



Benefits and limitations of regional multi-model ensembles for storm loss estimations

Markus G. Donat*, Gregor C. Leckebusch, Simon Wild, Uwe Ulbrich

Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6–10, 12165 Berlin, Germany

ABSTRACT: Spatial patterns of near-surface wind speeds and resulting loss potentials associated with severe winter storms were investigated in multi-model simulations with regional climate models (RCMs), driven by ERA40 re-analyses. The benefits and limitations of dynamical downscaling for windstorm loss calculations were explored, including a quantification of the performance of the multi-model ensemble as a whole and the systematic investigation of the influence of model selection on the ensemble results. A comparison of the wind fields in the different models revealed both systematic biases in individual RCMs and model-specific anomalies over mountainous regions. Further, a storm loss model was applied to the RCM wind fields, and the calculated losses were validated against observed annual insurance loss data available for Germany. Generally, a distinct advantage from dynamical downscaling was obvious. However, all RCMs failed in realistically simulating 1 specific major event. If this particular event was excluded from the considerations, almost all simulations revealed high correlations (above 0.8) with observed losses, comparable to losses calculated directly from the large-scale reanalysis wind field. For the best performing models, considerably higher loss correlations up to 0.95 were obtained, suggesting that the high-resolution RCMs exceeded the value of assimilation in the driving data for the area considered. Combining calculated losses from the individual RCMs into a multi-model ensemble, the performance of the ensemble mean was as good as the performance of the best single model. Examining all possible sub-ensembles, we found that generally a higher minimum performance was obtained with a larger number of ensemble members, whereas the maximum performance was hardly affected by the ensemble size.

KEY WORDS: Multi-model ensemble · Regional climate models · Dynamical downscaling · Winter storms · Storm losses · Impact model

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Severe winter storms frequently cause heavy damage in Europe, in particular to infrastructure and environments (e.g. buildings, electricity supply, forests). They can also entail restrictions to traffic and may lead to injury or even loss of lives. Thus, such extreme events are of high relevance to many parts of society. The strong wind fields generally affect large areas, contributing to high accumulated loss amounts. For example, the storm 'Kyrill' (affecting large parts of Europe between 17 and 19 January 2007, cf. Fink et al. 2009) caused estimated losses of about 4 to 4.5 billion Euro (Swiss Re 2008). A series of European winter storms in December 1999 ('Anatol', 'Lothar', 'Martin', cf. Ulbrich et al. 2001) and in early 1990 (e.g. 'Daria', 'Hertha', 'Vivian', 'Wiebke', cf. McCallum & Norris 1990) both caused total insur-

ance losses exceeding 10 billion Euro across Europe. Compared to loss amounts from other natural disasters, wind storms are the most relevant cause of losses in central Europe: in Germany, for example, they account for 53% of economic losses due to natural hazards and 64% of insured losses (Munich Re 1999, 2007).

An estimation of storm losses is important for a multitude of applications. Examples include the assessment of changed risk under climate change conditions as well as near-term loss estimations shortly after or even before a storm occurs (e.g. to effectively coordinate relief operations and financial planning) or the identification of regions with particular high risk of damage. Accurate loss estimates are thus desirable not only from the insurance companies' perspective, but also for e.g. political decision makers, emerging institutions or infrastructural planning.

*Email: markus.donat@met.fu-berlin.de

For the calculation of storm losses, different loss functions can be found in the literature. For example, Heneka et al. (2006) derived local empirical relations for vulnerability (i.e. the link between the storm event and the consequent damage) based on insured losses that occurred during historic storm events. Yet insurance loss information at a high spatial and temporal resolution is usually not available for public research. Common approaches for loss calculations are based on physical assumptions, considering the kinetic energy of the wind or the wind power, for instance. Thus, in general polynomial relations (square or cubic functions) between loss and wind speed are used (Palutikof & Skellern 1991, Munich Re 1993, Dorland et al. 1999, Klawka & Ulbrich 2003).

The storm loss model proposed by Klawka & Ulbrich (2003) is based on a cubic relationship between wind speed and losses. It properly calculates annual insurance losses in Germany and Great Britain not only based on observational wind data from synoptic weather stations, but also from reanalysis data on spatial scales of about 100 km (Leckebusch et al. 2007, Pinto et al. 2007). In these studies, it was also applied to global climate model data, working on spatial scales of about 200 km.

Dynamical downscaling with regional climate models (RCMs) is generally applied to obtain atmospheric information at a higher spatial resolution than that provided by global climate models, in particular for information on the influence of regional orographic characteristics. Additionally, physical processes acting on scales not resolved by the driving large-scale models can exert an effect on simulated regional wind patterns. The advantage of dynamical downscaling with respect to wind fields over complex terrain was demonstrated, e.g. by Goyette et al. (2001), Žagar et al. (2006), Hofherr & Kunz (2010), and Kund et al. (2010) obtaining wind speeds in good accordance with observations over mountainous regions. Such an improvement should in turn allow for a higher accuracy of storm loss calculations, as suggested by Heneka et al. (2006).

On the other hand, numerical climate model simulations (and thus also dynamical downscaling) are affected by different sources of uncertainty. The most important are uncertainties in parameterisations of physical processes and in the numerical formulation of the RCMs, sensitivity to the initial conditions or to boundary conditions and uncertainty due to internal variability generated within the model domain. Combining different models into a multi-model ensemble (MME) generally increases the skill, reliability and consistency of model projections, both in weather forecasting and in seasonal prediction applications (e.g. Hagedorn et al. 2005). Compared to a single model, an MME can also be favourable for climate–timescale simulations (cf. Palmer & Räisänen

2002, Collins 2007, Räisänen 2007, Ulbrich et al. 2008). MMEs primarily sample initial condition and model uncertainties (Tebaldi & Knutti 2007). However, a quantification of the MME advantages on the climate timescale is difficult, as, for example, reliable data for verification are often missing on this timescale. Further, the way the ensemble is constructed has an influence on the ensemble performance (Weigel et al. 2008).

This study addresses added values and limitations of dynamical downscaling and of combining different models into an MME for storm loss assessment applications. RCM simulations driven by reanalysis data should reproduce the observed chronology of weather situations. This offers the opportunity of a sound validation of these climate simulations, considering not only climate mean state and variability, but also the chronology of events. We used annual accumulated insurance loss data as a proxy describing the occurrence of severe storms. Loss potentials calculated from RCM outputs were then compared with the observed loss values to allow for an estimation of the added value of downscaling for each RCM. In a further step, the benefit from combining the different models into a multi-model ensemble of climate-scale simulations was investigated, and the effect of different possible model combinations included in an ensemble was examined systematically. Again, the agreement with time series of observed loss data was used for verification.

2. DATA AND METHODS

2.1. Data

All RCM simulations were carried out for a common domain including the whole continental European area (from approximately 10° W to 40° E and 30° to 65° N). Overall, a set of 11 different RCMs was considered (Table 1). All models were used for dynamical downscaling of ERA40 reanalysis data (Uppala et al. 2005). The horizontal resolution of ERA40 is approximately 1.125° (i.e. about 120 km in central Europe). Almost all RCMs were run at 2 different spatial resolutions of 0.44° (~50 km) and 0.22° (~25 km). Two models were run at only 1 resolution (GKSS-CLM: 0.44°, C4I-RCA3: 0.22°). The RCM simulations were generally driven at the domain's lateral boundaries; only the GKSS-CLM simulation makes use of a spectral nudging technique (von Storch et al. 2000).

This study was based on daily maxima of simulated 10 m wind speeds. This parameter was archived from all considered runs as the daily maximum speed value over all time steps. From ERA40, only 6-hourly instantaneous values of wind speed were available. Here, a daily maximum was calculated as the maximum of the

Table 1. Regional climate models (RCMs) included in this study (ENSEMBLES setup)

RCM	Institute	Resolution (km)	Gust	Source calculation
C4I-RCA3	Community Climate Change Consortium for Ireland	25	X	Kjellström et al. (2005)
CHMI-ALADIN	Czech Hydrometeorological Institute	25 and 50		Bubnova et al. (1995)
CNRM-RM4.5	Météo France/Centre National de Recherches Météorologiques	25 and 50		Gibelin & Déqué (2003)
DMI-HIRHAM	Danish Meteorological Institute	25 and 50		Christensen et al. (1996)
ETHZ-CLM	Swiss Federal Institute of Technology	25 and 50	X	Stappeler et al. (2003) Jaeger et al. (2008)
GKSS-CLM	GKSS Forschungszentrum Geesthacht, Germany	50		von Storch et al. (2000) Stappeler et al. (2003) Jaeger et al. (2008)
HC-HadRM3	UK Met Office, Hadley Centre	25 and 50		Jones et al. (1995)
KNMI-RACMO2	Royal Netherlands Meteorological Institute	25 and 50	X	Lenderink et al. (2003)
METNO-HIRHAM	Norwegian Meteorological Institute	25 and 50		Christensen et al. (1996)
MPI-REMO	Max Planck Institute for Meteorology, Hamburg, Germany	25 and 50	X	Jacob & Podzun (1997) Jacob (2001)
SMHI-RCA3	Swedish Meteorological and Hydrological Institute	25 and 50	X	Kjellström et al. (2005) Samuelsson et al. unpubl.

4 instantaneous wind speeds stored at 00:00, 06:00, 12:00 and 18:00 h UTC. This value is expected to be slightly smaller than the maximum over all time steps (cf. Pinto et al. 2007, Rockel & Woth 2007), inducing a small inhomogeneity in our data basis. The daily maximum wind speed data are hereafter referred to as WIMAX.

Some of the RCMs also featured a gust parameterisation; a daily maximum 10 m gust wind speed was available for 5 models. For 4 of these (ETHZ-CLM, SMHI-RCA3, KNMI-RACMO2 and MPI-REMO), the data were available at both resolutions, and 1 model with gust parameterisation (C4I-RCA3) was only run at 25 km resolution. Daily maxima of modelled gusts are hereafter referred to as GUST.

The 10 m wind speed (6-hourly, see above) and a 10 m wind gust were also available for the ERA40 dataset. As in the RCMs, the latter is a model diagnostic parameter, hence its calculation is based on parameterisation. Different approaches for calculating the gust wind speeds are used in the different models. In the RCA3 models (SMHI and C4I), a wind gust estimate method based on physical considerations (Brasseur 2001) is implemented. It assumes that surface gusts result from the deflection of air parcels from the upper boundary layer, brought down by turbulent eddies. This method takes into account the turbulence kinetic energy, the mean wind and the stability in the boundary layer. ERA40 and KNMI-RACMO2 use a scheme based on the similarity relation by Panofsky et al. (1977). Here, a standard deviation of the near-surface wind is calculated, again taking into account the static stability of the boundary layer. The maximum

gust speed is then estimated by adding a term including this standard deviation to the 10 m wind speed (cf. White et al. 2003). MPI-REMO and ETHZ-CLM calculate gusts based on empirical assumptions taking into account the turbulence kinetic energy in the lowest model layer (Schrodin 1995).

As almost all storm events causing severe damage occur during boreal winter (Klawa & Ulbrich 2003, Munich Re 2007), the analyses presented here concern the extended winter period from October to March.

Annual accumulated insurance loss data for Germany were provided by the Gesamtverband der Deutschen Versicherungswirtschaft e.V. (German Insurance Association, hereafter GDV) for the period 1970–2000. They included losses to residential buildings due to storm and hail events. Large losses are usually caused by intense winter storm events affecting a large area; losses due to convective events and hail only account for a small percentage of the total annual loss (Klawa & Ulbrich 2003, GDV pers. comm.). Annual loss values are given as loss ratios, i.e. the relationship between insured claims and totally insured values. An advantage of this measure is that inflation can be neglected, as it is included in both insured values and loss. The unit is given in Euros (€) per 1000 € (i.e. in ‰).

2.2. Calculation of storm-induced losses

Loss potentials for Germany were calculated using a storm loss model that was developed by Klawa & Ulbrich (2003), based on observational wind speed data from German weather stations. Recently, a ver-

sion of this model was also applied to reanalysis and global climate model data (Leckebusch et al. 2007, Pinto et al. 2007), well reproducing observed losses and further used for the estimation of future loss changes. A detailed description of the storm loss model can be found in the previous literature; here its basic features are briefly recapitulated. Annual loss ratios (the ratio of insured values affected by storm losses) were calculated by the following equation:

$$\text{Loss ratio} = A \times \sum_{\text{area}}^{\text{Germany}} \text{pop}(\text{area}) \times \sum_{\text{days}}^{\text{year}} \text{LLI} + B \quad (1)$$

with local loss index

$$\text{LLI} = \max \left[0; \left(\frac{v_{\max}(\text{area}, \text{day})}{v_{98}(\text{area})} - 1 \right)^3 \right] \quad (2)$$

In this function, v_{\max} is the daily maximum wind speed (i.e. WIMAX or GUST) in a grid box, and v_{98} is the local 98th percentile of daily maximum wind speeds. Thus, it is assumed that losses occur locally during the 2% of days with the strongest winds. Using a relative threshold, the loss function takes into account model biases of simulated wind speeds. A is the obtained regression coefficient from calibrating calculated losses with the insurance data and B is the axis intercept; $\text{pop}(\text{area})$ is the population density. Population density is regarded as a proxy for insured values, because information about spatial distribution of total insured values is not publicly available. For central European housing, the assumption is reasonable that insured values are proportional to the population density (Klawa & Ulbrich 2003). For the calculations in this study, we used gridded population density data for the year 2000, on a $0.25^\circ \times 0.25^\circ$ grid (CIESIN & CIAT 2005). In contrast to the RCMs, the population data were provided on a regular grid. The mapping between RCM winds and population data was performed by searching for the nearest neighbours.

The regression coefficient A could not be computed from event-based loss numbers, as the respective values were only available from a few events, at the same time being of limited reliability. Thus, we used a loss data set containing yearly accumulated loss ratios for the whole area of Germany, provided by the GDV. The annually integrated loss indices at the individual grid points in Germany were accumulated, weighted by the population density. A linear regression was then used to scale the ‘raw losses’ as calculated by the loss function towards the GDV loss ratios (Fig. 1 exemplifies the regression for losses calculated from ERA40). The regression between calculated ‘raw losses’ and observed loss ratios was calculated for each model individually. The ensemble mean of losses was calculated for each year by averaging the losses of the individual ensemble members for this year.

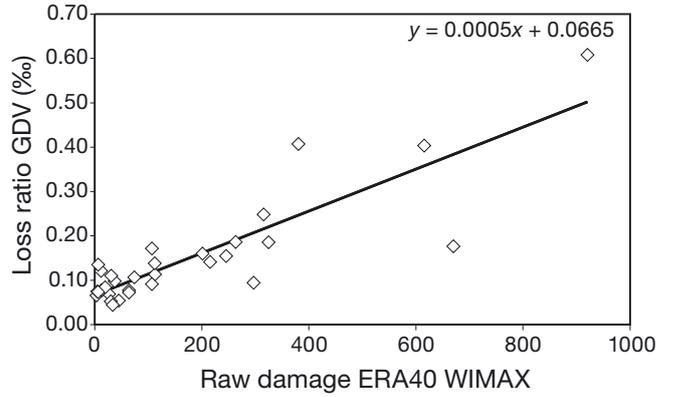


Fig. 1. Regression of calculated raw loss data from ERA40 with annual loss ratios provided by the German Insurance Association (GDV) for the years 1970–2000

3. RESULTS

3.1. Simulated wind fields

Even though the loss model takes systematic biases of extreme wind speeds into account, it is worthwhile considering characteristics in the RCM (extreme) wind fields. Comparison of the 98th percentile of WIMAX in the ERA40 reanalysis (Fig. 2) and the different ERA40-driven RCM simulations at 50 and 25 km resolution (Fig. 3) revealed considerable differences between the individual models and ERA40 itself. On the one hand, systematic biases occurred: some models simulated higher wind speeds than those in ERA40 in almost all regions (e.g. GKSS-CLM, DMI-HIRHAM), others somewhat lower wind speeds (e.g. CNRM-RM4.5, CHMI-ALADIN). On the other hand, there were differ-

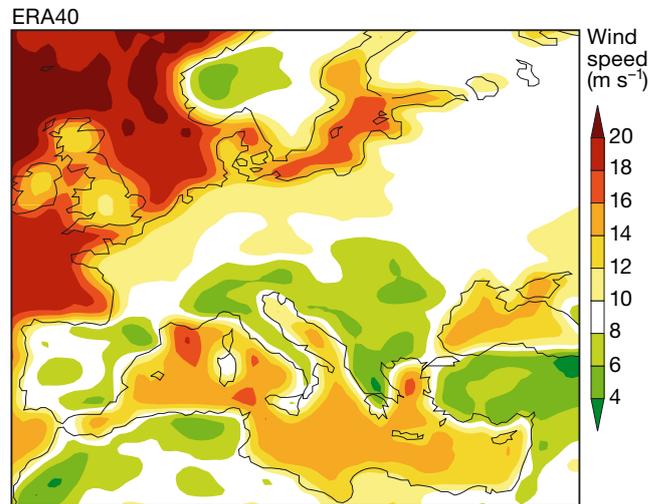


Fig. 2. 98th percentile of daily maximum wind speed (WIMAX) in the ERA40 reanalysis

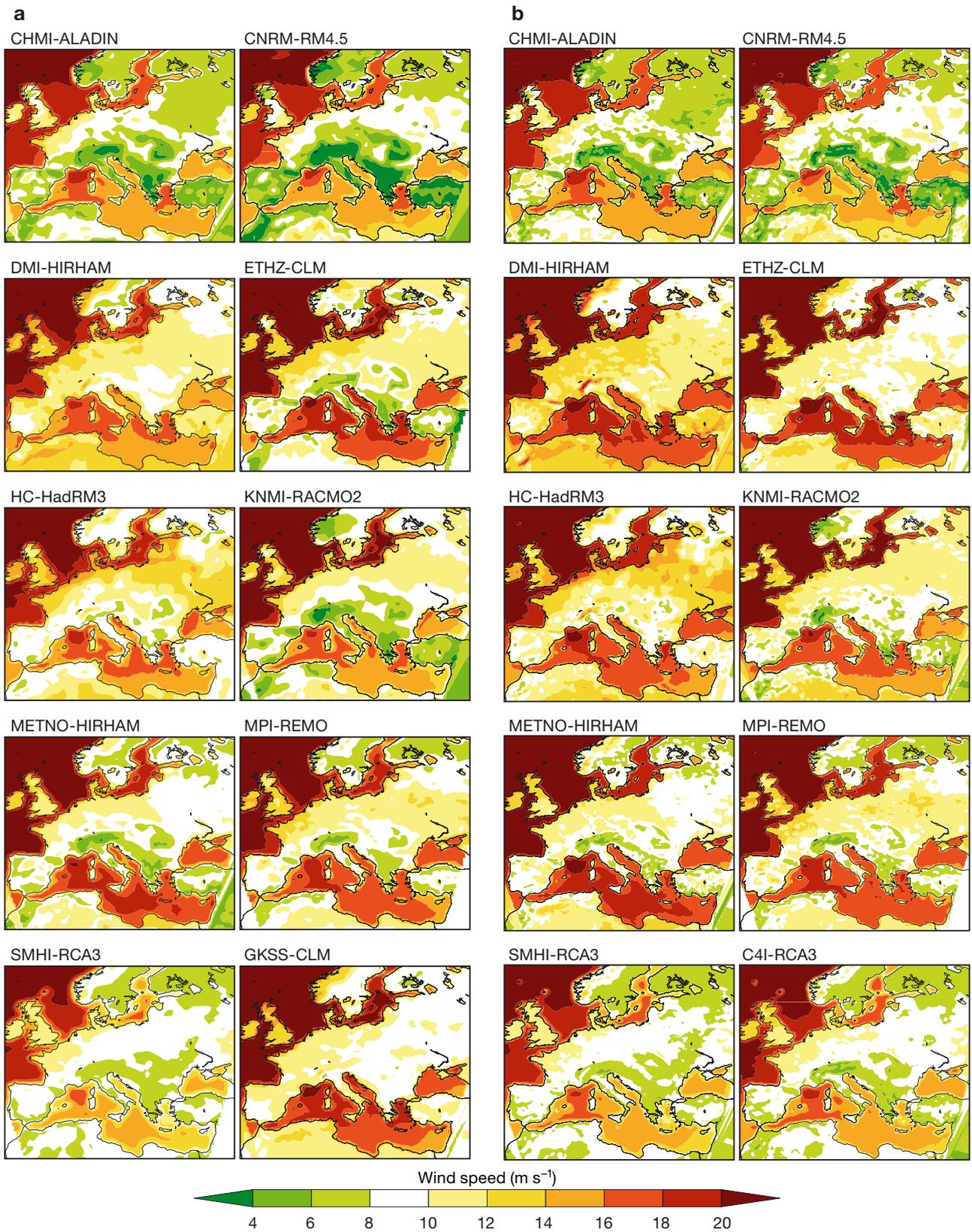


Fig. 3. 98th percentile of daily maximum wind speed (WIMAX) in the regional climate model (RCM) simulations driven by ERA40. RCMs at (a) 50 km and (b) 25 km resolution

ences in the spatial patterns, especially in regions with complex orography such as the Alps or the Scandinavian Mountains. Whereas for these mountainous regions some models and also ERA40 produced local minima of wind speed, other models (in particular DMI-HIRHAM, HC-HadRM3 and partly ETHZ-CLM) featured distinct maxima, irrespective of the resolutions. In the simulations with higher resolution (25 km) the wind maps revealed finer spatial structures, and some additional models featured maxima over mountainous regions (in particular over the Scandinavian Mountains; MPI-REMO, C4I-RCA3, SMHI-RCA3, METNO-HIRHAM). This phenomenon may be influenced by different sub-gridscale roughness parameterisations in the individual models; however, a detailed validation of these differences was beyond the scope of this study.

Considering GUST, the models reproduced the typical pattern found for WIMAX, but the absolute speed values were higher (not shown). Even regarding the model-specific pattern over complex orography, characteristics described above for WIMAX were also found for daily maximum wind gust values. For example, KNMI-RACMO2 revealed local minima over the Alps and the Scandinavian Mountains at both resolutions, whereas in other models (e.g. ETHZ-CLM, C4I-RCA3, SMHI-RCA3), local maxima were found in these regions. Della-Marta et al. (2009) documented that ERA40 reanalysis often yields unrealistic gust wind speeds over complex terrain.

The described differences in the simulation of wind speeds are incorporated by the loss model's approach, avoiding an absolute threshold. Applying a relative threshold (like the 98th percentile in this study) takes into account individual characteristics of the different models and also the local wind climatology.

3.2. Validation with loss data for Germany

Fig. 4 shows the time series of annual loss ratios for GDV data and loss model results based on ERA40 and the 50/25 km resolution ensembles of RCM simulations for the 1970–2000 period. Annual accumulated loss ratios for Germany derived from ERA40 reanalysis data revealed correlations with the GDV data of 0.86 for WIMAX and 0.89 considering gust calculations. The correlations for all individual RCM simulations versus GDV loss ratios are shown in Fig. 5. In both the 50 and 25 km simulations, METNO-HIRHAM yielded the best accordance with GDV loss ratios, with correlations of 0.87 (50 km) and 0.84 (25 km). However, most of the models showed reasonable agreement with observed losses at both resolutions. For the simulations at 25 km (50 km) resolution, correlations were higher than 0.7 for 7 (9) of the 10 models. A systematic statement on a preferable resolution cannot be given. In total, 9 models were run at both resolutions; for 2 of them, the correlation was higher at 50 km; for 4 models, the correlation was higher at 25 km; and for 3 models, correlations at both resolutions were roughly at the same level. The weakest performance was found for HC-HadRM3, revealing correlation values of 0.19 (50 km) and 0.42 (25 km). For HC-HadRM3 and C4I-RCA3, larger domain sizes were used, extending farther westwards over the Atlantic. This may at least partly contribute to the weaker correlation values for these models, as cyclones develop over a longer distance inside the RCM domain and may thus differ more from the (observation-based) reanalysis data.

Large deviations in modelled losses, in comparison to observations, were found for the years 1972 and 1993. The losses in 1972 were mainly caused by a

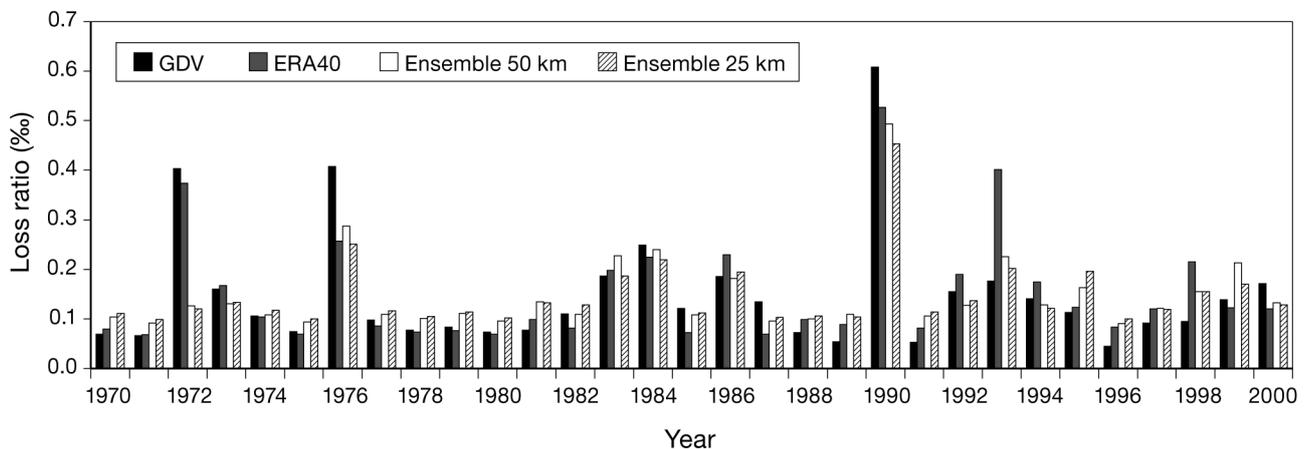


Fig. 4. Time series of annual loss ratios for German Insurance Association (GDV) data and as calculated from ERA40 reanalysis and the reanalysis-driven regional climate model (RCM) simulations. For the RCMs, the ensemble means of the simulations are shown at 50 km resolution (white bars) and 25 km (hatched bars)

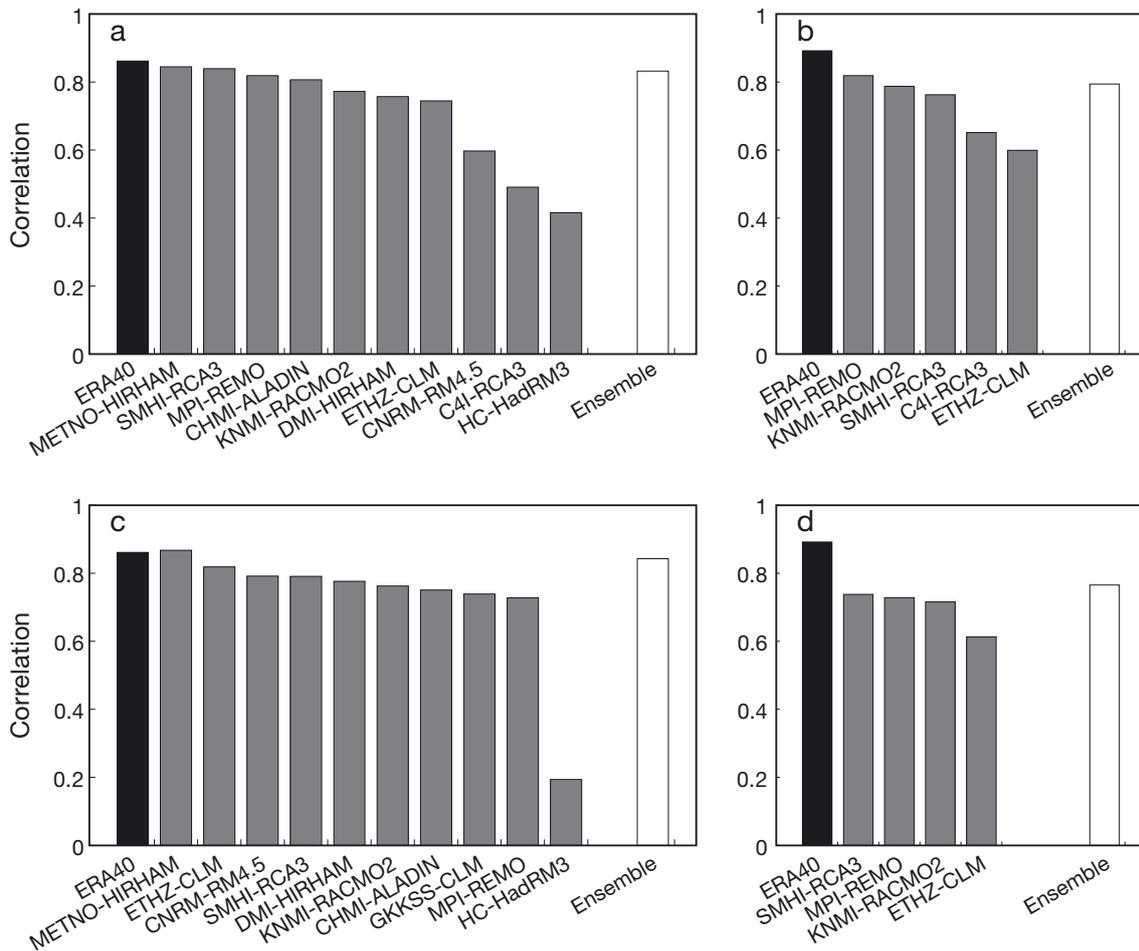


Fig. 5. Correlations of annual loss ratios for ERA40 (black) and the ERA40-driven regional climate model (RCM) simulations (grey) compared to GDV losses; in each figure, the RCMs are ordered according to the correlation value, from highest values (left) to lowest (right). The white bars represent the correlations of the ensemble means of annual losses, calculated from all individual RCMs in each group. (a) 25 km resolution daily maximum wind speed (WIMAX); (b) 25 km resolution daily maxima of modelled gusts (GUST); (c) 50 km WIMAX; (d) 50 km GUST

strong storm event on 13 November, the so-called ‘Lower Saxony Storm’ (Cappel & Emmerich 1975). None of the RCMs revealed intensive losses in Germany for this event, although it seems to be well captured by the ERA40 reanalysis. Thus, the correlation values of losses calculated from the RCMs were negatively affected by failing to capture the severe event in 1972 (cf. Section 3.4). For 1993, losses calculated from ERA40 added up to an unrealistically high amount, while losses based on the RCMs were of the same magnitude as the observed losses. In particular, 2 days in 1993 showed very high losses (14 January: storm ‘Verena’, and 24 January: ‘Barbara’). Losses calculated for these days based on ERA40 were 2 to 3 times higher than for most of the RCMs. Thus, the RCM ensemble agreed considerably better with observed losses, pointing to a benefit from using the RCM ensemble at least for these cases. Some larger differences were also

apparent for the years 1976 and 1990. Here, however, the ERA40 and RCM ensemble results revealed comparable magnitudes, both underestimating the losses in comparison with the insurance data. It is not obvious what caused the deviation in these specific years. There could be particular socio-economic factors (related to the series of storms in 1990, when large areas were affected by losses and specific regions were hit by storms several times within a few weeks) leading to enhanced loss, but an error in the reanalysis data also cannot be excluded. We note that a number of RCMs revealed loss ratios which approximately agreed with the insurance data. A more thorough investigation of the reasons for the deviations is necessary for these years, but is beyond the scope of the current paper.

In all models, the mean annual loss ratio for Germany was around 0.15‰, as in the GDV data (Table 2). This accordance is expected due to the fact that calcu-

Table 2. Mean \pm SD of calculated loss ratios (%) 1970–2000 for Germany in the ERA40 reanalysis, and the ERA40-driven RCM simulations. For comparison, the observed mean \pm SD loss ratio based on the German Insurance Association (GDV) data set is $0.15 \pm 0.12\%$. WIMAX: daily maximum wind speed; GUST: daily maxima of modelled gusts

	WIMAX	GUST
ERA40	0.15 ± 0.11	0.14 ± 0.10
RCMs		
50 km		
METNO-HIRHAM	0.14 ± 0.10	
ETHZ-CLM	0.16 ± 0.11	0.15 ± 0.07
CNRM-RM4.5	0.14 ± 0.09	
SMHI-RCA3	0.15 ± 0.10	0.15 ± 0.09
DMI-HIRHAM	0.15 ± 0.10	
KNMI-RACMO2	0.15 ± 0.10	0.14 ± 0.08
CHMI-Aladin	0.14 ± 0.08	
MPI-M-REMO	0.15 ± 0.10	0.15 ± 0.10
HC-HadRM3	0.15 ± 0.02	
GKSS-CLM	0.15 ± 0.10	
C4I-RCA3		
Ensemble mean	0.15 ± 0.08	0.15 ± 0.09
25 km		
METNO-HIRHAM	0.14 ± 0.10	
ETHZ-CLM	0.14 ± 0.08	0.15 ± 0.08
CNRM-RM4.5	0.14 ± 0.03	
SMHI-RCA3	0.15 ± 0.10	0.14 ± 0.08
DMI-HIRHAM	0.15 ± 0.10	
KNMI-RACMO2	0.15 ± 0.09	0.15 ± 0.10
CHMI-Aladin	0.16 ± 0.11	
MPI-M-REMO	0.14 ± 0.08	0.14 ± 0.07
HC-HadRM3	0.14 ± 0.05	
GKSS-CLM		
C4I-RCA3	0.15 ± 0.06	0.15 ± 0.08
Ensemble mean	0.15 ± 0.08	0.15 ± 0.08

lated raw losses were fitted towards observed loss ratios by linear regression specifically for each individual model. However, the simulated interannual variability in terms of the SD of annual loss ratios was lower compared to observed loss variability in almost all models. While the SD of annual losses in the GDV data was 0.12 and 0.10% in losses calculated from ERA40, it ranged between 0.02 and 0.11 in the RCM simulations. The highest — and almost realistic — variability of 0.11% was found in losses calculated from ETHZ-CLM (50 km) and CHMI-ALADIN (25 km); the lowest variability in losses was revealed by CNRM-RM4.5 (25 km) with 0.03% and HC-HadRM3 (50 km) with 0.02%. Those models yielding a low interannual variability in calculated losses were characterised by a relatively narrow distribution function of wind speeds, particularly in the upper tail. The magnitude of exceedances of the loss threshold was comparatively low in these models, and thus the interannual variability in losses was relatively low.

Ensemble mean loss ratios were calculated for each year as the average of the loss ratios calculated for the

10 different RCMs in each resolution, revealing correlations with GDV data of 0.84 (50 km) and 0.83 (25 km), respectively (Fig. 5a,b). At both resolutions, this value was only slightly below the correlation of the best single model and corresponded to the correlation of losses calculated from large-scale reanalysis. In the ensemble mean of all simulations, there was also a realistic mean loss amount, but the SD was about 30% too low compared with observed loss data (Table 2).

Losses calculated from GUST in the 50 km simulations (Fig. 5d) revealed generally weaker correlations with observed losses compared to losses calculated from WIMAX. The correlation was the same (0.73) for both wind parameters only in MPI-REMO. In the 25 km simulations (Fig. 5b), 2 models showed higher correlations compared to WIMAX (C4I-RCA3 and KNMI-RACMO2) and 2 had lower correlations (SMHI-RCA3 and ETHZ-CLM), and for MPI-REMO, it was again identical. These results indicate that the loss calculations were not fundamentally improved by using gust parameterisation instead of WIMAX. The 2 cases of higher correlations occurred in the simulations at the finer resolution. Still, representatives of all 3 different gust parameterisation schemes (MPI-REMO, KNMI-RACMO2 and SMHI-RCA3) revealed good performance in reproducing observed losses (correlations higher than 0.7) at both resolutions. Considering the MME of those RCMs with gust parameterisation (Fig. 5b,d), the ensemble mean was again close to the correlation of the best single model (and even higher/better than the best model for the 50 km simulations). In the ensemble mean, the correlations with observed insurance loss data were slightly higher in the 25 km simulations compared to the simulations at the coarser resolution.

3.3. Impact of different model combinations on the ensemble mean performance

To demonstrate the effect of the typically arbitrary model selection for MME studies (e.g. due to the availability of model simulations), the performance of the loss calculations was investigated with respect to the models considered in the ensemble. Therefore, sub-ensembles were constructed, consisting only of the 2, 3, 4, ..., 9 single models with the best correlations (not shown). Compared to the ensemble incorporating all models, a slightly increased correlation was achieved, with a maximum of 0.87 in the 25 km simulations, when only the best 4 models were considered, and of 0.86 in the 50 km simulations, considering the best 3 models. For the 25 km simulations, this optimised ensemble mean was even slightly better than the best single model, although it was also obvious that the differences between the correlations were marginal

when the weakest models were left out, compared to when the complete ensemble (i.e. all models) was considered.

The different possibilities of ensemble construction and the resulting performance were examined systematically, without taking individual model performances into account for the ensemble construction. Based on 10 RCMs available in each resolution, there were in total 1023 possible combinations containing 1 to 10 models (10 single models, 45 combinations each with 2 and 8 models, 120 combinations considering 3 or 7 models, 210 considering 4 or 6 models, 252 different combinations of 5 models, 10 combinations of 9 models and 1 ensemble mean containing all 10 models). For all of these combinations, the correlations of the (sub-) ensemble means with GDV loss data were calculated. For each group of sub-ensembles consisting of 1 to 10 models, the spread of correlation values is presented in Fig. 6, also indicating the range between the 10th and the 90th percentile of all solutions, as well as the median. As more models were included in the (sub-) ensemble, the minimum correlation increased, whereas the maximum correlation for all (sub-)ensemble sizes was slightly above the correlation of the complete ensemble with all 10 models included. Best correlation values between 0.85 and 0.87 were reached. The absolutely best performance ($r = 0.872$) was obtained in the 25 km simulations for an ensemble of 3 models (METNO-HIRHAM, MPI-REMO, CHMI-ALADIN, i.e. the single models with best, third and fourth best correlations). In the 50 km simulations, the highest correlation ($r = 0.871$) was found for a different combination of 3 models (METNO-HIRHAM, CNRM-RM4.5, DMI-HIRHAM, again the models with best, third and fourth best correlations). Note that these maximum correlations were not necessarily obtained for the 3 best single models (cf. Fig. 5). Nevertheless, regarding the best 10 correlations for each group of sub-ensembles according to the number of included models (not shown), it becomes obvious that the best 4 single models (METNO-HIRHAM, SMHI-RCA3, MPI-REMO, CHMI-ALADIN for 25 km; METNO-HIRHAM, ETHZ-CLM, CNRM-RM4.5, DMI-HIRHAM for 50 km) were most often among the ‘best ensembles’. Vice versa, for each of the weakest 10 correlations, the 4 weakest single models (ETHZ-CLM, CNRM-RM4.5, C4I-RCA3, HC-HadRM3 for 25 km; MPI-REMO, SMHI-RCA3, HC-HadRM3, GKSS-CLM for 50 km) were included most often. Considering, for example, the best 10 sub-ensembles of the 25 km simulations consisting of 5 models, the best performing single models were included in most of these (METNO-HIRHAM in 9 sub-ensembles, SMHI-RCA3 in 7, MPI-REMO in 9 and CHMI-ALADIN in 10). However, there were also some combinations among these best

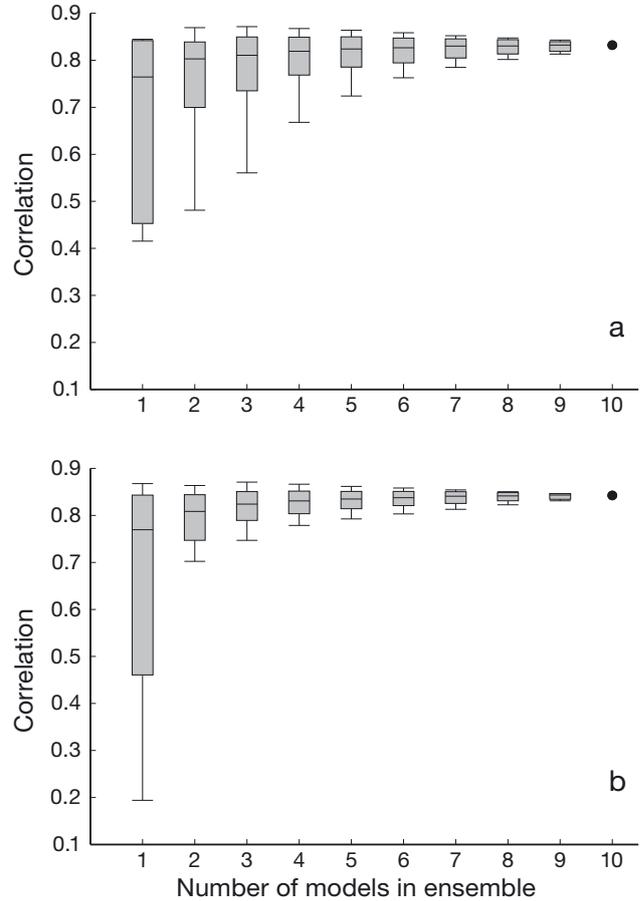


Fig. 6. Correlations of annual losses in Germany calculated from all possible model combinations in comparison with German Insurance Association (GDV) losses. The range of correlation values (y-axis) is indicated for each group of sub-ensembles consisting of 1 up to 10 models (x-axis). The whiskers at the top (bottom) of range show the highest (lowest) correlation value from all sub-ensembles in each group consisting of N models. The boxes indicate the range between the 10th and the 90th percentile of all correlations; the black line in the centre shows the median. The correlation for the ensemble of all 10 models (equivalent to the white bars in Fig. 5a,c) is presented by the black dots. (a) 25 km resolution daily maximum wind speed (WIMAX); (b) 50 km resolution WIMAX

sub-ensembles in which the weakest models were included (ETHZ-CLM in 4, CNRM-RM4.5 in 6 and C4I-RCA3 in 2). On the other hand, the weakest individual models were included in most of the 10 weakest sub-ensembles, for example, consisting of 5 models (ETHZ-CLM in 7, CNRM-RM4.5 in 8, C4I-RCA3 and HC-HadRM3 in all 10 model combinations). Nevertheless, the best individual models were found as members of some of the weakest sub-ensembles (METNO-HIRHAM in 1, SMHI-RCA3 in 2 and MPI-REMO in 3).

In summary, these results illustrate how multi-model ensembles may reveal a high performance even if weak models are included; they therefore support the use of large ensembles. The spread between the best

and weakest performing model combinations was considerably reduced when more models were included in the ensemble (Fig. 6). The performance could be increased only slightly by using an optimal choice of model combinations containing a subset of the whole ensemble.

3.4. Storm of 13 November 1972 and impacts on the RCM performance measure

On 13 November 1972, Europe was hit by one of the most devastating storm events during the 20th century, leading to immense losses, several fatalities and dozens injured. In Germany, this storm was named the 'Lower Saxony Storm', as the state of Lower Saxony was hit particularly heavily. The synoptic situation was characterised by a fast-moving, explosively developing secondary depression on the southern flank of a strong steering cyclone west of Norway. The secondary low travelled fast from the North Atlantic eastwards, crossing the British Isles and the northern part of Germany where it reached its maximum intensity with a core pressure below 960 hPa. Several weather stations in central Europe registered extraordinarily high wind speeds exceeding 40 m s^{-1} . On the Brocken (mountain station at 1142 m a.s.l.), wind gusts reached 68 m s^{-1} .

All RCMs failed to realistically simulate the explosive development of the small depression inside the European model domain. Although most simulations showed a strong wind field in relation to a (not sufficiently strong) secondary cyclone, none of the simulations revealed significant exceedances of the loss threshold (98th percentile of daily maximum wind speeds). By contrast, the explosive development seemed to be well contained in the ERA40 reanalysis and thus extreme wind speeds were obtained, contributing to high simulated losses. For illustration, the MSLP field and extreme wind field caused by this storm are shown in Fig. 7 for ERA40 and the MPI-REMO simulation at 25 km resolution as a representative for the RCMs. In this exemplary RCM, the position of the cyclone seemed to be simulated realistically. However, the core pressure was more than 10 hPa higher compared to ERA40, leading to a weaker horizontal pressure gradient, and thus the resulting wind speeds were considerably lower.

The failure of all RCMs to realistically simulate this particular storm event seems to document a limitation of dynamical downscaling. We assume that the explosive development of the small cyclone inside the domain cannot be triggered from the domain's lateral boundaries. However, the model with spectral nudging did not reproduce the observed development either; thus, it also failed to simulate high losses in central

Europe. In some simulations, extreme loss potentials were found, affecting eastern Europe (KNMI-RACMO2 at 25 km resolution) or Scandinavia (CNRM-RM4.5 at both resolutions). The DMI-HIRHAM run at 50 km resolution also revealed a region with wind speeds strongly exceeding the loss threshold which shifted northwards, over the North Sea. Thus, explosive development took place in some realisations, although affecting regions other than the one hit by the real storm. In line with the failure of all RCMs to simulate this specific event, it should also be considered that for the early 1970s, the uncertainty in the ERA40 upper troposphere is still relatively large, as e.g. only a few satellite observations are assimilated for this early period (Uppala et al. 2005). It is thus possible that atmospheric conditions favouring the explosive development of this storm were not sufficiently included in the reanalysis data used for driving the RCMs at the lateral boundaries. An analysis of the forecast runs of the ERA40 model started on 12 November 1972 00 UTC and 12 UTC (i.e. 18 and 30 h prior to the maximum intensity in the reanalysis data) also reveals less intensive developments (minimum core pressure +9 and +15 hPa, respectively), suggesting that the extreme growth conditions related to the explosive development were not sufficiently captured in the large-scale data. Owing to ongoing assimilation of observations, the extremely deep pressure values were finally included in the reanalysis data. Hence, the failure of the RCMs related to this particular storm should not necessarily be seen as a shortcoming of the RCMs, but possibly as a shortcoming of the driving data. A similar situation of an explosively developing secondary depression south of a strong steering cyclone on 25 and 26 December 1999 (storm 'Lothar', cf. Ulbrich et al. 2001) was well simulated by most RCMs.

To assess potential benefits of using RCMs for loss calculations apart from the obvious failure for 1 event, the model performance measure (i.e. correlation of simulated losses with insurance records) was recalculated excluding the Lower Saxony storm (Fig. 8). Although only a few days would have to be excluded, owing to the annual resolution of the insurance loss data the complete year 1972 had to be removed. This led to considerably higher correlations with observed losses in all simulations, and a clear benefit from dynamical downscaling became obvious: while the correlations for losses based on ERA40 remained on the same level (0.84 for WIMAX and 0.89 for GUST), the best 3 RCMs in both resolutions revealed correlations higher than 0.9, while the best correlations based on single models were 0.95 (0.94) at 50 km (25 km) resolution. For 9 (7) out of the 10 simulations at 50 km (25 km) resolution, the correlation was then above 0.8. The ensemble mean of all 10 RCMs re-

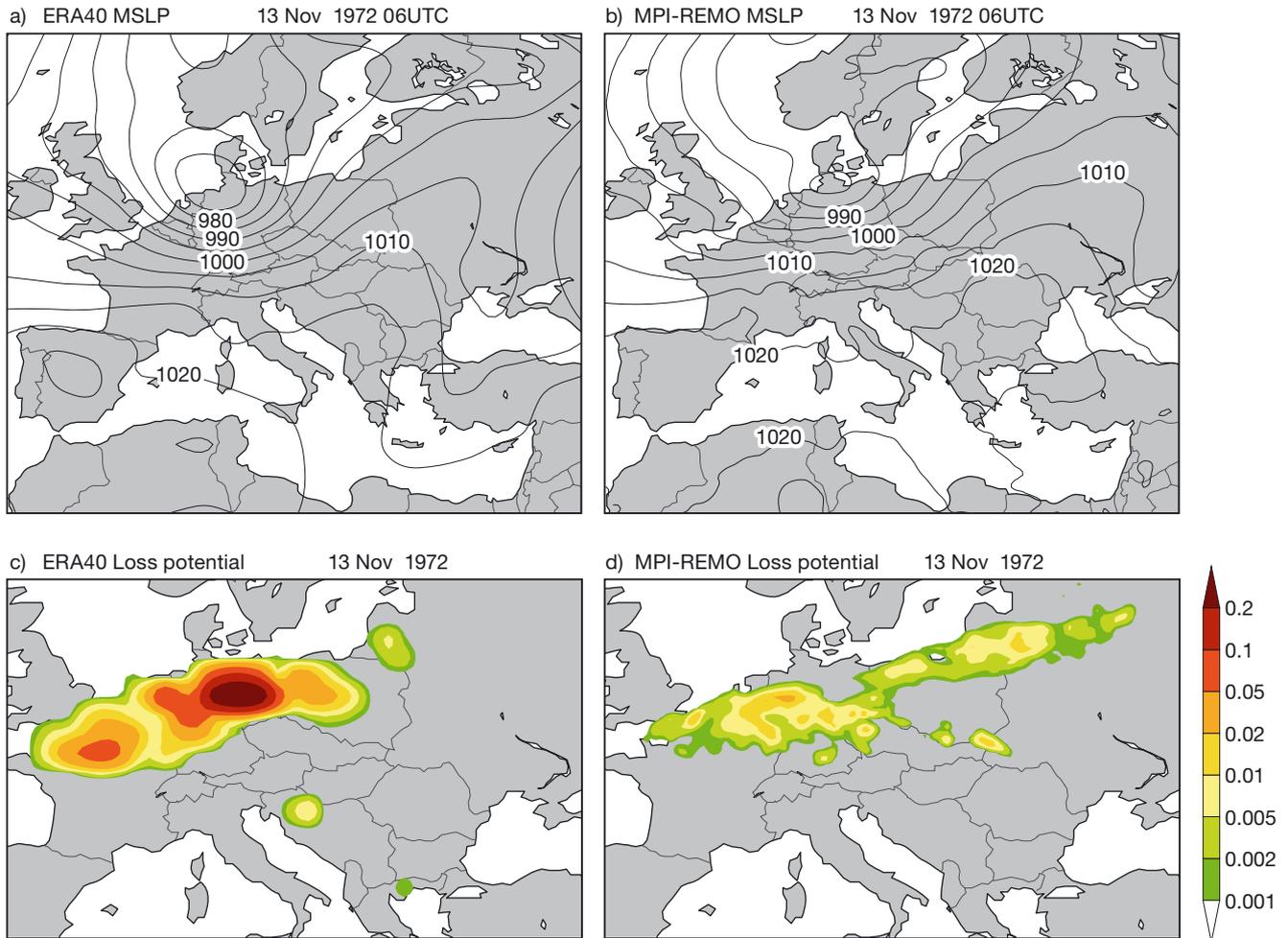


Fig. 7. Situation of the storm on 13 November 1972 as depicted by ERA40 (left) and the MPI-REMO simulation at 25 km resolution (right, exemplary for the reanalysis-driven regional climate model [RCM] simulations). (a) MSLP field for 06:00 UTC on 13 November 1972 in ERA40 (unit: hPa); (b) same as (a), but for MPI-REMO; (c) maximum wind speed on 13 November 1972 in ERA40, displayed as the normalised cubic exceedance of the 98th percentile of daily maximum wind speeds $[(v - v_{98})/v_{98}]^3$; (d) same as (c), but for MPI-REMO

vealed a correlation of 0.94 at both resolutions and thus performed considerably better than losses calculated on the basis of the large-scale reanalysis wind field. Hence, aside from 1 event that was not realistically simulated by all RCMs, a clear improvement of the loss estimates was obtained by dynamical downscaling.

4. SUMMARY, DISCUSSION AND CONCLUSIONS

Spatial patterns of extreme wind speeds and related loss potentials due to severe wind storms were validated in a multi-model ensemble of reanalysis-driven RCM simulations. The typical regional distribution of wind speeds was reproduced by all models, although considerable biases in the simulated speed values were identified for some models. Specific anomalies of

the wind fields were found particularly over complex terrain, where some models revealed local minima and others local maxima of simulated wind speeds. For some models, this phenomenon was found to depend on the spatial resolution, suggesting that part of this effect might be caused by the sub-orography in the grid boxes which are representative of larger areas, including mountains as well as valleys. Even if the grid box size can be considerably reduced by dynamical downscaling, it still covers an area of about 625 km² at the finest resolution considered here. By contrast, observational data for comparison are strongly influenced by the location of the station, which might be close to mountain tops or in valleys and thus affected by accelerating or decelerating effects. Therefore, a proper evaluation of RCM-simulated wind speeds over complex orography requires further systematic investigation and is beyond the scope of this study.

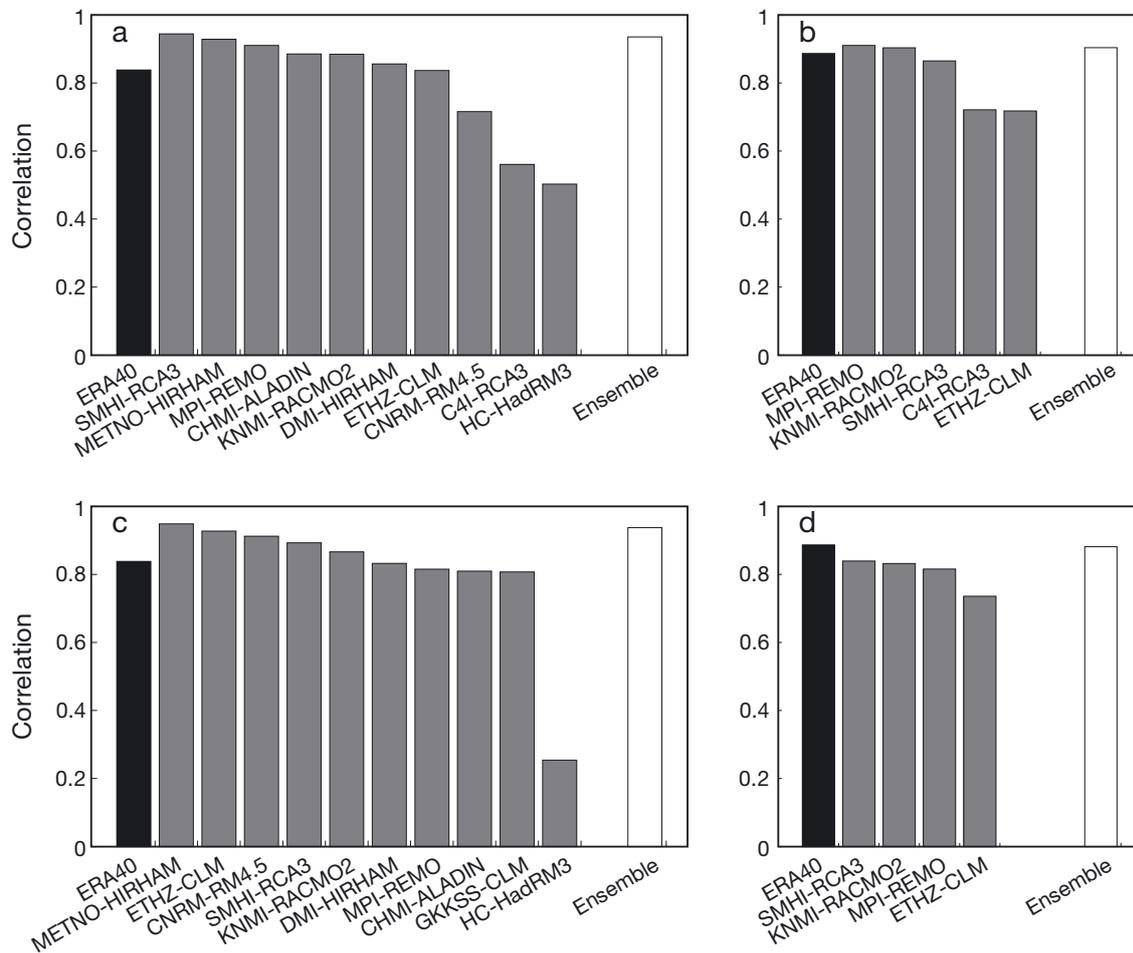


Fig. 8. Correlations between calculated and observed annual losses excluding the year 1972. Other details as in Fig. 5

The applied loss model provided reasonable results (as also shown in previous studies, cf. Leckebusch et al. 2007, Pinto et al. 2007). High correlations of 0.86 (using WIMAX) and 0.89 (using GUST) between observed annual losses for Germany and calculated losses based on the ERA40 reanalysis were obtained. All RCMs failed to realistically simulate one of the most loss-intensive storm events occurring in November 1972. Excluding this particular event from the considerations, loss correlation values above 0.8 were found for most of the reanalysis-driven RCM simulations. For the best performing models, a considerable improvement in loss calculations through dynamical downscaling was documented in comparison to losses calculated on the basis of the large-scale reanalysis data: correlations between calculated and observed annual losses increased up to 0.95. Thus, in most cases, a distinct advantage of applying RCMs became obvious. Considering the complete time series (i.e. including the 1972 event), the RCM benefits were masked by the failure with this specific event. The performance of losses cal-

culated from the RCM simulations was consequently on average 0.1 points lower and thus slightly below the performance of losses calculated on the basis of the reanalysis dataset. The deficiencies in realistically simulating the explosive development related to the 1972 event seemed to be due to shortcomings in the large-scale driving data: atmospheric conditions favouring an explosive cyclone development might not be sufficiently captured by the reanalysis data, as e.g. for this early period only a few satellite observations were available for assimilation. However, it should be noted that the ERA40 reanalysis is subject to data assimilation and is thus regularly 'corrected' towards observations, which accounts for its better performance. By contrast, the RCMs were generally driven at the domain boundaries only and developed their specific dynamics inside the domain. Thus, a better matching with observed losses would not be expected if the driving data are erroneous. However, most of the RCMs generally simulated realistic loss amounts for the majority of years considered and thus featured reason-

able results in terms of loss calculations. Calculated losses for 1993 were unrealistically high based on ERA40, but more realistic for the RCM ensemble, pointing to an advantage from downscaling. The RCM using spectral nudging did not perform better than those without. Additionally, the considered loss measure was integrated with respect to sampling of different events and regions. Thus, the possible benefit from a more realistic representation of regional wind features might be cancelled out. Availability of event-based and regionalised insurance data seems to be desirable for a specific examination of the added value of dynamical downscaling for loss calculations in comparison with observed losses at a higher temporal and spatial resolution, which is the subject of ongoing research.

Considering RCM simulations at 2 different horizontal resolutions, we did not find any evidence for better results with a finer resolution. However, a resolution of 25 km might still not be fine enough to obtain a significant improvement. Nevertheless, it is possible that positive effects might come from simulations at considerably finer resolutions, resolving meso-scale processes.

To our knowledge, this study is the first to investigate the benefits of dynamical downscaling for storm loss estimations. Previous studies have presented reasonable estimates of insured losses for Germany and Great Britain based on reanalysis data at a spatial scale of about 100 to 200 km (Leckebusch et al. 2007, Pinto et al. 2007). Benefits of dynamical downscaling with respect to a more realistic representation of winds over complex terrain were documented e.g. by Goyette et al. (2001), Žagar et al. (2006), Hofherr & Kunz (2010), Kunz et al. (2010). Another study partly analysing the same reanalysis-driven RCM simulations also revealed a wide range of model performance, comparing simulated monthly precipitation with observations (Sánchez et al. 2009). Further, Christensen et al. (2008) documented considerable biases in the ENSEMBLES RCMs compared to observational data with respect to temperature and precipitation. None of these studies investigated and quantified whether there is any benefit from dynamical downscaling in comparison to the performance of the large-scale driving data, as presented here.

The most serious damage is generally thought to be caused by severe gusts. Consequently, it is often assumed that gust wind speeds should be more suitable for the calculation of storm losses than maximum wind speeds without gusts. The results presented here, however, indicate that correlations of annual losses are generally not higher when simulated gusts are considered for the loss calculations (with ERA40 being a major exception). On the one hand, this might reveal

deficits with the gust parameterisations; on the other hand, it supports the approach of using a relative threshold for the calculation of losses, such as the 98th percentile. With this approach, calculations are independent of the absolute simulated speed values, and the loss model is applicable to different climate models, although biases compared to observed absolute values of wind speed may exist. Note that the loss correlations based on daily maximum gusts tended to reveal minimally higher correlations at the finer resolution. This was found in 3 of 4 models with gust parameterisation that were available at both resolutions and also in the ensemble mean. Three different gust parameterisation schemes were used in the RCMs considered here. Although they were of different complexity, and some were based on either more empirical or more physical assumptions, a favourable scheme could not be identified. Representatives of all 3 schemes performed well in reproducing observed losses. Still, a considerable improvement could not be achieved in comparison to loss estimates based on daily maximum wind speeds without gust parameterisations.

The reanalysis-driven RCM simulations offer a chance to demonstrate the profit of combining different models to an MME also for climate simulations. On shorter timescales (such as weather and seasonal prediction), the superiority of MMEs compared to single models was demonstrated by various studies (cf. Hagedorn et al. 2005), whereas on climate timescales, a verification and quantification of the MME advantages is less straightforward, as skill measures are generally not applicable and appropriate data for comparison are often lacking. Here, owing to the large-scale forcing which depicts observed atmospheric developments, the RCMs were expected to reproduce a realistic chronology of events, and the results could be compared with observational time series. We used insurance loss data as an independent proxy for the occurrence of severe storms. Despite identical large-scale forcing, losses calculated from the reanalysis-driven RCMs revealed a large spread regarding their ability to reproduce observed annual losses. Combining the different RCM simulations led to good correlations with observed losses, comparable to the best single model. Thus, excluding (including) the 1972 storm event, high correlation values of about 0.94 (0.85) were obtained, even though models with weak correlations smaller than 0.3 (0.2) were included. Sub-ensembles leaving out the weaker models could lead to slightly increased correlation values, in some cases higher than for the best single model. A systematic investigation of the influence of the MME construction on the performance demonstrated a lower spread of results for large ensembles, mainly caused by an increased minimum performance.

Although the multi-model performance was not significantly better than the best single model, the largest benefit was the consistently better performance of the multi-model ensemble, even when weakly performing models were included. This is particularly important because, in the case of climate simulations, it is generally difficult to measure the performance of the individual models (which would be necessary for obtaining objective quality measures, serving as a criterion for selecting the models to be included in the MME). Further, the best models identified for a specific application or time period do not necessarily have to be the best for a different application or time period (cf. Reichler & Kim 2008). Not only was it difficult to identify 1 best model, but the results further demonstrated that it might also be difficult to identify a best sub-ensemble, as some of the models were included in both the best and the weakest model combinations. Thus, the findings of this study suggest that an MME consisting of all available models should lead to reasonable results. Weighting of models or use of sub-ensembles might still improve the performance. However, in the cases presented here, a selection of the best models yielded only a marginal improvement compared to the inclusion of all models, supporting the use of large ensembles. The results suggest that as more models are included in an ensemble, the obtained consistency of performance increases: for larger ensemble sizes, the spread between the best and the weakest performing model combination becomes considerably lower. The use of a large MME should be a good choice for obtaining stable and reliable results and at the same time also allows for an estimation of the (model) uncertainty.

Loss calculations based on RCM simulations for future climate scenarios, driven by global climate models, will be considered in subsequent studies. An interesting question will be how the change signals from the RCMs compare to the change signals in the driving large-scale model.

Acknowledgements. This study was supported by the ENSEMBLES project, funded by the European Commission's 6th Framework Programme through contract GOCE-CT-2003-505539. We particularly thank the modelling groups for performing the simulations and the Danish Meteorological Institute (DMI) for archiving the model results. We further thank ECMWF, the German Weather Service (DWD) and the German Climate Computing Centre (DKRZ) for ERA40 data use and availability. The German Insurance Association (GDV) is acknowledged for conditioning and providing the loss data used. We are much obliged to P. Lorenz (Max Planck Institute for Meteorology, Hamburg) for fruitful discussions and are also grateful to 3 anonymous reviewers and to editor E. Kjellström, whose constructive comments helped to focus this paper.

LITERATURE CITED

- Brasseur O (2001) Development and application of a physical approach to estimating wind gusts. *Mon Weather Rev* 129: 5–25
- Bubnova R, Hello G, Benard P, Geleyn JF (1995) Integration of the fully elastic equations cast in the hydrostatic-pressure terrain-following coordinate in the framework ARPEGE/ALADIN NWP System. *Mon Weather Rev* 123: 515–535
- Cappel A, Emmerich P (1975) Zwei Wetterkatastrophen des Jahres 1972: der Niedersachsen-Orkan und das Gewitterunwetter von Stuttgart. *Rep German Weather Serv* 135, Offenbach
- CIESIN & CIAT (Center for International Earth Science Information Network Columbia University; and Centro Internacional de Agricultura Tropical) (2005) Gridded population of the world, Version 3 (GPWv3): population density grids. Socioeconomic Data and Applications Center (SEDAC), Columbia University, Palisades, NY. Available at <http://sedac.ciesin.columbia.edu/gpw>
- Christensen JH, Christensen OB, Lopez P, van Meijgaard E, Botzet B (1996) The HIRHAM4: regional atmospheric climate model. *DMI Sci Rep* 96-4
- Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys Res Lett* 35:L20709 doi:10.1029/2008GL035694
- Collins M (2007) Ensembles and probabilities: a new era in the prediction of climate change. *Philos Trans R Soc Lond A* 365:1957–1970
- Della-Marta PM, Mathis H, Frei C, Liniger MA, Kleinn J, Appenzeller C (2009) The return period of wind storms over Europe. *Int J Climatol* 29:437–459
- Dorland C, Tol RSJ, Palutikof J (1999) Vulnerability of the Netherlands and northwest Europe to storm damage under climate change. *Clim Change* 43:513–535
- Fink AH, Brücher T, Ermer E, Krüger A, Pinto JG (2009) The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Nat Hazards Earth Syst Sci* 9: 405–423
- Gibelin AL, Déqué M (2003) Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Clim Dyn* 20:327–339
- Goyette S, Brasseur O, Beniston M (2003) Application of a new wind gust parameterisation: Multiple scale studies performed with the Canadian regional climate model. *J Geophys Res* 108(D13), 4374, doi:10.1029/2002JD002646
- Hagedorn R, Doblas-Reyes FJ, Palmer TN (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting. I. Basic concept. *Tellus Ser A Dyn Meteorol Oceanogr* 57:219–233
- Heneka P, Hofherr T, Ruck B, Kottmeier C (2006) Winter storm risk of residential structures—model development and application to the German state of Baden-Württemberg. *Nat Hazards Earth Syst Sci* 6:721–733
- Hofherr T, Kunz M (2010) Extreme wind climatology of winter storms in Germany. *Clim Res* 41:105–123
- Jacob D (2001) A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorol Atmos Phys* 77:61–73
- Jacob D, Podzun R (1997) Sensitivity studies with the regional climate model REMO. *Meteorol Atmos Phys* 63:119–129
- Jaeger EB, Anders I, Lüthi D, Rockel B, Schär C, Seneviratne SI (2008) Analysis of ERA40-driven CLM simulations for Europe. *Meteorol Z* 17:349–367

- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate-change over Europe using a nested regional-climate model. Part I: assessment of control climate, including sensitivity to location of lateral boundaries. *QJR Meteorol Soc* 121:1413–1449
- Kjellström E, Barring L, Gollvik S, Hansson U and others (2005) A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3). *Reports Meteorology and Climatology*, 108, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
- Klawa M, Ulbrich U (2003) A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Nat Hazards Earth Syst Sci* 3:725–732
- Kunz M, Mohr S, Rauthe M, Lux R, Kottmeier Ch (2010) Assessment of extreme wind speeds from Regional Climate Models. 1. Estimation of return values and their evaluation. *Nat Hazards Earth Syst Sci* 10:907–922 doi:10.5194/nhess-10-907-2010
- Leckebusch GC, Ulbrich U, Fröhlich L, Pinto JG (2007) Property loss potentials for European midlatitude storms in a changing climate. *Geophys Res Lett* 34:L05703 doi:10.1029/2006GL027663
- Lenderink G, van den Hurk B, van Meijgaard E, van Ulden A, Cuijpers H (2003) Simulation of present-day climate in RACMO2: first results and model developments. *KNMI Tech Rep* 252:24
- McCallum E, Norris WJT (1990) The storms of January and February 1990. *Meteorol Mag* 119:201–210
- Munich Re (1993) Winterstürme in Europa. Publication of the Munich Re, Order Number 2041-E-d. Available at www.munichre.com
- Munich Re (1999) Naturkatastrophen in Deutschland: Schadenerfahrungen und Schadenpotentiale. Publication of the Munich Re, Order Number 2798-E-d. Available at www.munichre.com
- Munich Re (2007) Zwischen Hoch und Tief — Wetterrisiken in Mitteleuropa, Edition Wissen. Publication of the Munich Re, Order Number 302-05481. Available at www.munichre.com
- Palmer TN, Räisänen J (2002) Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature* 415:512–514
- Palutikof JP, Skellern AR (1991) Storm severity over Britain, a report to Commercial Union General Insurance. Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich
- Panofsky HA, Tennekes H, Lenschow DH, Wyngaard JC (1977) The characteristics of turbulent velocity components in the surface layer under convective conditions. *Boundary-Layer Meteorol* 11:355–361
- Pinto JG, Fröhlich EL, Leckebusch GC, Ulbrich U (2007) Changing European storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM. *Nat Hazards Earth Syst Sci* 7:165–175
- Räisänen J (2007) How reliable are climate models? *Tellus Ser A Dyn Meteorol Oceanogr* 59:2–29
- Reichler T, Kim J (2008) How well do coupled models simulate today's climate? *Bull Am Meteorol Soc* 89:303–311
- Rockel B, Woth K (2007) Extremes of near-surface wind speed over Europe and their future changes as estimated from an ensemble of RCM simulations. *Clim Change* 81:267–280
- Sánchez E, Romera R, Gaertner MA, Gallardo C, Castro M (2009) A weighting proposal for an ensemble of regional climate models over Europe driven by 1961–2000 ERA40 based on monthly precipitation probability density functions. *Atmos Sci Lett* 10:241–248 doi:10.1002/asl.230
- Schrodin R (ed) (1995) *Dokumentation des EM/DM Systems*. Deutscher Wetterdienst, Offenbach a. Main
- Stappeler J, Dom G, Schättler U, Bitzer HW, Grassmann A, Damrath U, Gregoric G (2003) Mesogamma scale forecasts using the nonhydrostatic model LM. *Meteorol Atmos Phys* 82:75–96
- Swiss Re (2008) *Natur- und Man-made-Katastrophen im Jahr 2007: hohe Schäden in Europa*. Sigma, Nr. 1/2008, Swiss Re publishing, Zürich. Available at www.swissre.com
- Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. *Philos Trans R Soc Lond A* 365:2053–2075 doi:10.1098/rsta.2007.2076
- Ulbrich U, Fink A, Klawa M, Pinto JG (2001) Three extreme storms over Europe in December 1999. *Weather* 56:70–80
- Ulbrich U, Pinto JG, Kupfer H, Leckebusch GC, Spangehl T, Reyers M (2008) Northern Hemisphere storm tracks in an ensemble of IPCC climate change simulations. *J Clim* 21:1669–1679
- Uppala SM, Allberg K, Simmons PW, Andrae AJ and others (2005) The ERA-40 re-analysis. *QJR Meteorol Soc* 131:2961–3012
- von Storch H, Langenberg H, Feser F (2000) A spectral nudging technique for dynamical downscaling purposes. *Mon Weather Rev* 128:3664–3673
- Weigel AP, Liniger MA, Appenzeller C (2008) Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *QJR Meteorol Soc* 134:241–260
- White PW, ECMF Research Department (2003) *IFS documentation CY23r4. Part IV: physical processes (CY23r4)*. Available at www.ecmwf.int/research/ifsdocs/CY23r4/
- Žagar N, Žagar M, Cedilnik J, Gregorič G, Rakovec J (2006) Validation of mesoscale low-level winds obtained by dynamical downscaling of ERA40 over complex terrain. *Tellus Ser A Dyn Meteorol Oceanogr* 58:445–455

Submitted: October 15, 2009; Accepted: June 15, 2010

Proofs received from author(s): November 2, 2010