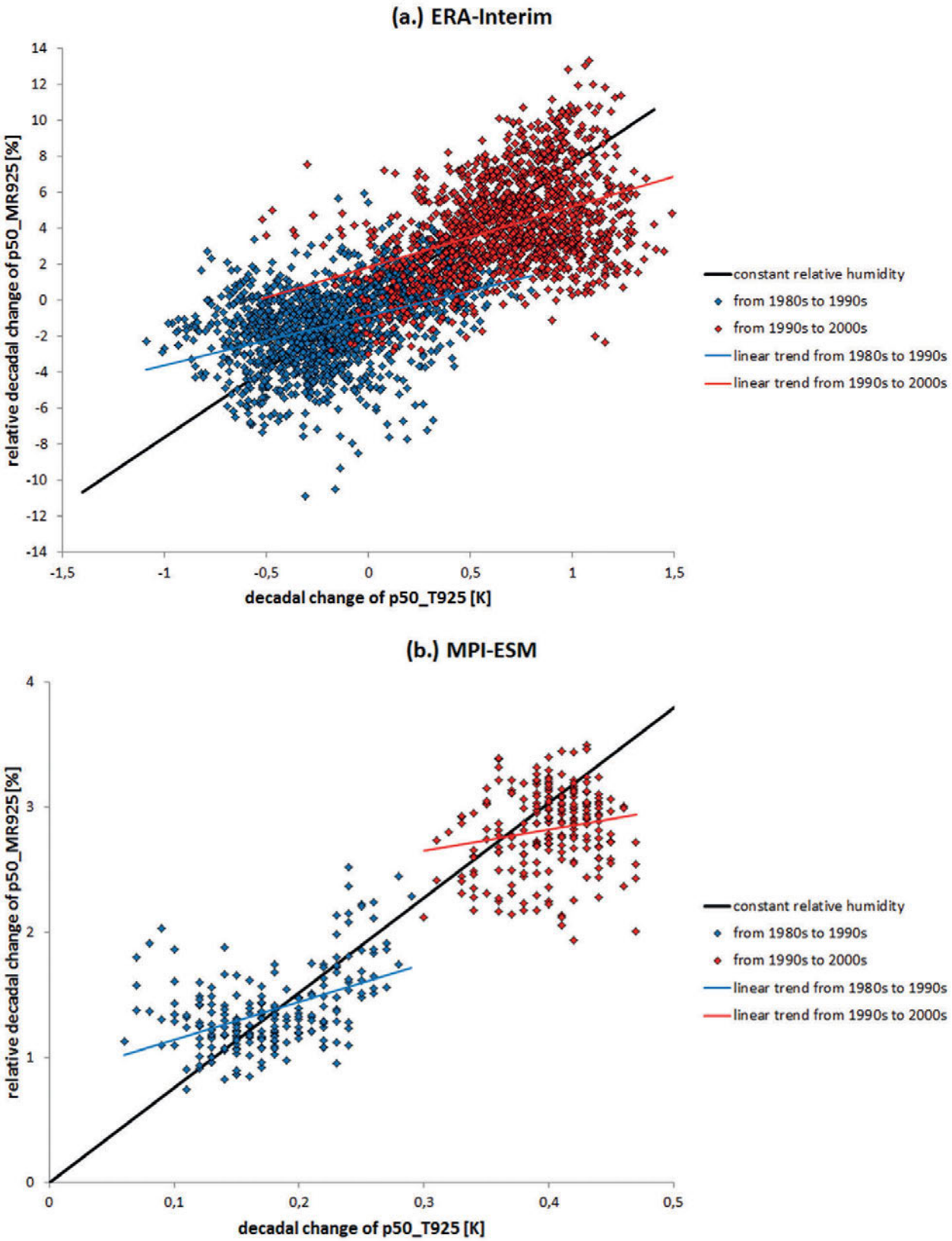

Figure 8: Relative decadal changes of p50_MR925 [%] as a function of decadal changes of p50_T925 [K] from 1981–1990 to 1991–2000 (blue) and from 1991–2000 to 2001–2010 (red). Each diamond represents one gridpoint on the European continent. Colored lines are the best linear regression fits. Relative humidity remains constant along the black lines, which represent a relative p50_MR925 change of 7.6 %/K, valid for a temperature of 0 °C according to the Clausius-Clapeyron equation. (a.) ERA-Interim reanalyses; (b.) MPI-ESM hindcasts. MPI-ESM gridpoints are fewer due to its coarser resolution. Note the different scaling of the axes.

by a reduced resolution. It appears that the accuracy of physical parameterizations does not keep pace with the refinement of model resolution from LR to MR.

The present quality of the simulation of thunderstorm proxies by MPI-ESM makes it difficult to use MPI-ESM results for estimating the future development of severe storm events on a decadal basis. For the time being there are at least two possible solutions, either waiting for model improvements or applying a bias correction. As the bias pattern is rather stationary, a linear regression may suffice to remove biases of MPI-ESM convective parameters. Naturally the added skill of such a bias corrected model needs to be carefully demonstrated.

Acknowledgments

The present study was funded by the German Bundesministerium für Bildung und Forschung (BMBF) through the research program "Mittelfristige Klimaprognosen" (MiKliP), grant FKZ: 01LP1117A. Cordial thanks are expressed to Anja T. Westermayer for her help with input/output issues in the NetCDF data format. The comments of two anonymous reviewers helped to improve the manuscript.

References

Balmaseda, M.A., K. Mogensen, A.T. Weaver, 2012: Evaluation of the ECMWF Ocean Reanalysis System ORAS4. – Quart. J. Roy. Meteor. Soc., **139**, 1132–1161.

Böhm, U., M. Kücken, W. Ahrens, A. Block, D. Hauffe, K. Keuler, B. Rockel, A. Will, 2006: CLM – The climate version of LM: Brief description and long-term applications. – COSMO Newsletter **6**, pp. 225–235, Consortium for Small-Scale Modelling, Offenbach am Main, Germany.

Brands, S., S. Herrera S., J. Fernández, J. Gutiérrez, 2013: How well do CMIP5 Earth System Models simulate present climate conditions in Europe and Africa? – Climate Dyn., **41**, 803–817.

Brooks, H.E., 2013: Severe thunderstorms and climate change. – Atmos. Res., **123**, 129–138.

Brooks, H.E., N. Dotzek, 2008: The spatial distribution of severe convective storms and an analysis of their secular changes. – In: Diaz. H.F., R. Murnane (Eds.): Climate Extremes and Society. – Cambridge University Press, Cambridge, UK, 33–53.

Brooks, H.E., J.W. Lee, J.P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. – Atmos. Res., **67–68**, 73–94.

Collins, W.D., C.M. Bitz, M.L. Blackmon, G.B. Bonan, C.S. Bretherton, J.A. Carton, P. Chang, S.C. Doney, J.J. Hack, T.B. Henderson, J.T. Kiehl, W.G. Large, D.S. McKenna, B.D. Santer, R.D. Smith, 2006: The Community Climate System Model Version 3 (CCSM3). – J. Climate **19**, 2122–2143.

Craven, J.P., H.E. Brooks, 2004: Baseline climatology of sounding derived parameters associated with deep, moist convection. – Nat. Wea. Digest **28**, 13–24.

Craven, J.P., R.E. Jewell, H.E. Brooks, 2002: Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. – Wea. Forecast. **17**, 885–890.

Dee, D.P., S.M. Uppala, A.J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J. -J. Morcrette, B. -K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. -N. Thépaut, F. Vitart, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. – Quart. J. Roy. Meteor. Soc. **137**, 656, 553–597.

Del Genio, A.D., M.-S. Yao, J. Jonas, 2007: Will moist convection be stronger in a warmer climate? – Geophys. Res. Lett., **34**, L16703.

Devis, A., N.P.M. van Lipzig., M. Demuzere, 2014: A height dependent evaluation of wind and temperature over Europe in the CMIP5 Earth System Models. – Climate Res. **61**, 41–56.

Diffenbaugh, N.S., M. Scherer, R.J. Trapp, 2013: Robust increases in severe thunderstorm environments in response to greenhouse forcing. – Proc. Natl. Acad. Sci., **110**, 16361–16366.

Doswell, C.A. III, H.E. Brooks, R.A. Maddox, 1996: Flash flood forecasting: an ingredients-based methodology. – Wea. Forecast. **11**, 560–580.

Dotzek, N., 2003. An updated estimate of tornado occurrence in Europe. – Atmos. Res., **67–68**, 153–161.

Dotzek, N., P. Groenemeijer, B. Feuerstein, A.M. Holzer, 2009: Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. – Atmos. Res., **93**, 575–586.

Eccel, E., P. Cau, K. Riemann-Campe, F. Biasioli, 2012: Quantitative hail monitoring in an alpine area: 35-year climatology and links with atmospheric variables. – Int. J. Climatol., **32**, 503–517.

Giorgetta, M.A., J.H. Jungclaus, C.H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, K. Glushak, V. Gayler, H. Haak, H.-D. Hollweg, T. Ilyina, S. Kinne, L. Kornblueh, D. Matei, T. Mauritsen, U. Mikolajewicz, W.A. Müller, D. Notz, F. Pithan, T. Raddatz, S. Rast, R. Redler, E. Roeckner, H. Schmidt, R. Schnur, J. Segschneider, K.D. Six, M. Stockhause, C. Timmreck, J. Wegner, H. Widmann, K.-H. Wieners, M. Claussen, J. Marotzke, B. Stevens, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project Phase 5. – J. Adv. Model. Earth Syst., **5**, 572–597.

Goddard, L., A. Kumar, A. Solomon, D. Smith, G. Boer, P. Gonzalez, V. Kharin, W. Merryfield, C. Deser, S J. Mason, B.P. Kirtman, R. Msadek, R. Sutton, E. Hawkins, T. Fricker, G. Hegerl, C.A.T. Ferro, D.B. Stephenson, G.A. Meehl, T. Stockdale, R. Burgman, A.M. Greene, Y. Kushnir, M. Newman, J. Carton, I. Fukumori, T. Delworth, 2013: A verification framework for interannual-to-decadal predictions experiments. – Climate Dyn., **40**, 245–272.

Graf, M., M. Sprenger, R.W. Moore, 2011: Central European tornado environments as viewed from a potential vorticity and Lagrangian perspective. – Atmos. Res., **101**, 31–45.

Groenemeijer, P., T. Kühne, 2014: A climatology of tornadoes in Europe: Results from the European Severe Weather Database. – Mon. Wea. Rev., **142**, 4775–4790.

Jury, M.W., A.F. Prein, H. Truhetz, A. Gobiet, 2015: Evaluation of CMIP5 models in the Context of Dynamical Downscaling over Europe, – J. Climate **28**, 5575–5582.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, R. Jenne, D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. – Bull. Amer. Meteor. Soc., **77**, 437–472.

Kapsch, M.-L., M. Kunz, R. Vitolo, T. Economou, 2012: Long-term trends of hail-related weather types in an ensemble of regional climate models using a Bayesian approach. – J. Geophys. Res., **117**, D15107, DOI:10.1029/2011JD017185.

Keenlyside, N.S., J. Ba, 2010: Prospects for decadal climate predictions. – WIREs Clim Change **1**, 627–635.

Kunz, M., J. Sander, Ch. Kottmeier, 2009: Recent trends of thunderstorm and hailstorm frequency and their relation to atmospheric characteristics in southwest Germany. – Int. J. Climatol. **29**, 2283–2297.

Latif, M., 2011: Uncertainty in climate change projections. – J. Geochemical Exploration **110**, 1–7.

Latif, M., N.S. Keenlyside, 2011: A perspective on decadal climate variability and predictability. – Deep Sea Research Part II: Topical Studies in Oceanography **58**, 1880–1894.

Liu, Z., 2012: Dynamics of interdecadal climate variability: A historical perspective. – J. Climate **25**, 1963–1995.

Markowski, P., Y. Richardson, 2010: Mesoscale Meteorology in Mid-Latitudes. – Wiley-Blackwell, 430 pp.

Marsh, P.T., H.E. Brooks, D.J. Karoly, 2007: Assessment of the severe weather environment in North America simulated by a global climate model. – Atmos. Sci. Lett. **8**, 100–106.

Marsh, P.T., H.E. Brooks, D.J. Karoly, 2009: Preliminary investigation into the severe thunderstorm environment of Europe simulated by the Community Climate System Model 3. – Atmos. Res., **93**, 607–618.

Meehl, G.A., L. Goddard, J. Murphy, R.J. Stouffer, G. Boer, G. Danabasoglu, K. Dixon, M.A. Giorgetta, A.M. Greene, E. Hawkins, G. Hegerl, D. Karoly, N. Keenlyside, M. Kimoto, B. Kirtman, A. Navarra, R. Pulwarty, D. Smith, D. Stammer, T. Stockdale, 2009: Decadal prediction – can it be skillful? – Bull. Amer. Meteor. Soc., **90**, 1467–1485.

Mohr, S., M. Kunz, 2013: Trend analysis of convective indices relevant for hail events in Germany and Central Europe. – Atmos. Res., **123**, 211–228.

MunichRe, 2006: Highs and lows: Weather risks in central Europe. – Munich Re Group publication.

Müller, W.A., J. Baehr, H. Haak, J.H. Jungclaus, J. Kröger, D. Matei, D. Notz, H. Pohlmann, J.S. Von Storch, J. Marotzke, 2012: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. – Geophys. Res. Lett., **39**, L22707.

Pattantyús-Ábrahám M., W. Steinbrecht, 2015: Temperature Trends over Germany from Homogenized Radiosonde Data. – J. Climate **28**, 5699–5715, DOI: 10.1175/JCLI-D-14-00814.1.

Pohlmann, H., W.A. Müller, K. Kulkarni, M. Kameswarrao, D. Matei, F.S.E. Vamborg, C. Kadow, S. Illing, J. Marotzke, 2013: Improved forecast skill in the tropics in the new MiKliP decadal climate predictions. – Geophys. Res. Lett., **40**, 5798–5802.

Rasmussen, E.N., D.O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. – Wea. Forecast. **13**, 1148–1164.

Sander, J., 2011: Extremwetterereignisse im Klimawandel: Bewertung der derzeitigen und zukünftigen Gefährdung. – PhD Thesis, University of Munich, 125 pp.

Schaefer, J., R. Edwards, 1999: The SPC Tornado/Severe Thunderstorm Database. – Preprints, 11th Conf. on Applied Climatology, Dallas, TX, Amer. Meteor. Soc., 603–606.

Smith, D.M., S. Cusack, A.W. Colman, C.K. Folland, G.R. Harris, J.M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. – Science **317**, 796–799.

Stevens, B., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast, M. Salzmann, H. Schmidt, J. Bader, K. Block, R. Brokopf, I. Fast, S. Kinne, L. Kornblueh, U. Lohmann, R. Pincus, T. Reichler, E. Roeckner, 2013: Atmospheric component of the MPI-M Earth System Model: ECHAM6. – J. Adv. Model. Earth Syst., **5**, 146–172.

Trapp, R.J., N.S. Diffenbaugh, H.E. Brooks, M.E. Baldwin, E.D. Robinson, J.S. Pal, 2007: Changes in severe thunderstorm frequency during the 21st century due to anthropogenically enhanced global radiative forcing. – Proc. Natl. Acad. Sci., **104**, 19719–19723.

Trapp, R.J., N.S. Diffenbaugh, A. Gluhovsky, 2009: Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. – Geophys. Res. Lett., **36**, L01703.

Uppala, S.M., P.W. Kållberg, A.J. Simmons, U. Andrae, V. Da Costa Bechtold, M. , Fiorino, J.K. Gibson, J. Haseler, A. Hernandez, G.A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R.P. Allan, E. Andersson, K. Arpe, M.A. Balmaseda, A.C.M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. McNally, J.F. Mahfouf, J-J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljević, P. Viterbo, J. Woollen, 2005: The ERA-40 re-analysis. – Quart. J. Roy. Meteor. Soc., **131**, 2961–3012.

Van Klooster, S.L., P.J. Roebber, 2009: Surface-based convective potential in the contiguous United States in a business-as-usual future climate. – J. Climate **22**, 3317–3330.

Verbout, S.M., H.E. Brooks, L.M. Leslie, D.M. Schultz, 2006: Evolution of the U.S. Tornado Database: 1954–2003. – Wea. Forecast. **21**, 86–93.

Westermayer, A.T., P. Groenemeijer, G. Pistotnik, R. Sausen, E. Faust, 2016: Identification of favroable environments for thunderstorms in reanalysis data. – Meteorol. Z., DOI:10.1127/metz/2016/0754.