Dear editor:

We have tried to address all the comments raised by the anonymous reviewers. We hope the changes implemented will make it more suitable for an eventual publication in the journal. Below, we offer a detailed answer to all the comments issued by the reviewers.

Juan José Gómez-Navarro (on behalf of all co-authors)

Detailed response to anonymous reviewer #1

The present work assesses the ability of a mesoscale model (WRF) on reproducing the wind over Switzerland during 24 major wind storms. The paper is focused on two different questions. The first is identifying the optimum configuration of the model, by testing four different PBL schemes (YSU, YSU, MYJ and ACM2) and three different running modes (free, analysis nudging and spectral nudging). The second part is focused on the important question of the added value, this is, the improvement of the results when using a higher horizontal resolution. Both are relevant questions within the scope of GMD, and especially the second one is nowadays the subject of a lot of interest and debate in the regional modelling community. For the evaluation, the papers uses a comprehensive observational dataset containing more than a hundred station records over a very interesting domain, featuring an extremely complex and mountainous terrain (Switzerland). Thus, in my opinion, it represents a substantial advance in modelling science. The overall methodology and statistics used are robust enough to support the results: using boxplots to represent the sample uncertainty is a good decision. Furthermore, using an open source model as WRF does guarantee the reproducibility of the results. The writing is also very good, with precise explanations and almost no typos. I think that this is a pretty good manuscript and I only have minor comments. Consequently, I support publication after minor review.*

We thank the anonymous reviewer for carefully reading the manuscript and his/her positive view on it. We believe his/her very constructive comments will allow us to improve the current version of the manuscript. We have tried to address his/her comments, and the changes implemented in the new version of the manuscript are discussed below.

Figures: The line colours in the figures containing lines (figures 4, 6, 7, 8 and 10) are for me very difficult to distinguish. Some of them are also difficult to print or to see in a projector. I recommend using a different colour scheme, for example the Paired scheme in the qualitative schemes in colorbrewer2.org is a good choice.

We have changed the colour scheme according to this suggestion, and we have also increased the thickness of the lines in order to increase its readbility. We hope the new scheme improves the quality of the figures.

Page 5438. Lines 11-12: I would say "the lack of representation of the unresolved topography", rather than the unresolved topography itself.

The text has been changed according to the reviewer's suggestion.

Page 5440. Line 24: Some of the regions studied in the cited papers could be categorized as "complex topography", it is even mentioned in the title of one of them (Jimenez et al. 2010). But of course nothing like Switzerland. Maybe replace by "not focus on areas with an extremely complex topography as Switzerland" or equivalent.

The text has be modified to account for this detail

Page 5442. Lines 18-19: I think that it would be more correct to use a power law instead of linear interpolation. Anyway, the difference is likely small for small heigh differences, as most of them likely are.

This point has been also raised by Reviewer #2. Although we do not discuss in the manuscript, we indeed started using a power law to account for different heights in the observations, and translate them to an equivalent level of 10 m that could be compared to model output. However, later on the authors decided to use the raw observations without further manipulations, and compare this data with the modelled wind at the same height, using the 3-dimensional fields of simulated wind and using a linear interpolation whenever necessary. But as guessed by the reviewer, this change didn't have any appreciable impact in the results, so it was not discussed in the paper. We have in any case clarified this point in the revised version of the manuscript.

Page 5445. Lines 5-6: As I see it, more than a different PBL scheme, it is the YSU scheme itself, with a parameterization of the unresolved terrain implemented in its surface layer scheme. Also, it is not mentioned in the paper so far but, in the scheme that Jimenez and Dudhia (2012) implemented, there is also a correction that effectively removes the roughness when the laplacian of the (resolved) terrain is below -30m. This feature is intended to remove a negative bias found by Jimenez and Dudhia (2012) in the wind speed over summits, but it can be a source of bias in some resolutions and terrains, as shown by García-Díez et al. (2015). This may be have a relevant influence on the results of the present paper.

This explanation was indeed missing in the former version, and has been improved according to the suggestion by the reviewer. This argument is indeed an interesting explanation for some biases reported by our study, and has been used accordingly through the manuscript.

Page 5446. Lines 13-14. Another advantage of the re-forecast running mode is the low computational cost. As WRF scalability falls fast for large number of processors, splitting the simulations in time can save great amounts of run-time if a large number of cores is available. Even the extra spin-up required does not compensate this effect.

This remark has been included in the new version of the manuscript.

Page 5449. Lines 5-8: Looking at all the results, I won't say that YSU's non-local approach is clearly the best. YSU is clearly superior, but it all seems to depend on the proper representation of the surface roughness and the orographic drag. Maybe ACM2 and MYJ approaches could reach YSU* skill with a similar parameterization of the unresolved topography effects.*

This is a fair observation that has been pointed out in the new version of the manuscript.

Page 5450. Lines 5-7. This is interesting. There is also a similar effect of in GarcíaDíez et al (2015), a lower correlation when using Jimenez and Dudhia (2012) correction, that is corrected when removing the laplacian-dependent part of the correction. Though it is very small. Maybe something is going on in the momentum budget?

By the time when writing the first version of this manuscript, the authors were not aware of the Study by Garcia-Diez et al. (2015). It is indeed interesting to see how such subtle but anomalous behaviour is reproduced in very a different region and under different driving conditions. Unfortunately the authors have no a good candidate to explain such an anomaly. We however believe that finding the source of such bias in the PBL parametrization is beyond the scope of the present study, and is thus relegated as a starting point for future studies.

Page 5450. Line 28: Mentioning the label of Lothar in the paper here (S13) makes easier find it in the figure.

This change has been implemented.

Page 5451. Line 7: If I get it right, this means that the WRF winds are too ageostrophic, which could be related to too much roughness. But looking at the wind speeds it seems the opposite, there *is a positive bias. I wonder what is going on in the model.*

This concern was also raised by Reviewer #2, as it is an intriguing, yet systematic error. Unfortunately we do not have a satisfactory explanation for such a behaviour. As argued by the reviewer, the overestimation of wind does not seem to be related to wind direction. It is a bit difficult to find a satisfactory physical reason for such a bias, since the limited number of cases, and its exceptional nature (wind storms), precludes a reliable estimation of the wind roses in the simulations that can be compared to the observations. Such an assessment could be carried out once a long and continuous run is available, something that is being carried out by the authors of this manuscript, and thus is planned as a matter for future research.

The most likely candidate is, in the opinion of the authors, the representativity error. The 2 km resolution is not fine enough to resolve the complex topography of the Alps, particularly over valleys, which have a clear potential to channel wind into preferred directions. Thus, the misrepresentation of an observational site in the model can lead to this type of biases.

Page 5454. Lines 9-10: This is not surprising given the cases under study. The strength of the flow crossing the domain during wind storms gives small room for internal variability to develop. I think that if the simulations covered a longer period especially summers, with a weaker large-scale circulation, the difference between free and nudged simulations would be more important.

This is a valid remark that has been included in the description of the results.

Page 5454. Line 28: RMSE is in Fig. 7B, no 7C.

This change has been implemented.

Section 4.3: The role of horizontal resolution. I find the analysis to be good. However, I am going to suggest the citation of a couple more references. Mass et al. (2002) is relevant for explaining the lower correlation found in the 2 km simulations respect to the 6 km. They explain how the RMSE can be lower in higher resolution simulations, despite increased realism. In this case, the RMSE is lower in the 2 km run because of the great reduction of the bias, but the lower correlation found is directly related to the smaller intrinsic predictability of the finest scales, in the mesoscale-gamma and microscale (Lorentz, 1969). García-Díez et al (2015) is very relevant here. In contrast with the present work, these authors do not find added value (or very small) in 9 km resolution simulations over Spain, when comparing with the driving model (GFS). The very different topography, with the extreme mountainous environment of Switzerland, can be an explanation of the different results. I think that this should be discussed in the text.

This is indeed an interesting point of view. We have improved the discussion of the results based on the above mentioned references.

Page 5458. Line 12. "p" missing in topography.

This change has been implemented.

Page 5459. Line 12-14: As noted before, likely the strong large-scale circulation related to the wind storms is making the effect of nudging smaller than it would be for a generic climate run.

This remark has been pointed out according to the reviewer suggestion.

Page 5459. Lines 22-23: Commas after "scheme", and "expected"?

These changes have been implemented.

Detailed response to anonymous reviewer #2

The authors analyze the influence of different WRF settings in the wind simulation over complex terrain during wind storms. This kind of sensitivity experiments have not extensively applied to the simulation of extreme winds. I can see the value of the contents of the article in this direction. More specifically, I find the article informative and useful to 1) optimize the model performance and 2) guide model developments during these extreme conditions.

We appreciate the time devoted by the reviewer to carefully read the manuscript and provide interesting notes and insights that clearly will help to improve it. We discuss below the main changes applied to the text according to his/her comments.

My main concern is related to the analysis of the role of the horizontal resolution. The interaction of the domains seem to be two-way which does not provide a clean comparison of the simulations at different horizontal resolutions. This needs to be addressed before the manuscript could be accepted for publication. More specific comments are provided below.

The reviewer is right in this comment: using two-way simulations precludes the real evaluation of the role of the spatial resolution, since coarser domains are artificially improved by the skill in the inner domains. Indeed we had this into account in the design of the simulations, but unfortunately we failed to explain it properly. The reason is that although simulations were firstly carried out in a two-way set-up, they were repeated for specifically addressing the added value of the spatial resolution, once we realised that it would be an important added value of the paper. Unfortunately, the model set-up description in the paper was not updated to reflect this change. Thus, the results we present in the paper regarding the role of the spatial resolution (Section 4.3) correspond to simulations carried out in one-way configuration, and thus the Fig. 9 already reflects the changes demanded by the reviewer. Obviously made changes in this section to emphasise this important detail of the model set-up and avoid further misunderstandings.

1. Page 5445, line 11. Are you activating the Jimenez and Dudhia scheme in all the domains? Note that the scheme also accounts for the speed up of the wind over hills and mountain tops. 2.

We activate it for all domains for consistency reasons. We realise however that the explanation of the YSU* was incomplete (as it was also pointed out by Reviewer #1). Thus, we have improved the description, including the important remark regarding the changes in the YSU scheme to improve the simulation of wind on mountain tops.

2. Page 5455, Section 4.3 The role of horizontal resolution. It is pointed out in the model set up (Section 3.1) and in the caption of Fig. 2 that the exchange of information between domains is two-way. If this is correct, the inner domains provide feedback to their parent domains. This means that the wind at a given resolution contains information of the wind at higher resolutions. The comparison presented does not show a clean analysis of the role of horizontal resolution. The best way to analyze the impact of horizontal resolution is to run WRF with the domains exchanging information in one-way (from the mother domain to the finer domain). 3.

As pointed out above, this is indeed the way we carried out the simulations, although it was not properly described in the manuscript. We have taken the measures to minimise the risk of further misunderstandings.

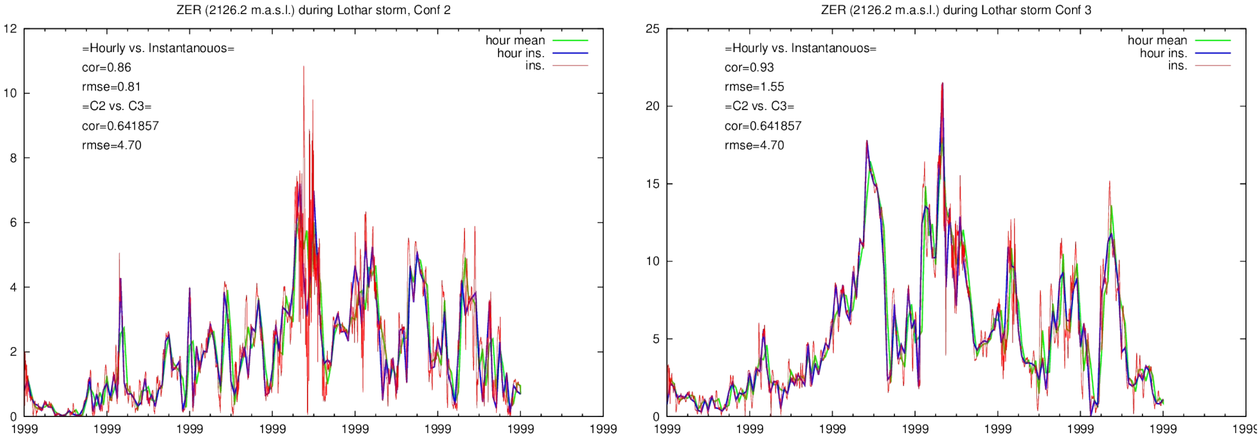
3. Page 5442, lines 10-12: Are you using hourly averages to compare with 10-min obs? Why not average hourly the observations or use instantaneous WRF outputs to compare with the 10-min observations?

No, this is not what we do, and we apologise for the bad explanation that lead to this misunderstanding. We have tried to reformulate the explanation in the new version of the manuscript.

The observations are available at 10-min resolution in some cases. However, Meteoswiss provides though its data portal hourly mean series since the 80's in most stations, so this is the temporal resolution of the observations that we are using in this analysis, directly provided from Meteoswiss without further manipulations from our side. We compare this product to the instantaneous wind obtained every hour from the model. Clearly there is a mismatch between both variables, since we are comparing an instantaneous variable (simulated wind) with an averaged value (observed wind). However, we argue that the error introduced by this mismatch is low, and does not preclude obtaining robust conclusions regarding model performance and the role of different configurations.

We tested this by obtaining the instantaneous wind in a number of locations (WRF has an option that allows us to obtain a series of instantaneous output in every time step for certain locations, whereas this output would be prohibitive for the whole grid). We obtained this output, and compared it with its hourly mean, and with the sub-sampling that consists of the selecting one instantaneous value per hour. We tested this for four locations with different orographic conditions, and all reproduce similar results. The results for the station ZER (Located close to Zermatt) for storm Lothar and two model configurations are shown as an example in Figure 1 in this document. The differences between the hourly mean (green line) and the hourly instantaneous (blue line) are negligible compared to the differences between the two model configurations. The instantaneous wind speed follows the other two on average, and although it exhibits a larger variability (as expected for having higher temporal resolution), it does not play a role in the comparison, since we only analyse hourly resolution observations (comparable to the green and blue series). Thus, this figure demonstrates how the error introduced by not using hourly averaged values of simulated wind speed is not an obstacle to disentangle the differences produced by different model configurations, so the results of the paper do not become compromised by such an approximation.

Figure



1: Simulated time series of wind speed at the closest grid point to the ZER station for stormLlothar. The series indicate instantaneous (red), sampled hourly instantaneous (blue) and hourly mean (green) values of wind speed in that location. The figure in the left (right) corresponds to the model configuration C2 (C3) in the manuscript.

4. Page 5442, lines 17-18: It is mentioned that the measurement height of the sensors differs and that the modeled winds are linearly interpolated to the height of the sensors. The authors should provide a description of the height of the wind sensors and what is the location of the lowest model levels to have an idea of how accurate is the interpolation. How accurate is a linear interpolation *at sensors located very close to the ground?*

This has been discussed above regarding a comment by Reviewer #1. The height of the sensors is not homogeneous, but the number of observations whose height is not 10 meters is small (below 10%), and where it is not, the heights differ up to 60 metres at most. The first three eta levels are located on average at 1.3, 54.37 and 130.78 meters above terrain. Thus, the vertical interpolation is in all cases located between these first three levels, and the effect of these deviations is rather small. In contrast to the linear interpolation used in the manuscript, we also corrected the observations with a power law method. The results of the model performance are similar. This point was however not discussed in the former version of the manuscript, so in the new version we have tried to clarify this aspect.

5. Page 5452, line 18. The negative bias in the wind directions is systematic. Do the authors have an explanation for this?

As argued in response to reviewer #1, “The most likely candidate is, in the opinion of the authors, the representativity error. The 2 km resolution is not enough to solve the complex topography of the Alps, particularly over valleys, which have a clear potential to channel wind into preferred directions. Thus, the misrepresentation of an observational site in the model would introduce systematic errors in the topography of the simulation that potentially can lead to this type of persistent errors. However, this is just a guess that can hardly be proved or disproved with the available data, so we have decided to not include this discussion in the paper.”

6. ERA-Interim is not properly capitalize several times.

This has been reviewed carefully.

7. It is very difficult to identify the line corresponding to each experiments in Fig. 4-7. The authors should find new style conventions to avoid confusion.

We have addressed this in response to Reviewer #1

8. It is probably a good idea to indicate in the title that the focus is on wind storms. The conclusions of the article are valid for these extreme conditions. Conclusions may differ for other synoptic situations.

We have changed the title to narrow the focus of the paper towards storms.

9. Closely related to the previous comment, the conclusions apply for cases of extreme winds. This should be stated clearer in the conclusions. The best WRF configuration may not be the same for weak wind conditions for instance.

We have changed the conclusions to narrow the focus of the paper towards storms.

Sidlerstrasse 5, 3012 Bern, Switzerland Sidlerstrasse 5, 3012 Bern, Switzerland Brandschenkestrasse 90, 8002 Zurich, Switzerland

Manuscript prepared for Geosci. Model Dev.
with version 2014/07/29 7.12 Copernicus papers of the LaTeX class copernicus.cls.
Date: 13 October 2015

Sensitivity of the WRF model to PBL parametrizations and nesting techniques: Evaluation of ~~surface~~ wind storms over complex terrain

Juan José Gómez-Navarro[1,2], Christoph C. Raible[1,2], and Silke Dierer[3]
[1]Climate and Environmental Physics, Physics Institute, University of Bern
[2]Oeschger Centre for Climate Change Research
[3]Axis RE
Correspondence to: gomez@climate.unibe.ch

Abstract. Simulating surface wind over complex terrain is a challenge in regional climate modelling. Therefore, this study aims at identifying a setup of the WRF model that minimizes systematic errors of surface winds in hindcast simulations. Major factors of the model configuration are tested to find a suitable setup: the horizontal resolution, the PBL parameterization scheme and the way WRF is nested to the driving dataset. Hence, a number of sensitivity simulations at a spatial resolution of 2 km are carried out and compared to observations. Given the importance of wind storms, the analysis is based on case studies of 24 historical wind storms that caused great economic damage in Switzerland. Each of these events is downscaled using eight different model setups, but sharing the same driving dataset. The results show that the lack of representation of the unresolved topography leads to a general overestimation of wind speed in WRF. However, this bias can be substantially reduced by using a PBL scheme that explicitly considers the effects of non-resolved topography, which also improves the spatial structure of wind speed over Switzerland. The wind direction, although generally well reproduced, is not very sensitive to the PBL scheme. Further sensitivity tests include four types of nesting methods: nesting only at the boundaries of the outermost domain, analysis and spectral nudging, and the so-called re-forecast method, where the simulation is frequently restarted. These simulations show that restricting the freedom of the model to develop large-scale disturbances slightly increases the temporal agreement with the observations, at the same time that it further reduces the overestimation of wind speed, especially for maximum wind peaks. The model ~~skill~~ performance is also evaluated in the outermost domains, where the resolution is coarser. The results demonstrate the important role of horizontal resolution, where the step from 6 km to 2 km significantly improves model performance. In summary, the combination of a grid size of 2 km, the

non-local PBL scheme modified to explicitly account for non-resolved orography, as well as analysis or spectral nudging, is a superior combination when dynamical downscaling is aimed at reproducing real wind fields.

1 Introduction

Prominent features of the North Atlantic and European climate are cyclonic disturbances, which may be intensified and lead to severe storms (von Storch and Weisse, 2008). Several severe wind storms have hit Central Europe during the last decades (Schiesser et al., 1997; Etienne et al., 2013). These situations, although rare, produce considerable economical cost and are listed as an important natural hazard in Europe (Beniston et al., 2007). Ongoing economic and demographic growth, as well as climate change may imply even stronger impacts in the future, which has raised concerns among reinsurance companies, since isolated events such as storm Lothar in December 1999 caused damages of up to US$ 12 billion (MunichRe, 2001). A better understanding of the mechanisms leading to severe wind storms, and a reliable projection of their characteristics under climate change conditions is important to minimize the impact of such events in contemporary and future societies (Muskulus and Jacob, 2005; Goyette, 2010). However, wind is still not as widely studied as temperature or precipitation (e.g. Schär et al., 2004; Kjellström et al., 2007; Rajczak et al., 2013). For example in areas of complex terrain like Switzerland the main focus of high resolution simulations with respect to wind is on case studies (Goyette, 2008; Etienne et al., 2013). Recently, simulations of about 90 storms over Switzerland are combined to a storm climatology (Stucki et al., 2015).

The fundamental problem regarding surface wind is its intrinsically complex nature, particularly over areas of complex terrain like the Alps (Whiteman, 2000). This complexity precludes its realistic simulation with coarse-resolution models, but also hampers the extrapolation of local observations onto regular grids, which could be used for impact studies. Dynamical downscaling is a common tool that allows bridging the gap between the coarse resolution of Global Circulation Models (GCMs) or reanalysis products and the local terrain characteristics that influence temperature, precipitation or wind (e.g., Kotlarski et al., 2014). Thereby, this method employs Regional Climate Models (RCMs), which, driven at the boundaries by a global dataset, simulate the climate in a limited area domain. This reduces computational costs, which in turn allows implementing simulations with higher spatial resolutions. RCMs driven by GCMs have been used for a variety of applications: from climate change projections (Kjellström et al., 2007; van der Linden P. and Mitchell, 2009; Gómez-Navarro et al., 2010; Jerez et al., 2013; Rajczak et al., 2013; Jacob et al., 2013) to paleoclimatology (Gómez-Navarro et al., 2013, and references therein). Besides, they are used in the so-called hindcast simulations, which blend the reliability of reanalysis products with the high-resolution provided by RCMs. Studies focusing specifically on wind have been one of the applications of such type of simulations ~~(Jiménez et al., 2010; Jerez and Trigo, 2013; Etienne et al., 2013; García-Díez et al., 2013; Menendez et al., 2014; Lorente-Plazas e~~

RCMs, however, contain various sources of uncertainties, like deviations in the driving dataset, numerical approximations, as well as parametrizations of the sub-grid processes. A number of studies in different locations assessed the sensitivity of the model performance due to different model configurations. Dierer et al. (2005) studied the dependency of wind speed on the Planetary Boundary Layer (PBL) parameterizations implemented in the model MM5 as well as the atmospheric stability in different European countries. More recently, García-Díez et al. (2013) focused on the role of different PBL schemes. Other studies have investigated the role of the PBL schemes in the ability of simulating surface wind of typhoons (Kwun et al., 2009), along the coast of the Mediterranean sea (Menendez et al., 2014), or in Southern Spain (Santos-Alamillos et al., 2013). Recently, García-Díez et al. (2015) investigated the added value of dynamical downscaling compared to the driving dataset, and found that over the Iberian Peninsula, 9-km resolved simulations are barely able to add value compared to the driving dataset. However, this result might critically depend on the area under study, as in relatively flat areas the added value of dynamical downscaling becomes less apparent.

~~Although the studies discussed above tackle the problem of the uncertainties in the model configuration regarding wind, they do not focus on areas of very complex topography~~Thus, this study focuses on the performance of very high resolution RCM simulations over an area of extremely complex topography such as the Alpine area. As suggested by Jiménez et al. (2008), in such areas the spatial resolution becomes a major challenge, and the conclusions drawn from coarser resolution simulations ~~can not~~ cannot be generalised without caution. Thus, the present study aims at finding a model setup that minimises systematic errors in hindcast simulations in storms with the purpose of reproducing mean and maximum surface winds in complex terrains. Thereby, potentially important sensitivities of the model setup are explored, which encompass spatial resolution, PBL parameterizations and the use of nudging techniques. The sensitivity with respect to the driving data set is not investigated in order to concentrate on the sensitivity within the model. ~~Geographically the study focuses on the Alpine area.~~ Since the simulations at 2 km grid size require significant computational power, the study is based on a case study approach, rather than continuous simulations. Hence, a total of 24 historical wind storms is simulated for each model setup.

The study is structured as follows: Section 2 presents the reanalysis product used to drive the RCM and the observational network. Section 3 describes the model setup including the different nesting options tested in this study. It further presents the set of sensitivity experiments carried out. The results are discussed in section 4 focusing first on the role of the PBL scheme and the nesting method applied. Then, the role of the horizontal resolution is discussed, including how errors are spatially distributed over different areas of the Alps. Finally, section 5 draws main conclusions.

2 Data

2.1 Reanalysis dataset

The dataset providing the initial and boundary conditions for the RCM is the ERA-Interim reanalysis (Dee et al., 2011). It spans the period from 1979 to today, and is used in its highest resolution of $0.75° \times 0.75°$.The ERA-Interim dataset is generated by running the Integrated Forecast System model (version 2006) of the European Centre for Medium-Range Weather Forecasts (ECMWF). The horizontal resolution of the model is T255 (approximately 80 km). The model has 60 vertical levels up to a pressure level of 0.1 hPa. Observational data are assimilated with a 4-dimensional variational analysis (4D-Var) in a 12-hour analysis window. A number of observational datasets are used ranging from satellite data to surface pressure observations and radiosonde profiles (Dee et al., 2011). However, the assimilation system does not take into account observations of surface wind, which is important to avoid circularity given that this study uses wind observations in the validation part, neither observations of pressure and humidity over high terrain (typically elevations higher than 1500 m).

2.2 Observational data

To evaluate the model's ability in dynamically downscaling wind storms, a reliable set of observations is required. In particular, this is the case in areas of complex terrain, where wind speed and direction can vary within distances of tens of meters. The Swiss Federal Office of Meteorology and Climatology (MeteoSwiss) provides such observations from a dense network of weather stations. ~~This dataset contains 10-min~~ Hourly mean values of wind speed and direction ~~. The model simulations are evaluated hereafter by using hourly means of weather station wind measurements calculated from their 10-minute mean.~~ are used in this analysis. Those are compared to the hourly model output of wind speed and direction at the nearest grid point to each station. Although this introduces a mismatch between hourly averaged and instantaneous values, its influence in the results was investigated, and the results (not shown) show that it is negligible compared to uncertainties attributable to the use of different model configurations.

Some basic data checks are carried out before using the data in the evaluation. Following an approach similar to Lorente-Plazas et al. (2014b), all series are visually inspected. Simple plausibility checks are performed, such as calculating and plotting the running mean and standard deviation to search for anomalies. Stations showing spurious jumps or gaps are discarded from the analysis hereafter. The measurement heights above ground differ in some stations, ranging from 10 to 60 metres. Therefore, the simulated wind is linearly interpolated to the measurement height for the comparison with observations. Additionally, and as a measure to check for the sensitivity of this approximation, the observed wind has been converted to an equivalent height of 10 metres, the level provided by the model output. The differences in the performance metrics obtained with both methods in all the

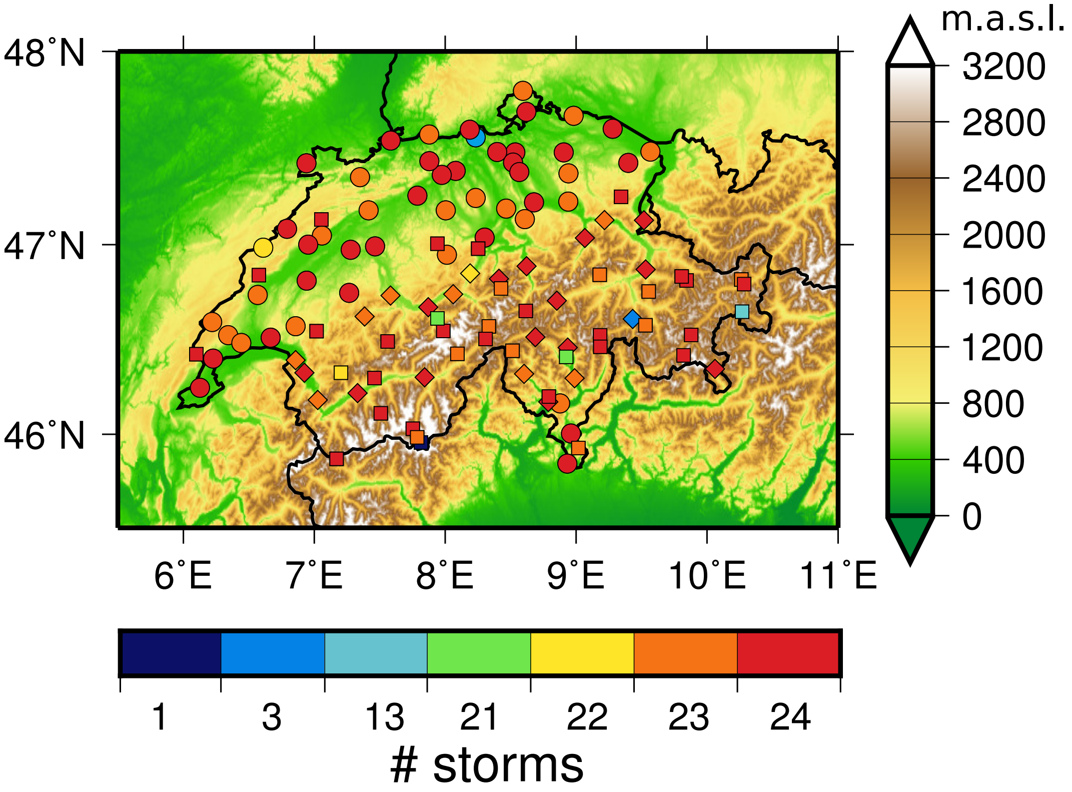


Figure 1. Network of observational sites for wind speed and direction run by MeteoSwiss. The orography of the area is illustrated by the color shading, whereas each symbol indicates the location of an observational site. The filling colour of the symbols indicates the number of storms (Table 1) entering the analysis that are registered by each station. Symbol shape indicates whether the stations are on a plain (circle), a mountain (square) or in a valley (diamond).

analysis below are negligible, and thus are not discussed here for the sake of brevity. After the quality checks, the remaining weather station network still sufficiently covers Switzerland (Fig. 1). We consider in the analysis all the stations that recorded each individual storm. Note that this number increases with time as the observational network has been growing. Thereby, the first storm selected took place in February 1990, and was recorded by a total of 68 stations and the last storm in February 2010 was recorded by 112 stations. 65 stations capture all 24 storms, 36 missed just one storm, whereas only 4 sites captured less than 20 storms. The weather stations cover a wide variety of geographical conditions: plains, valleys and mountainous areas with a minimum (maximum) height of 197 (3580) metres above sea level (m.a.s.l). Thus, this dataset allows evaluating surface wind simulations under different geographical and climatic conditions. Still, as suggested by Gómez-Navarro et al. (2012), the errors and uncertainties contained in the observations shall not be neglected, but need to be kept in mind when using it to draw conclusions about simulation performance.

3 Model and experimental design

3.1 Model setup

The study is based on the Weather Research and Forecasting Model (WRF, version 3.5), aired in September 2013 (Skamarock et al., 2008). WRF is a limited area meteorological model used for weather forecasting and climatic purposes. It employs an Eulerian mass-coordinate solver with a non-hydrostatic approach, and a terrain following eta-coordinate system in the vertical. It is a state-of-the-art mesoscale model used in a variety of studies also for hindcast simulations (Kwun et al., 2009; Jiménez et al., 2010; Awan et al., 2011; Jiménez et al., 2012; García-Díez et al., 2013; Jiménez and Dudhia, 2013; Santos-Alamillos et al., 2013; Menendez et al., 2014, among others).

A first decision in regional climate modelling concerns the selection of the domain to be simulated. Although this selection is susceptible of introducing uncertainties, this study employs just one domain setup, and hence the sensitivity of the performance to the model domain is not investigated here. There are a number of reasons for this. First, there is not much freedom, in the sense that the domain is primary selected according to the area of interest, in this case the Alpine area. The number of domains is conditioned by the resolution of the driving data set and the final resolution of 2 km aimed in our study. So only one is used in all simulations (Fig. 2). It consists of four two-way nested domains with grid size of 54, 18, 6 and 2 km for the domains D1 to D4, respectively. All domain use a Lambert conformal projection, which conserves the spatial distances in both directions. The analysis hereafter evaluates the model ~~skill~~ performance with the focus set on the innermost domain, although the model performance in the coarser domains is also investigated to assess the role of the horizontal resolution. Vertically, WRF does not allow using a varying number of levels in nested domains. Hence, the number of vertical levels has been set to 40 in every domain. This number is similar to the number in recent literature, which ranges between 30 and 46 (Miguez-Macho et al., 2004; Lo et al., 2008; Kwun et al., 2009; Jiménez and Dudhia, 2012; Santos-Alamillos et al., 2013; Etienne et al., 2013). The vertical resolution ensures that several eta levels lie below the PBL height at any time. Naturally, the number of levels within the PBL varies at each grid point according to the PBL height due to different meteorological situations. In the simulations carried out, a minimum (maximum) of 3 (7) levels vertical layers lie within the PBL at any time.

Another source of uncertainty is related to the choice of the physical parameterizations, such as microphysics, convection, radiation and the formation of the PBL, among others (Stensrud, 2007). Since the latter is the parameterization that is most relevant for the surface winds, a number of sensitivity tests are conducted and analysed in order to find the most appropriate PBL scheme (see next section). The other parameterizations remain unchanged in all simulations, i.e., the Microphysics WRF Single Moment 6-class scheme (Hong and Lim, 2006), the Kain-Fritsch scheme of cumulus (Kain, 2004), which is implemented only in the two outermost domains, the Rapid and accurate Ra-

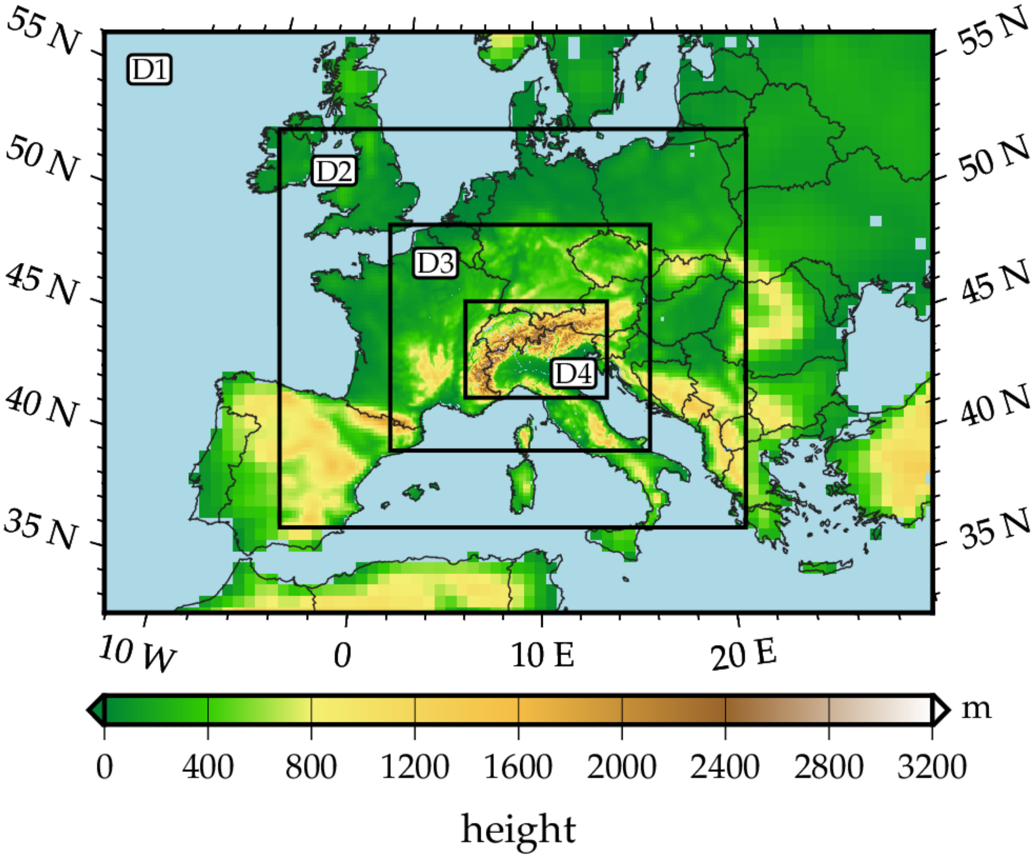


Figure 2. Configuration of the four two-way nested domains. The spatial resolutions are 54, 18, 6 and 2 km, for the domains D1 to D4, respectively. The figure depicts the actual orography and land mask implemented in the simulations.

diative Transfer Model (RRTM) (Mlawer et al., 1997), the short-wave radiation scheme by Dudhia (1989), and the Noah land soil model (Chen and Dudhia, 2001).

3.2 PBL schemes

The PBL plays a major role in simulating surface winds (Stensrud, 2007; Kwun et al., 2009; Santos-Alamillos et al., 2013; Menendez et al., 2014). Nowadays, there are many different approximations to account for the relevant subgrid processes that lead to different PBL formations. In this study four different schemes are implemented, which capture a considerable range of different approaches possible. Similar to García-Díez et al. (2013), we use the fully non-local scheme developed in the Yonsei University (hereafter YSU) (Hong and Lim, 2006), the local closure scheme by Mellor-Yamada-Janjic (MYJ) (Mellor and Yamada, 1982; Janjić, 2001), and the Asymmetric Convective Model 2 (ACM2), which combines local and non-local transport depending on the atmosphere conditions (Pleim, 2007a, b). García-Díez et al. (2013) described the different approaches in detail, therefore not repeated here. The fourth scheme ~~is~~ consists of a subtle modification of the YSU scheme that accounts for the unresolved orography by introducing a correction term in the momentum equation (Jiménez and Dudhia, 2012). This scheme aims at correcting a general problem of WRF with simulating wind, namely its tendency to overestimate wind speed (Cheng and Steenburgh, 2005; Mass and Ovens, 2011; Jiménez and Dudhia, 2012). This is in particular a problem in areas of complex terrain, where topographic features not explicitly considered by the coarse resolution of the model introduce further friction. This scheme also effectively removes the roughness when the Laplacian of the resolved terrain falls below -20 m^{-1}. Although the latter aims at removing the negative bias reported by Jiménez and Dudhia (2012) over mountain tops, it can be a source of bias in other resolutions, and can indeed explain part of the biases found in this study. This scheme is referred hereafter as YSU*.

3.3 Nesting approach

RCMs are nested in a global dataset, which drives the simulation by providing the initial and boundary conditions. Dynamical downscaling is hence mostly an initial value problem in the first days of the simulation, which evolves into a boundary value problem when the initial state has been "forgotten" by the atmosphere. However, how to specify the lateral boundary conditions is a mathematically ill-posed problem, since they become over-specified (Staniforth, 1997). A solution to this problem, widely adopted in state-of-the-art RCMs, consists of newtonially relaxing the driving fields in a buffer zone around the borders of the grid (Davies, 1976). This relaxation damps small scale discrepancies, but does not handle large-scales properly and generates disturbances in the large-scale circulation (Miguez-Macho et al., 2005)). Several methods have been proposed to deal with this problem.

The first approach basically consists of using Newtonian relaxation at the boundaries without any correction inside the domains. This is referred hereafter as "free simulations". In favour of this approach, it is argued that simulations benefit from a better representation and the undisturbed development of regional processes. Another argument is that RCMs are often used to downscale climate change projections or paleo-simulations. Such simulations are performed with relatively coarse GCMs, so that modifications of the large-scale circulation maybe beneficial, as potential biases from the GCMs may be partly corrected by the RCMs.

In case of reanalysis data used at the boundaries, it may be desirable that the RCM simulation stays close to the large-scale situation of the driving data. A first method to achieve this is the so-called reforecast simulation. The method consists of splitting a long simulation in shorter simulation periods of one to few days, running each period separately and finally merging them. This method effectively minimises the impact of the boundaries, transforming the problem into a mostly initial-value problem. The reforecast method is regularly applied (Jiménez and Dudhia, 2012; García-Díez et al., 2013; Menendez et al., 2014, among others), and the increased skill of this method compared to continuous runs has been reported (Lo et al., 2008). A major advantage of this nesting method is its simplicity. Further, it introduces an important computational advantage, since it effectively allows splitting long simulations into a number of independent tranches, providing a natural and very efficient parallelization of the problem. However, it has the undesirable side effect of ~~producing a large number of independent simulations, each of which~~ requiring a spin-up period for each tranche that has to be ~~discarded.~~ run but discarded. Thus, depending on the length of each tranche and the spin-up period, the convenience of this method has to be carefully addressed. In this study we test this approach by simulating every single day independently ~~,~~ with a spin-up period of 12 hours for each run, which results in 1/3 of computational time not exploited.

A more sophisticated method is to force the RCM to follow the driving large-scale conditions. This is implemented by additional terms in the dynamic equations that restrict the degrees of freedom of the simulation. This is the so-called nudging nesting, of which two versions are available. The 3D analysis nudging introduces a Newtonian relaxation term in the prognostic equations of the model, and was first introduced by Charney et al. (1969). This addition corrects some variables by an artificial tendency term based on the difference between the original state produced by the model and the driving dataset (Lo et al. (2008)). WRF provides a number of options that allow selecting which variables and vertical levels should be affected by the correction term. In the current study, horizontal wind, temperature and humidity are nudged in every level but in the boundary layer. The intensity of the correction depends on a nudging factor, which is set here to the default value of $3 \times 10^{-4}\ \mathrm{s}^{-1}$, which is also used in similar studies (Lo et al., 2008).

A variation of this method is spectral nudging, introduced by von Storch et al. (2000). In this approach the variables are Fourier-transformed prior to the nudging. Then, only selected parts of the spectrum are nudged in a similar fashion as the 3D analysis nudging approach, i.e., by introducing a Newtonian relaxation term in the equations. In doing so, the model is forced to mimic the long waves of the driving input data, which contain the large-scale pattern of the atmospheric circulation, whereas it is free to add value in the smaller scales (Miguez-Macho et al., 2005). As in the sensitivity experiments using 3D analysis nudging, the simulations carried out in this study nudge horizontal wind, temperature and humidity only in the levels above the PBL, and with the same nudging factors, $3 \times 10^{-4}\ \mathrm{s}^{-1}$. Unlike 3D analysis nudging, this configuration requires setting the number of waves to be considered in the Fourier analysis, which controls the spatial variability from the input dataset that is being preserved. This number is set to 4 and 2 for the domains D1 and D2, respectively, which correspond to a wave length of about 1000 km. Due to their small size, no nudging is applied in the two innermosts domains.

3.4 Overview of the experiments

This section summarises the set of simulations carried out to investigate the sensitivities of the different settings. Following the approach by Etienne et al. (2013), a total of 24 historical wind storms is

Table 1. List of 24 historical wind storms and the prevailing synoptic flow conditions. This list is adapted from Table 1 in Etienne et al. (2013).

Storm (given name)	Date	Synoptic flow
S01 (Vivian)	1990-02-27	Westerly flow
S02	1993-12-21	Westerly flow
S03	1994-01-28	Westerly flow
S04	1995-01-26	Westerly flow
S05	1995-02-16	Westerly flow
S06	1996-02-13	Westerly flow
S07	1996-11-11	Southerly flow (Foehn)
S08	1997-02-13	Westerly flow
S09	1997-12-17	Southerly flow (Foehn)
S10	1998-01-05	Westerly flow
S11	1998-01-19	Westerly flow
S12	1999-12-12	Westerly flow
S13 (Lothar)	1999-12-26	Westerly flow
S14	2000-02-16	Westerly flow
S15	2000-11-06	Southerly flow (Foehn)
S16	2001-12-14	North-easterly flow (Bise)
S17	2003-01-02	Westerly flow
S18	2004-01-12	Westerly flow
S19	2005-11-23	North-easterly flow (Bise)
S20	2007-01-01	Westerly flow
S21	2007-12-03	Westerly flow
S22 (Klaus)	2009-01-23	Westerly flow
S23	2009-02-10	Westerly flow
S24 (Xynthia)	2010-02-28	Southerly flow (Foehn)

considered (Table 1). The selected storms appear between November and February and are embedded in different synoptic scale flow conditions. Each storm is simulated using eight different model configurations encompassing the sensitivity due to PBL parameterization, nesting method and horizontal resolution (Table 2). Thus, a total of 192 simulations are performed. Each simulation spans 6 days with the corresponding storm in the middle of the simulation, and discarding a spin-up period of 12 hours.

The comparison of observations and simulation results is performed for hourly values at each observational site. For this, the simulation result at the closest grid point to the observational site is selected. Although this can lead to representativeness errors (Jiménez et al., 2010), such errors

Table 2. Summary of the eight model configurations used in the sensitivity studies.

Configuration	PBL Scheme	Nesting
C1	YSU	Free run
C2	YSU*	Free run
C3	MYJ	Free run
C4	ACM2	Free run
C5	YSU	3D Analysis
C6	YSU*	3D Analysis
C7	YSU*	Reforecast
C8	YSU*	Spectral

are systematic in all simulations, and do not play a significant role in the assessment of the relative model performance across model configurations.

4 Results

4.1 The role of the PBL scheme

To evaluate the sensitivity of the model result due to different PBL schemes, the setups C1 to C4 are compared with each other (Table 2). The analysis concentrates on results of the innermost domain. All storms are analysed in an identical way, but for the sake of brevity most of the discussion is based on the results for storm Lothar (storm number 13 in Table 1). Still, similarities and deviating characteristics found in other storms are discussed subsequent to the analysis of Lothar storm.

Fig. 3 shows the situation during the 24 hours around the most mature phase of storm Lothar. This situation was characterized by a intense upper-level zonal jet and strong baroclinicity. The storm formed over the western Atlantic and moved through the Atlantic with moderate amplitude until it reached the French Atlantic coast. There, it experienced an explosive growth as it travelled poleward across the upper-level jet axis (Rivière et al., 2010). The synoptic scale was dominated by a strong north-south gradient in geopotential height that produced strong large-scale winds with a western component.

The surface winds over Switzerland during storm Lothar are presented in Fig. 4 showing the wind speed averaged over 109 weather stations during a 6-day period. The temporal agreement of the sensitivity simulations with the observations is remarkable. The most severe winds peaked on the 26th December 12 UTC, but also the secondary peaks in the time series are generally well captured by all sensitivity simulations. Despite the good timing, an overestimation of wind speed becomes apparent. This overestimation of wind speed is in agreement with the results reported by other studies in different locations and synoptic circulations (Cheng and Steenburgh, 2005; Mass and Ovens, 2011; Jiménez and Dudhia, 2012). The comparison of the sensitivity simulations with

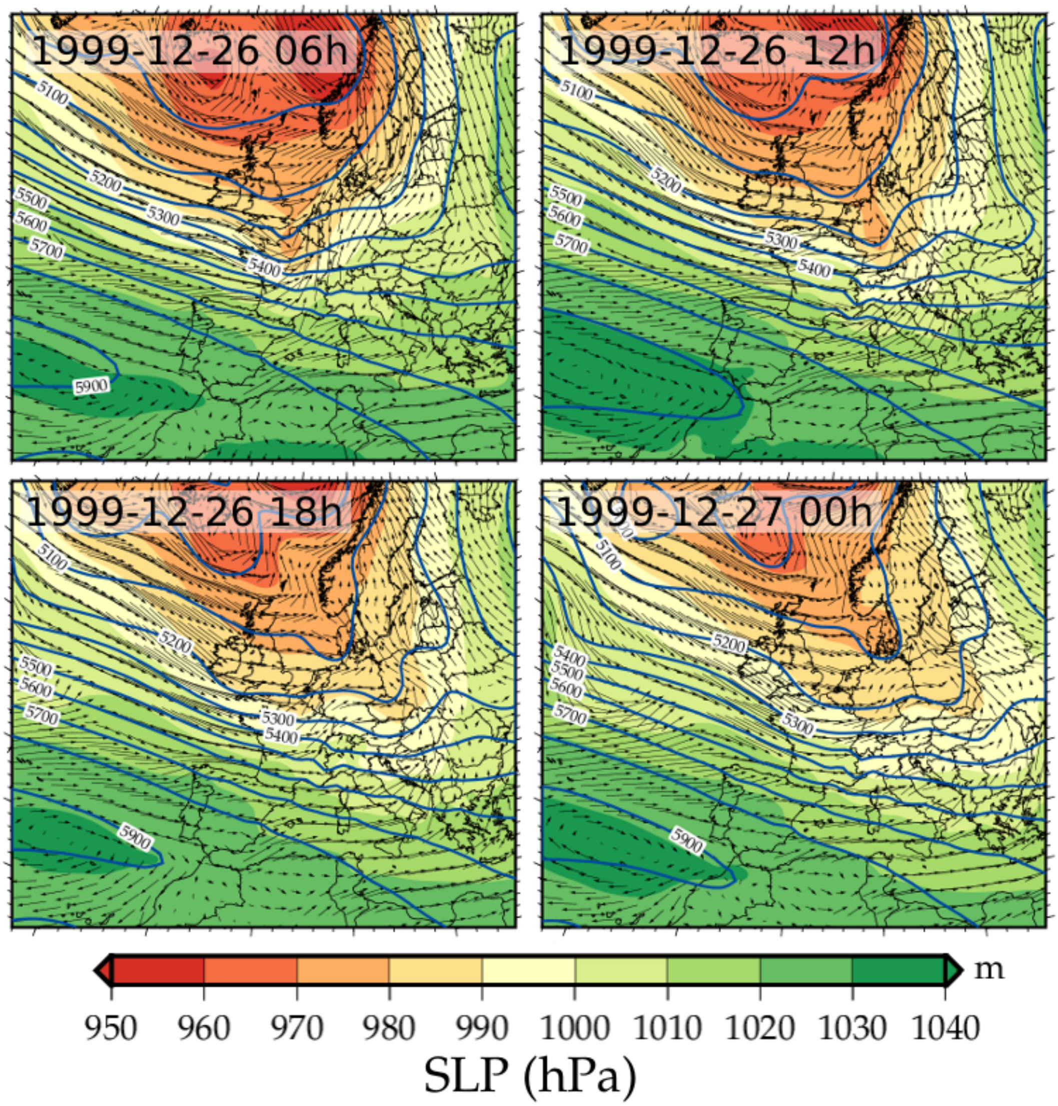


Figure 3. Synoptic situation of the storm Lothar: Colour shading depicts the sea level pressure, whereas blue contours indicate the geopotential height at 500 hPa. The arrows represent 10 m/s wind. The fields represent consecutive snapshots for the period of most severe wind speeds between 26th December, 06 UTC and 27th December, 00 UTC in steps of 6 hours (see Fig. 4). The data fields are obtained from the ~~Era-Interim~~ ERA-Interim reanalysis, the same used to drive the RCM simulations.

different PBL schemes (configurations C1 to C4 in Fig. 4) shows that the YSU* scheme (C2) substantially reduces such overestimation of wind speed. The setup showing the strongest overestimation of surface winds is the fully local scheme, MYJ (C3), followed by the hybrid approach, ACM2 (C4). This ~~suggests~~ indicates that the non-local approach used in ~~both the YSU and YSU* schemes is~~ YSU schemes is slightly more suited to reproduce wind speed over complex terrain~~confirming~~

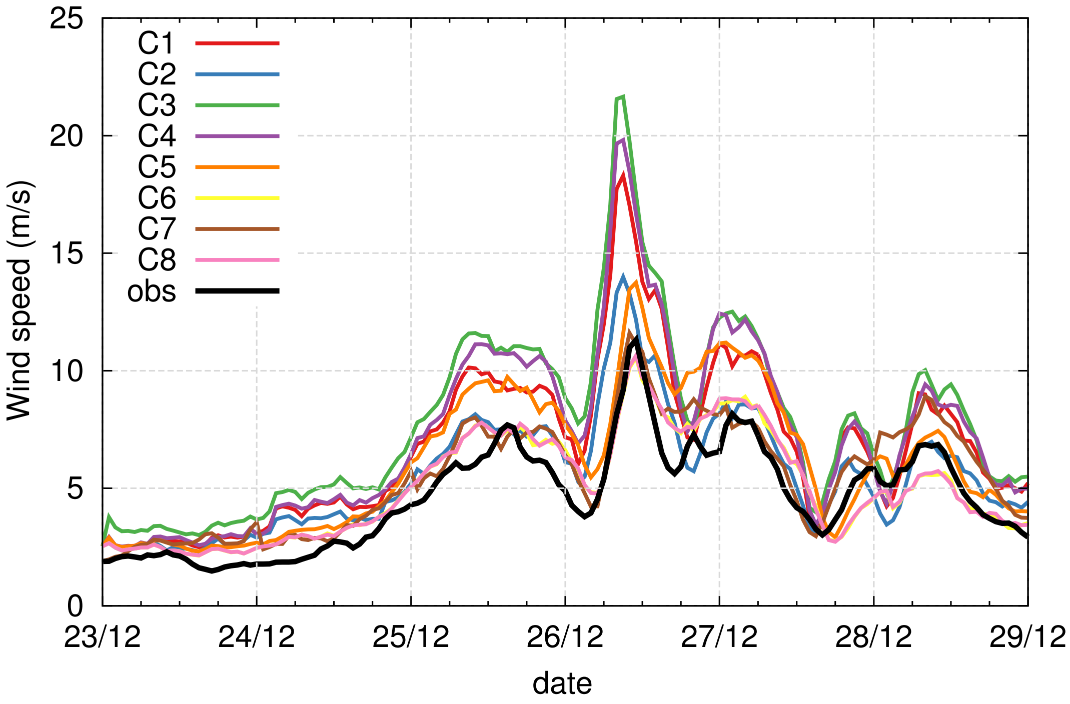


Figure 4. Time series of wind speed for a 6-day period around storm Lothar averaged for 109 stations. Thick black line depicts the series corresponding to the observations, whereas the coloured lines correspond to the simulation results with different model setups (Table 2).

~~findings of Jiménez and Dudhia (2012)~~ . The extra drag factor introduced in the YSU* scheme leads to significantly better simulation of wind speed, and suggests that the inclusion of similar approaches in the ACM2 and MYJ may lead to similar improvements also in these schemes for the representation of surface wind speed.

Fig. 4 provides a first glance of the model-observation comparison. Still, many details are lost by averaging over all stations, in particular the evaluation of the model performance to reproduce the spatial distribution of the most severe winds. Therefore, additional statistics are performed which are presented for storm Lothar in Fig. 5. The boxplots and diamonds show the temporal and the spatial performance, respectively (Fig. 5). Thereby, the boxplots represent the distribution of 109 stations for four statistical metrics that evaluate the temporal performance of the simulation: correlation, Root Mean Square Error (RMSE), bias of the mean wind speed and bias of the maximum wind speed. These four metrics are included since they allow evaluating if the model generally tends to over- or under-estimate wind speed, but also if the simulation is able to mimic the temporal evolution. Note that considering the maximum wind speed is important for scientific questions on extremes on wind speed. To assess the spatial ~~skill~~ performance of the sensitivity simulations (illustrated by diamonds in Fig.5) the wind speed is averaged over the 6 days of each storm simulation at each location. This is done separately for the model and the observations resulting in two spatial patterns of mean wind. Finally, the spatial correlation, spatial RMSE, and spatial biases are calculated. Note that this calculation is not meaningful for maximum wind speed (therefore omitted).

The temporal metrics (shown by boxplots) resemble the findings of the time series in Fig. 4, showing that all model configurations tend to overestimate wind speed. Compared to other PBL schemes, the YSU* scheme (C2) is able to reduce the bias of the mean wind, although it still shows slight positive biases for more than 75% of the stations. Additionally, the RMSE is lower in the sensitivity simulation with the YSU* (C2). The temporal correlation ranges between 0.8 and 0.2, depending on the weather station. The median value is about 0.6 in all configurations. There is a rather low variation between the sensitivity simulations in the temporal correlation, since this metric is dominated by the accuracy of the driving dataset, which is common to all simulations. There is however a lower correlation in the C2 configuration compared to the other configurations. This is a consistent feature across different storms (see discussion below). The ~~skill~~ performance of the maximum wind speed behaves very similar to the mean bias for all sensitivity simulations, although errors become more pronounced: the MYJ scheme exhibits a strong overestimation of the maximum wind speed that is above 10 m/s for 50% of the locations, whereas the YSU* scheme simulates values closer to the observations, although with deviations above 3 m/s in 50% of the locations.

The spatial metrics show that the biases behave similar to the ones of the temporal scale (Fig. 5). This is expected as the spatial bias is identical to the median of the temporal bias if the wind distribution is symmetric. The spatial bias again dominates the spatial RMSE, although in this case the RMSE is significantly lower across all simulations. The sensitivity simulation with the YSU* scheme (C2) shows the lowest spatial RMSE, highlighting the scheme's ability in reducing the overestimation of wind speed. The overall higher spatial correlations than temporal correlations indicate that the model generally is able to simulate the spatial structure of wind independent from the scheme applied. Again, the sensitivity simulation using the YSU* scheme is superior in this metric. Interestingly, the spatial correlation when using this scheme contrasts with the sensitivity simulation with original YSU scheme (C1), as the latter ranks worst among all sensitivity simulations with respect to the PBL scheme. Thus, the spatial metrics show that the YSU* scheme of Jiménez and Dudhia (2012) improves the surface wind simulation by taking into account unresolved orography.

Next, the sensitivity of the model to the PBL scheme is assessed with respect to wind direction. Thereby, the wind rose of the storm Lothar is shown (S13 in Fig. 6). The synoptic situation of storm Lothar shows a very intense westerly flow (Fig. 3). As expected, this situation dominates the wind rose showing primarily south-western directions at the surface. There are additional peaks in the observations in other directions, although due to the pooling process, it could be due to a systematic biases in certain stations (typically valleys, where the 2 km resolution could not be sufficient and lead to representativity errors), rather than a general change in wind direction during the life-time of the storm. Regardless of the PBL scheme, the simulations of storm Lothar are able to capture the main wind direction with a slight systematic bias towards southern directions. The major difference among

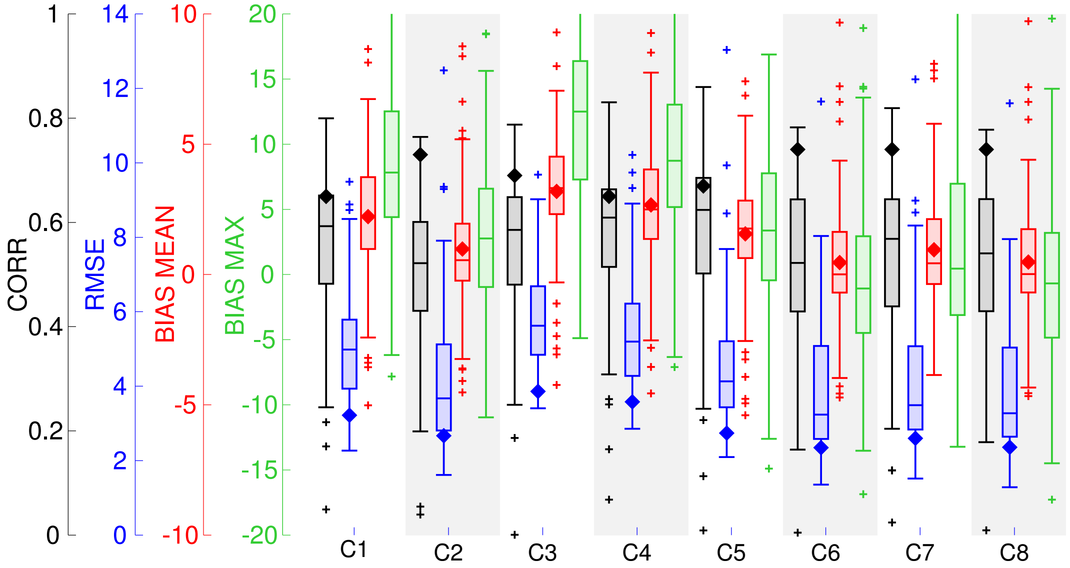


Figure 5. Different ~~skill~~ performance metrics for the comparison of observations and simulation results for storm Lothar. Each column represents the results of one sensitivity simulation in Table 2. The temporal ~~skill~~ performance is illustrated by boxplots which represent the distribution of four metrics over 109 weather stations: correlation (black), RMSE (blue), model-observation mean bias (red) and model-observation maximum bias (green). Four different scales are employed, which match the different colours of the symbols. The diamonds represent the spatial ~~skill~~performance. They are calculated for the mean wind speed of each location. The boxes represent the second and third quartiles, whereas the whiskers extend from the ends of the box to the most distant point whose y value lies within 1.5 times the interquartile range. Outliers are depicted by crosses.

configurations is a slightly less pronounced preferred direction in the YSU* scheme, although it does not contribute to reduce the ~~southern bias~~ bias of a too southern direction. Thus, the simulation of wind direction seems to be mostly insensitive to the PBL scheme selected.

Most of the conclusions drawn from the analysis of the storm Lothar about the PBL schemes are consistent through the various storms simulated. This is illustrated in a comprehensive although summarised way in Fig. 7. Hereby, the "temporal" series show the median temporal correlation (i.e., the centre of the boxplots in Fig. 5) whereas "spatial" series indicate the mean spatial correlation (i.e., the diamonds in Fig. 5). The temporal correlation seems to be mostly insensitive to the choice of PBL scheme (C1 to C4 in Fig. 7a) with the exception of the YSU* scheme (configuration C2), which shows a slightly lower temporal correlation for almost all storms. These slightly lower correlations are in agreement with similar findings by García-Díez et al. (2015) . Although the authors could not find reasons for the reduced temporal correlation, this phenomenon becomes ameliorated when nudging is used (see next section), rendering this caveat less relevant for the sake of the identification a suitable model setup. Still, the reduction of the overestimation of wind speed by this configuration leads to lower temporal RMSE across the storms (Fig. 7b). This improvement becomes especially

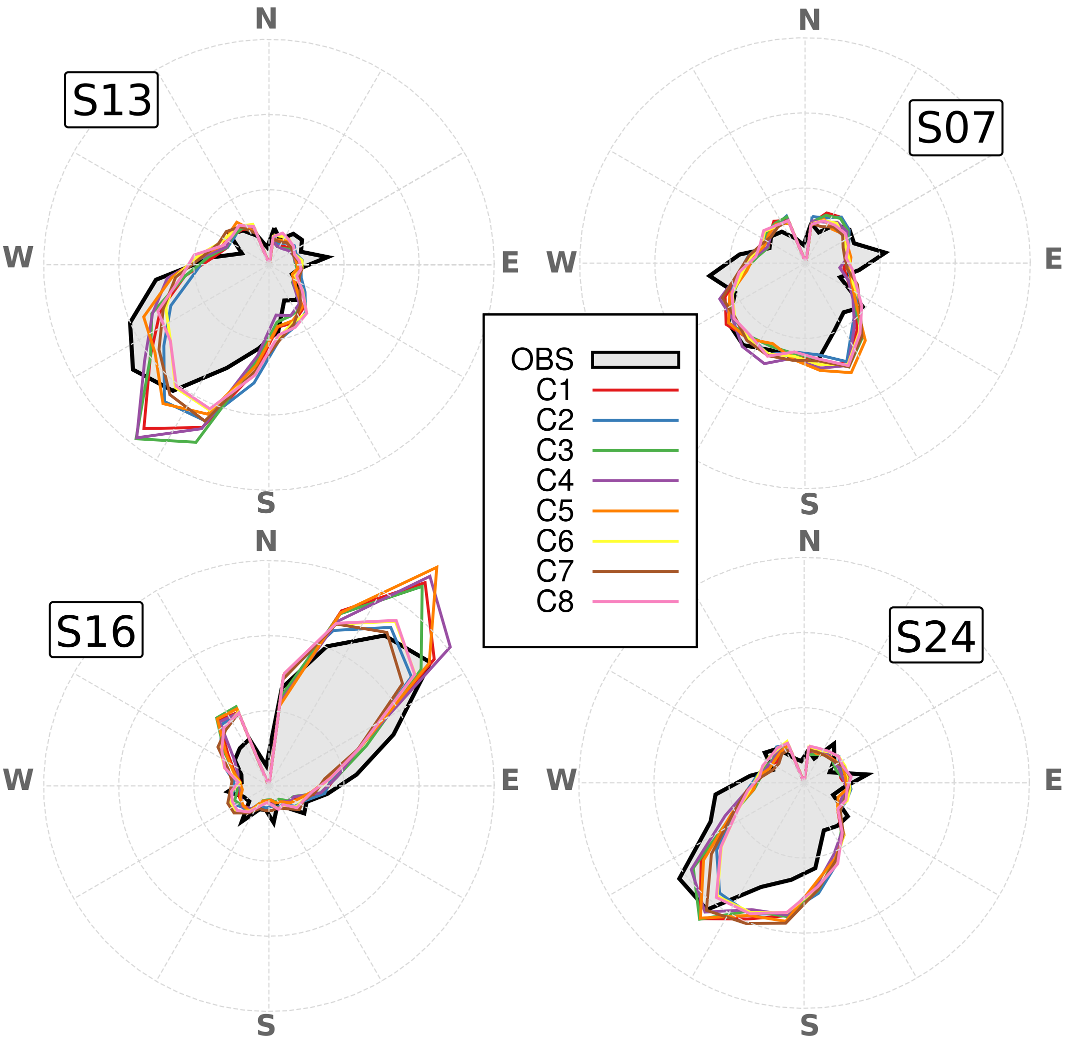


Figure 6. Wind roses corresponding to four different storm cases selected from Table 2. The number of the case is indicated in each panel, and corresponds to storm Lothar (S13), as well as other three storms with different synoptic conditions. The colours correspond to observations and simulations as in Fig. 4. For the calculation of the histograms the hourly wind direction during the entire period registered in each location is pooled, and the number of times that a wind direction lies within each 15° bin is counted and finally normalised.

noticeable in the maximum wind speed bias (Fig. 7c) where the sensitivity simulations with the YSU* scheme show that the bias fluctuates around zero whereas it is significantly larger for the other setups (C1, C3, and C4). In space, the comparison of the PBL schemes demonstrates that the YSU* scheme exhibits systematically higher spatial correlations (Fig. 7d). Thus, the simulations of all storms using the YSU* scheme are able to better allocate the wind speeds at the right locations than the simulations using other PBL schemes. The simulations using the YSU* scheme further

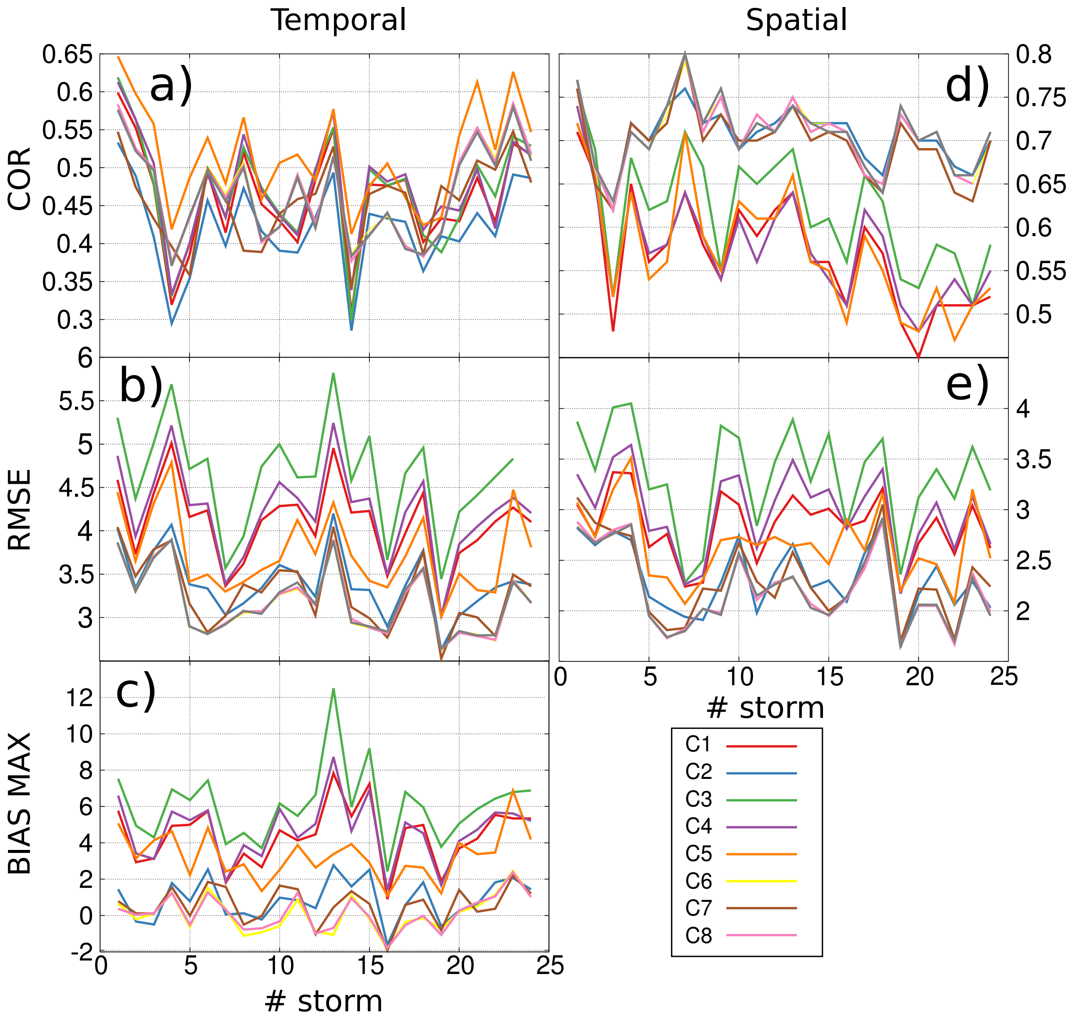


Figure 7. Model performance for the comparison of observed and simulated wind speed across the 24 storms defined in Table 1. Each coloured line corresponds to a model setup. Correlation (panels a and d), RMSE (panels b and e) and bias in the maximum wind speed (panel c) are shown on the temporal (panels a, b and c) and spatial (d and e) domains, respectively. The spatial values correspond directly to the diamonds in Fig. 5, whereas the time statistics show the median value, this is the centre of the boxplots in the same figure.

exhibit lower spatial RMSEs, due to higher correlations and a reduction of the overestimation of wind speed (Fig. 7e).

Wind direction performance across all storms is analysed in a similar fashion. However, this variable has to be treated differently, taking into account the problems associated to its circularity. Thus,

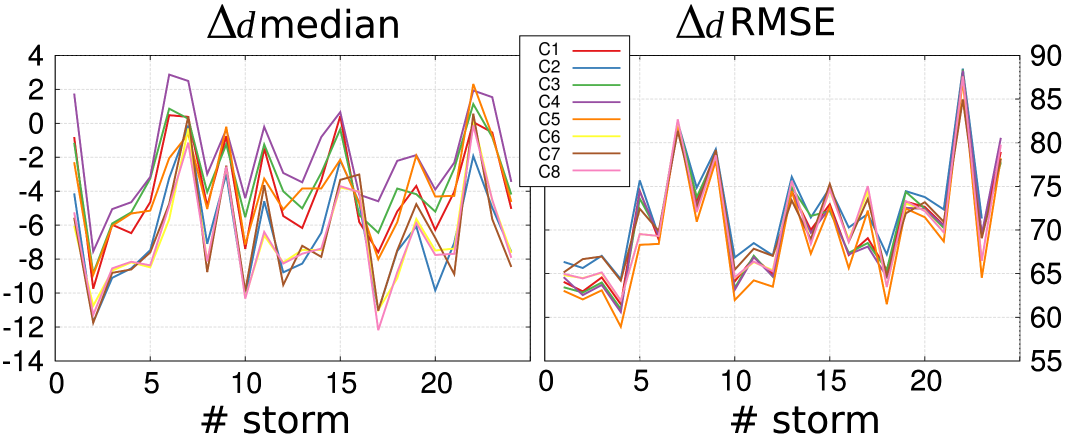


Figure 8. Performance metrics for wind direction for all storms. Left (right) panel shows the model ~~skill~~ performance evaluated ~~through~~ using the median (RMSE) of Δd, as defined in the main text.

similarly to Jiménez and Dudhia (2013) the Δd parameter is calculated:

$$\Delta d = \begin{cases} d_{\mathrm{WRF}} - d_{\mathrm{obs}} & \text{if} \quad |d_{\mathrm{WRF}} - d_{\mathrm{obs}}| \leq 180 \\ d_{\mathrm{WRF}} - d_{\mathrm{obs}} - 360 & \text{if} \quad d_{\mathrm{WRF}} - d_{\mathrm{obs}} > 180 \\ d_{\mathrm{WRF}} - d_{\mathrm{obs}} + 360 & \text{if} \quad d_{\mathrm{WRF}} - d_{\mathrm{obs}} < -180. \end{cases}$$

This definition produces positive (negative) biases when simulated wind direction is orientated clockwise (counter-clockwise) with respect to observations. Once this parameter is calculated for each site in each time step, its distribution is obtained. For this, all values are pooled, so the temporal and spatial details are lost in the discussion hereafter. A RMSE that accounts for the deviations between the simulation and the observations in every location and time step for each storm is derived from the distribution of this bias using

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}(\Delta d_i)^2\right]^{1/2}.$$

The sensitivity simulations show a RMSE of about 70° regardless of the PBL scheme (Fig. 8). Similarly, the median of Δd exhibits a negative bias, again independent from the PBL scheme. For both metrics, the inter-case variation is larger than the variation between the different sensitivity simulations. Thus, this confirms the finding of the storm Lothar that the PBL scheme plays a minor role in reproducing wind direction. Still, it is noteworthy that the C3 and C4 configurations perform better than C1 and C2 as they exhibit lower Δd and lower RMSE. This result is expected, and in good agreement with the findings by Jiménez and Dudhia (2013), who pointed out that the model's ability to reproduce wind directions is inversely related to wind speed.

To assess whether the results of wind direction and the minor role of the PBL scheme may depend on the storm selected, wind roses of three additional storms are shown in Fig. 6. The storm S07 corresponds to a typical Foehn storm. Unlike for the Lothar storm, the wind rose does not show preferred directions, as expected since Foehn storms affect only part of Switzerland. S16 corresponds to a ~~bise~~ Bise storm. In this configuration there is again a clear preferred direction, but is exactly opposite to Lothar. In this storm there is a clear second maximum towards $-30°$. Finally, S24 corresponds to the Xynthia storm in February 2010. This is a west-wind storm, although its particular trajectory when traveling towards Switzerland induces a Foehn-like situation, and thus has been catalogued as such by Etienne et al. (2013). In these examples, all sensitivity simulations show remarkable ~~skill~~ performance in identifying the most dominant wind directions. WRF is clearly able to capture the different nature of these storms, and simulate the surfaces wind regime accordingly. However, none of the four PBL schemes stands out in reproducing the wind direction, resembling the minor role of the PBL scheme and showing that this result is independent from the specific type of the storm.

4.2 The role of the nesting technique

The analysis carried out in the former section indicates that the YSU* scheme is superior compared to the other PBL parameterizations, so this scheme is used in the sensitivity experiments hereafter (see Table 2). The next choice pertains how the RCM is nested to the driving dataset. To assess the sensitivity of nesting approach, the focus is set on the sensitivity simulations C2, C6, C7 and C8, where the free simulations (C2) are compared with analysis nudging (C6), reforecast (C7) and spectral nudging (C8). Fig. 4 illustrates that the nesting techniques further reduce the systematic overestimation of the wind in the case of storm Lothar. For the wind speed maxima, the configurations C6 and C7 better reproduce the intensity and precise timing compared to the C2 configuration.

Fig. 5 presents spatial and temporal performance metrics for the storm Lothar. Spatially, the sensitivity simulations, which include nudging techniques, exhibit slightly higher correlations than the free simulation (C2). In contrast, spatial bias show that the analysis nudging reduces bias compared to C2, whereas the reforecast method increases the bias. The latter is due to the last day of the simulation, where this sensitivity simulation exhibits a strong bias (Fig. 4). For the temporal performance metrics, all nudged simulations (C6-C8) tend to increase the correlation compared to C2, although the improvements are small and rise concerns regarding the robustness of this finding. ~~A~~ Note that the choice of cases could be masking the importance of the nesting technique. The reason is that under wind storms the strength of the flow crossing the domain leaves little freedom for the model to develop deviations from the driving dataset. Thus, the role of the nesting technique could be larger in regular situations not considered by this study. Still, a more clear improvement introduced by nudging is however found for the maximum wind. The original YSU* scheme without nudging shows systematic positive biases in this variable, which are reduced when reforecast, but especially analysis or spectral nudging, are used (Fig. 5).

The analysis of wind direction delivers similar results as in the sensitivity to different PBL schemes (Fig. 6). The role of the nudging approach in correctly simulating the wind direction is minor regard-

Table 3. Representativity error in different model domains. The mean and standard deviation of the horizontal distance s between the station and the closest grid point and height difference between both Δh (both in meters) is shown for the domains D2 to D4. The values are calculated considering all stations shown in Fig. 1.

	$\overline{s}$	$\sigma(s)$	$\overline{\Delta h}$	$\sigma(\Delta h)$
D2	6592	2392	-139.76	660.49
D3	2075	801	-106.97	537.44
D4	747	296	-2.37	322.33

ing the storm Lothar. So, it is not possible to identify any nesting configuration that outperforms the others. Instead, all simulations behave similarly, and the main wind direction seems to be equally reproduced across sensitivity simulations according to the synoptic characteristic of the storm.

As before the analysis is extended to all 24 storms. The mean temporal correlation obtained for different storms is shown in Fig. 7a. This figure illustrates that the temporal agreement is slightly but consistently improved when some nudging is applied, rendering the temporal agreement with the observations comparable to the other schemes, as argued for the storm Lothar. The analysis and spectral nudging (C6 and C8) systematically improve the simulations compared to the free simulation (C2). The reforecast (C7) exhibits the improvement of temporal correlation to be highly dependent on the storm considered. This becomes even more obvious in Fig.7~~e~~ b where the C6 and C8 schemes exhibit lower RMSE than C2, and also generally lower RMSE than C7. Moreover, nudging reduces the bias in maximum wind speed consistently, and makes analysis and spectral nudging equally suitable to improve the maximum wind speed compared to free simulations (C2). Regarding the improvement in wind direction the model performance varies erratically depending on the storm (Fig. 8). The role of the nesting scheme with respect to the median error Δd is even smaller than of the PBL schemes. A very similar result can be drawn from the analysis of the RMSE, although in this case the nesting setups generally exhibit lower RMSE than the free simulation. However, these differences between the schemes are small, so that an identification of a nesting setup that significantly outperforms the others is not possible when considering wind direction.

4.3 The role of horizontal resolution

As argued above, the horizontal resolution has a profound impact on the ability of the model to simulate wind speed. In particular this is the case if the closest grid point of the model to the weather station is used in the analysis. ~~We note~~ Note that this simple approach neglects the fact that the model averages subgrid terrain properties, and leads to so-called representativity errors. It is beyond the scope of this study to assess these errors and to address a method to minimise them, since they introduce systematic biases that only depend on the domain configuration, which is fixed across simulations ~~,~~ and thus play a secondary role in the evaluation of the relative skill of different model configurations. Still, such errors, and the model performance in general, critically depend on the model resolution, so the importance of model resolution and the type of station is discussed in more detail. Note that for the analysis shown in this section, a subset of the simulations had to be repeated to set the nesting configuration to one-way. This allows evaluating the actual performance of the model in coarser domains, since otherwise the two-way approach artificially increases the performance in coarser domains based on the simulation in the inner ones.

The representativity error is quantified by calculating the horizontal distance (s) and difference of height (Δh) between the stations and the closest grid point (identical to the model ~~skill~~ performance assessment above). The mean representativity error over all weather stations as well as the standard deviation are given in Table 3 for the domains D2 to D4 with resolutions of 18 km, 6 km and 2 km, respectively. Obviously the 2 km resolution is closest to the real locations of the observations, with an average distance of 747 m. The horizontal errors become more severe when a coarser resolution is implemented, and reach a mean of about 6.6 km in the 18 km resolution setup. As expected, the height bias is close to zero, but there is a large standard deviation from station to station, indicating that the error is pronounced in areas of complex topography. The model topography is too smooth even at 2 km grid size and lacks to reproduce the real topography, which explains the high standard deviations.

The influence of the horizontal resolution on the model performance is investigated using the C6 configuration as an example (Fig. 9). Considering all stations, spatial correlations are 0.74, 0.39 and 0.22 for the resolutions from 2, 6 to 18 km, respectively. Similarly, the bias increases from 0.46, 2.19 and 3.24, respectively. This increase in bias is explained by the fact that a coarser resolution implies smoother orography, which eventually leads to an excess of wind speed due to the underestimated terrain roughness. The smoothness is also a reason why the RMSE monotonously increases.

For temporal metrics, a somewhat unexpected behaviour is found. Although the temporal correlation drops to a median value of zero in the coarsest domain analysed, the model exhibits a remarkable high correlation in the 6 km resolution domain. To better understand this high correlation the site-averaged wind speed in different domains for the case Lothar is compared to observations (Fig. 10). Although the series corresponding to D4 is more realistic and reproduces the timing and intensity of the most severe wind speed, the simulation in 6 km (D3) captures the phasing of secondary peaks in the time series better than in the 2 km resolution. Indeed, the RMSE reproduces the expected result of a reduction in performance when successive coarser domains are used~~, and demonstrates~~ . This behaviour is an instance of the more general problem of the intrinsically lower predictability of features at smaller scales, as described by Mass et al. (2002) . They showed how, although high-resolution simulation have the potential to improve the simulation of physical processes with respect to coarser ones, they are more severely affected by timing and spatial errors, as well as to deficiencies of the observational network used for verification. Thus, our example of winds over complex terrain illustrates how the validation of high-resolution models is a cumbersome task, and that the use of of several statistics allows more robust assessments of ~~the~~ model performance.

The role of the representativity error can be explored through the separation of the observational sites in subcategories such as stations in plains, mountain or valleys, as shown in Fig. 9 labelled PL, MO and VA, respectively. Although the temporal correlation is not dramatically dependent on the geographical category, it is slightly higher over plains, as expected by the fact that the model resolution is more suitable for simple terrain, and it is worse over valleys, where important terrain features remain unresolved. The performance measurements behave similar in D3, when separating in PL, MO and VA. In D2 the correlation is too low that it precludes drawing any conclusion. Although biases are generally close to zero in the innermost domain, there is a larger variation between the stations in the mountains, because the differences between the station height and model topography can be large and indeed RMSE is significantly larger in this locations.

The spatial correlation in the innermost domain shows a low value of 0.31 over the plains, which contrasts with the value of 0.78 obtained for mountains. This can be explained as a signal-to-noise artefact. The problem is that in plains the mean wind is not as strongly modulated by height as it is in mountains, where there is a larger difference among stations. Thus, small variations in mean wind lead to large variations in the spatial correlation, since the mean wind speed is not a good predictor of the location of a station within plains. Additionally, the correlation is calculated according to only the 46 stations that corresponds to plains in the Lothar storm. Such low number leads to large variance of the estimator of correlation, which further contributes to the signal-to-noise problem. Thus, spatial correlation of mean wind patterns over homogeneous terrain is not a meaningful measure of model ~~skill~~ performance and should be treated with care.

5 Summary and conclusions

This paper analyses a number of sensitivity experiments aimed at identifying a model setup for WRF that minimises systematic errors in hindcast simulations of wind storms over areas of complex topography. The simulations use the ~~Era-Interim~~ ERA-Interim reanalysis for initial and boundary conditions. These data are downscaled to a resolution of 2 km over the Alps in a series of consecutive nested domains. Due to the high demand of computational resources, the analysis is based on case studies, rather than continuous simulations over several years. Therefore, 24 different simulations lasting 6 days and containing prominent historical storms in Switzerland (Etienne et al., 2013) are simulated and analysed. This selection is motivated by their relevance in risk assessments and impact studies, which are two typical applications of dynamically downscaled datasets. To identify a suitable setup to realistically simulate wind over complex terrain, 8 different sensitivity experiments are performed for each case study taking into account different PBL parameterizations, nudging techniques and horizontal resolutions.

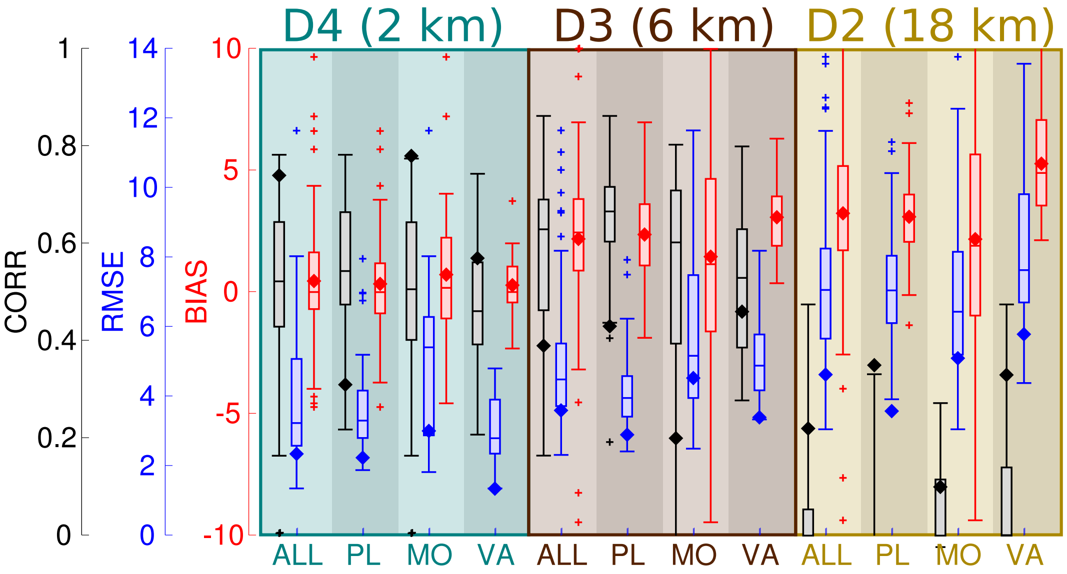


Figure 9. Influence of the grid size ~~for~~ on the ~~simulation skill~~different performance measures, based on the configuration C6 (Table 2). The figure depicts temporal and spatial correlation, RMSE and bias in a similar way as Fig. 5, but different spatial resolutions (grid size 18 km, 6 km and 2 km in domain D2, D3 and D4) and locations (ALL for all stations; PL for plains; MO for mountains; defined as those locations whose height exceeds 1200 m; and VA for valleys) are shown separately. The locations of each type of station are indicated in Fig. 1.

The sensitivity tests designed to evaluate the role of the PBL parameterization show that WRF systematically overestimates wind speed compared to observations. The overestimation occurs in all ~~type~~ types of locations (plains, valleys or mountains), and is exacerbated in coarser domains. This result confirms previous studies pointing out the overestimation of wind speed in simulations with WRF and its relation with ~~unresolved topograhy~~ the lack of representation of the unresolved topography (e.g Cheng and Steenburgh, 2005; Mass and Ovens, 2011; Jiménez and Dudhia, 2012). For the MYJ scheme wind speeds that are up to 100% larger than in the observations are found. The overestimation becomes even stronger when focusing on maximum wind speed, a variable especially relevant in impacts studies. However, this drawback can be significantly reduced by choosing the YSU* scheme which, being based on the non-local YSU scheme, explicitly accounts for unresolved orography. These results resemble findings by Jiménez and Dudhia (2012) and (Gonçalves-Ageitos et al., 2015), who tested this PBL parameterization in a small area of relatively complex topography in the North of the Iberian Peninsula and in the Pyrenees, respectively. It is noteworthy that this improvement is not produced by a trivial reduction of wind speed in every location, but this reduction is applied where complexity of topography is more severely underestimated, yielding a remarkable increase in the model's ability to reproduce the spatial structure of wind speed. As a minor caveat, this scheme tends to slightly reduce the temporal correlation of the simulated

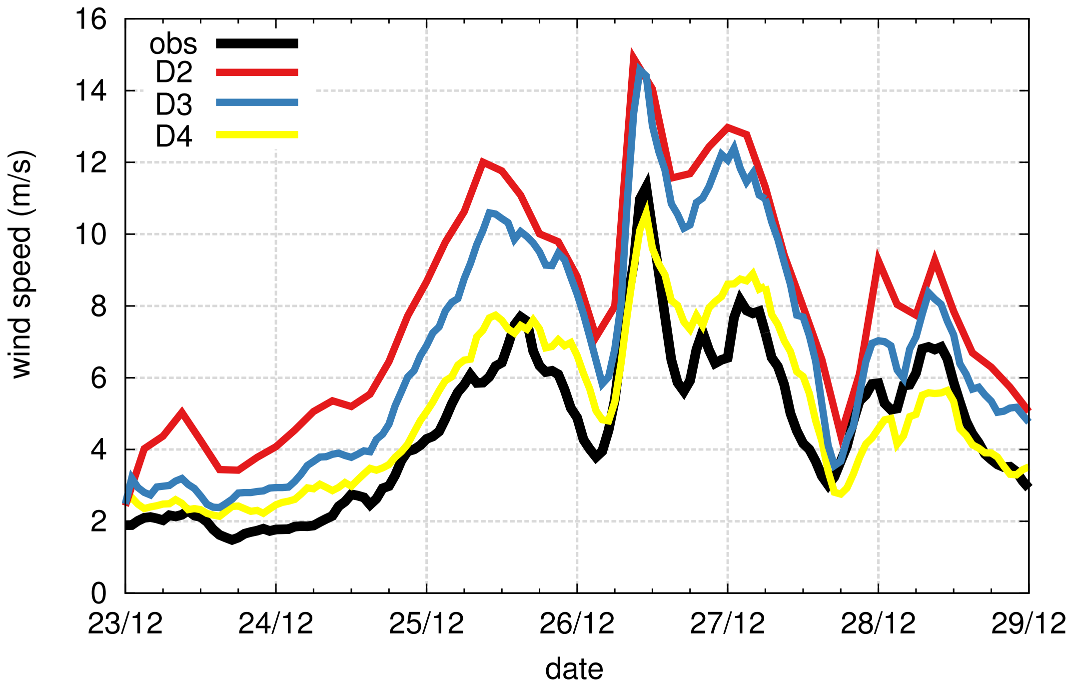


Figure 10. Site-averaged series of wind speed in the 6-day case study containing the storm Lothar. Black and yellow lines correspond to the observations and simulation in the domain D4 corresponding to C6 in Table 2 (same series as in Fig. 4). The ~~green~~ red and blue lines correspond to the result with the same model configuration, but in the domains D2 and D3, respectively.

wind compared to observations, being an unexpected side effect that has also been reported by García-Díez et al. (2015) in simulations over the Iberian Peninsula. The authors do not have however a satisfactory explanation for this behaviour, whose detailed analysis shall be addressed in future studies.

The model is qualitatively able to reproduce the leading wind directions generated by very different synoptic conditions. However, the simulations still exhibit systematic biases in wind direction that cannot be improved through a suitable model configuration~~. Generally,~~, since the model performance in reproducing wind direction exhibits little sensitivity to all the evaluated model configurations. This fact points toward the prominent of representativity errors produced by the channelling of wind in valleys that cannot be properly reproduced by the 2 km resolution of the simulations. Thus, the model performance regarding wind direction is dominated by other factors such as the driving conditions, insufficient resolution, or representativity errors.

Additionally, the sensitivity with respect to the nesting technique is explored by comparing free simulations to analysis and spectral nudging, as well as the so-called reforecast approach. The use of nudging techniques slightly improves several aspects of the simulation, like reducing the mean wind overestimation discussed above and improving the spatial pattern of mean wind (in particular

3D analysis nudging). Further, the free simulations generally show a lower temporal agreement with observations than nudged simulations, a feature that is consistent across storms. Analysis nudging yields a significant improvement for maximum wind speed, for which the overestimation is reduced and leads to values closer to zero on average than when no nudging is applied. These results indicate that preserving the large-scale circulation via nudging slightly improves the simulation of wind at regional scales, at least for hindcast simulations where the driving dataset is generally reliable, and whose aim is to be as close to the observations as possible. ~~We note however that for~~ Still, the particular choice of cases considered in this study could be underestimating the actual effect of nudging, since the strong flow crossing the domain in strong wind storms leaves little freedom for the model to develop disturbances when no nudging is applied. However, for other scientific questions, a free simulation setup could be more appropriate, as atmospheric processes and their interactions with regional scale features are able to develop desirable disturbances that add value to RCM simulations. Typical examples are climate change projections (van der Linden P. and Mitchell, 2009; Gómez-Navarro et al., 2010; Jacob et al., 2013), paleosimulations (Gómez-Navarro et al., 2013) but also classical sensitivity and process studies (Kilic and Raible, 2013; Cipagauta et al., 2014).

Using the setup with analysis nudging and the YSU* scheme, the role of the spatial resolution and the representativity error is assessed. As expected, horizontal resolution is critical for a realistic wind simulation in very complex terrain. A reduction from 6 to 2 km shows a clear improvement in simulating the mean wind pattern as well as maximum winds. This contrasts with the results reported by García-Díez et al. (2015) , that found little added value at 9 km resolution simulations driven by WRF over the Iberian peninsula, and indicates that the ability of RCMs to add value to the driving datasets depends critically on the complexity of the area of interest. In particular, this study demonstrates the ability of WRF to add value in simulations up to 2 km over the Alpine region. The results for the 18 km configuration show barely any ~~skill~~performance, with negligible spatial and temporal correlations. Thus, the overestimation of wind speed becomes exacerbated in coarser resolution domains, further indicating that the main source of wind overestimation is the unresolved orography. Separating in plain, mountain and valley areas the temporal agreement is slightly higher over flat terrain, and reduced in valleys. The mean biases are similar, although showing more spatial variability in the mountains, driven by the larger variability of height biases. More remarkable differences are seen in the RMSE values which show relatively high values of about 6 m/s in the mountains compared to 3 m/s in the flat regions and valleys.

In summary, this study suggests two setups ~~depending~~ for the simulation of wind storms over complex topography. They depend on the scientific question: (i) the configuration C6 with the YSU* scheme that reduces wind overestimation and increases spatial correlations. It further uses 3D analysis nudging, that improves the temporal agreement with respect to observations, and at the same time further reduces the overestimation of maximum wind speed and improves the spatial distribution of wind speed. Thus, this combination is the most suitable for running hindcast simulations aimed at achieving a reliable surface wind simulation over areas of complex orography and in synoptic situations leading to severe storms. (ii) When the timing is not so relevant but an undisturbed development of regional processes is needed, the configuration using the YSU* scheme and free simulations delivers a realistic simulation of surface winds over complex terrain.

Acknowledgements. The authors are grateful for the support provided by the Oeschger Centre for Climate Change Research and the Mobiliar lab for climate risk and natural hazards (Mobilab). Thanks are due to the Swiss National Supercomputing Centre (CSCS) for providing the supercomputing facilities and technical support required to perform the RCM simulations. The ~~ERA-interim~~ ERA-Interim reanalysis was kindly provided by the ECMWF. ~~Finally, we~~ We would like to thank Martina Messmer for sharing her great knowledge of the Swiss geography and the stimulating discussions. Finally, thanks to the two anonymous reviewers that greatly contributed to correct errors and improve the discussion of the results.

References

Awan, N. K., Truhetz, H., and Gobiet, A.: Parameterization-Induced Error Characteristics of MM5 and WRF Operated in Climate Mode over the Alpine Region: An Ensemble-Based Analysis, Journal of Climate, 24, 3107–3123, doi:10.1175/2011JCLI3674.1, 2011.

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A. T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K.: Future extreme events in European climate: an exploration of regional climate model projections, Climatic Change, 81, 71–95, doi:10.1007/s10584-006-9226-z, 2007.

Charney, J., Halem, M., and Jastrow, R.: Use of Incomplete Historical Data to Infer the Present State of the Atmosphere, Journal of the Atmospheric Sciences, 26, 1160–1163, doi:10.1175/1520-0469(1969)026<1160:UOIHDT>2.0.CO;2, 1969.

Chen, F. and Dudhia, J.: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity, Monthly Weather Review, 129, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2, 2001.

Cheng, W. Y. Y. and Steenburgh, W. J.: Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States, Weather and Forecasting, 20, 812–821, doi:10.1175/WAF885.1, 2005.

Cipagauta, C., Mendoza, B., and Zavala-Hidalgo, J.: Sensitivity of the surface temperature to changes in total solar irradiance calculated with the WRF model, Geofísica Internacional, 53, 153–162, doi:10.1016/S0016-7169(14)71497-7, 2014.

Davies, H. C.: A laterul boundary formulation for multi-level prediction models, Quarterly Journal of the Royal Meteorological Society, 102, 405–418, doi:10.1002/qj.49710243210, 1976.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Dierer, S., de Paus, T., Durante, F., Gregow, E., Lange, B., Lavagnini, A., Strack, M., and Tammelin, B.: Predicting Wind Speed for Wind Energy; Progress of the WINDENG Project, Wind Engineering, 29, 393–408, doi:10.1260/030952405775992616, 2005.

Draxl, C., Hahmann, A. N., Peña, A., and Giebel, G.: Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes, Wind Energy, 17, 39–55, doi:10.1002/we.1555, 2014.

Dudhia, J.: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model, Journal of the Atmospheric Sciences, 46, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2, 1989.

Etienne, C., Goyette, S., and Kuszli, C.-A.: Numerical investigations of extreme winds over Switzerland during 1990–2010 winter storms with the Canadian Regional Climate Model, Theoretical and Applied Climatology, 113, 529–547, doi:10.1007/s00704-012-0800-1, 2013.

García-Díez, M., Fernández, J., Fita, L., and Yagüe, C.: Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe, Quarterly Journal of the Royal Meteorological Society, 139, 501–514, doi:10.1002/qj.1976, 2013.

García-Díez, M., Fernández, J., San-Martín, D., Herrera, S., and Gutiérrez, J. M.: Assessing and Improving the Local Added Value of WRF for Wind Downscaling, Journal of Applied Meteorology and Climatology, 54, 1556–1568, doi:10.1175/JAMC-D-14-0150.1, 2015.

Gómez-Navarro, J. J., Montávez, J. P., Jimenez-Guerrero, P., Jerez, S., García-Valero, J. A., and González-Rouco, J. F.: Warming patterns in regional climate change projections over the Iberian Peninsula, Meteorologische Zeitschrift, 19, 275–285, doi:10.1127/0941-2948/2010/0351, 2010.

Gómez-Navarro, J. J., Montávez, J. P., Jerez, S., Jiménez-Guerrero, P., and Zorita, E.: What is the role of the observational dataset in the evaluation and scoring of climate models?, Geophysical Research Letters, 39, L24 701–L24 701, doi:10.1029/2012GL054206, 2012.

Gómez-Navarro, J. J., Montávez, J. P., Wagner, S., and Zorita, E.: A regional climate palaeosimulation for Europe in the period 1500–1990 – Part 1: Model validation, Clim. Past, 9, 1667–1682, doi:10.5194/cp-9-1667-2013, 2013.

Gonçalves-Ageitos, M., Barrera-Escoda, A., Baldasano, J. M., and Cunillera, J.: Modelling wind resources in climate change scenarios in complex terrains, Renewable Energy, 76, 670–678, doi:10.1016/j.renene.2014.11.066, 2015.

Goyette, S.: Development of a model-based high-resolution extreme surface wind climatology for Switzerland, Natural Hazards, 44, 329–339, doi:10.1007/s11069-007-9130-5, 2008.

Goyette, S.: Synoptic conditions of extreme windstorms over Switzerland in a changing climate, Climate Dynamics, 36, 845–866, doi:10.1007/s00382-010-0820-4, 2010.

Hong, S. and Lim, J.: The WRF single-moment 6-class microphysics scheme (WSM6), Journal of Korean Meteorology Society, 42, 129–151, 2006.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., Meijgaard, E. v., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, doi:10.1007/s10113-013-0499-2, 2013.

Janjić, Z.: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model, Tech. Rep. Officiale Note 437, National Centers for Environmental Prediction, Boulder, Colorado, USA, 2001.

Jerez, S. and Trigo, R. M.: Time-scale and extent at which large-scale circulation modes determine the wind and solar potential in the Iberian Peninsula, Environmental Research Letters, 8, 044 035, doi:10.1088/1748-9326/8/4/044035, 2013.

Jerez, S., Montavez, J. P., Gomez-Navarro, J. J., Lorente-Plazas, R., Garcia-Valero, J. A., and Jimenez-Guerrero, P.: A multi-physics ensemble of regional climate change projections over the Iberian Peninsula, Climate Dynamics, 41, 1749–1768, doi:10.1007/s00382-012-1551-5, 2013.

Jiménez, P. A. and Dudhia, J.: Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model, Journal of Applied Meteorology and Climatology, 51, 300–316, doi:10.1175/JAMC-D-11-084.1, 2012.

Jiménez, P. A. and Dudhia, J.: On the Ability of the WRF Model to Reproduce the Surface Wind Direction over Complex Terrain, Journal of Applied Meteorology and Climatology, 52, 1610–1617, doi:10.1175/JAMC-D-12-0266.1, 2013.

Jiménez, P. A., García-Bustamante, E., González-Rouco, J. F., Valero, F., Montávez, J. P., and Navarro, J.: Surface Wind Regionalization in Complex Terrain, Journal of Applied Meteorology and Climatology, 47, 308–325, doi:10.1175/2007JAMC1483.1, 2008.

Jiménez, P. A., González-Rouco, J. F., García-Bustamante, E., Navarro, J., Montávez, J. P., de Arellano, J. V.-G., Dudhia, J., and Muñoz-Roldan, A.: Surface Wind Regionalization over Complex Terrain: Evaluation and Analysis of a High-Resolution WRF Simulation, Journal of Applied Meteorology and Climatology, 49, 268–287, doi:10.1175/2009JAMC2175.1, 2010.

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., and García-Bustamante, E.: A Revised Scheme for the WRF Surface Layer Formulation, Monthly Weather Review, 140, 898–918, doi:10.1175/MWR-D-11-00056.1, 2012.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, Journal of Applied Meteorology, 43, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kilic, C. and Raible, C. C.: Investigating the sensitivity of hurricane intensity and trajectory to sea surface temperatures using the regional model WRF, Meteorologische Zeitschrift, 22, 685–698, doi:10.1127/0941-2948/2013/0472, 2013.

Kjellström, E., Bärring, L., Jacob, D., Jones, R., Lenderink, G., and Schär, C.: Modelling daily temperature extremes: recent climate and future changes over Europe, Climatic Change, 81, 249–265, doi:10.1007/s10584-006-9220-5, 2007.

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

Kwun, J. H., Kim, Y.-K., Seo, J.-W., Jeong, J. H., and You, S. H.: Sensitivity of MM5 and WRF mesoscale model predictions of surface winds in a typhoon to planetary boundary layer parameterizations, Natural Hazards, 51, 63–77, doi:10.1007/s11069-009-9402-3, 2009.

Lo, J. C.-F., Yang, Z.-L., and Pielke, R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, Journal of Geophysical Research: Atmospheres, 113, D09 112, doi:10.1029/2007JD009216, 2008.

Lorente-Plazas, R., Montávez, J. P., Jerez, S., Gómez-Navarro, J. J., Jiménez-Guerrero, P., and Jiménez, P. A.: A 49 year hindcast of surface winds over the Iberian Peninsula, Journal of Climate, p. Publised online, doi:10.1002/joc.4189, 2014a.

Lorente-Plazas, R., Montávez, J. P., Jiménez, P. A., Jerez, S., Gómez-Navarro, J. J., García-Valero, J. A., and Jiménez-Guerrero, P.: Characterization of surface winds over the Iberian Peninsula, International Journal of Climatology, p. Publised online, doi:10.1002/joc.4034, 2014b.

Mass, C. and Ovens, D.: Fixing WRF's high speed wind bias: A new subgrid scale drag parameterization and the role of detailed verification, in: 24th Conf. on Weather and Forecasting/20th Conf. on Numerical Weather Prediction, vol. 9B.6 [Available online at http://ams.confex.com/ams/91Annual/webprogram/Paper180011.html], Amer. Meteor. Soc., 2011.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A.: Does Increasing Horizontal Resolution Produce More Skillful Forecasts?, Bulletin of the American Meteorological Society, 83, 407–430, doi:10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2, 2002.

Mellor, G. L. and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Reviews of Geophysics, 20, 851–875, doi:10.1029/RG020i004p00851, 1982.

Menendez, M., García-Díez, M., Fita, L., Fernández, J., Méndez, F. J., and Gutiérrez, J. M.: High-resolution sea wind hindcasts over the Mediterranean area, Climate Dynamics, 42, 1857–1872, doi:10.1007/s00382-013-1912-8, 2014.

Miguez-Macho, G., Stenchikov, G. L., and Robock, A.: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations, Journal of Geophysical Research: Atmospheres, 109, D13 104, doi:10.1029/2003JD004495, 2004.

Miguez-Macho, G., Stenchikov, G. L., and Robock, A.: Regional Climate Simulations over North America: Interaction of Local Processes with Improved Large-Scale Flow, Journal of Climate, 18, 1227–1246, doi:10.1175/JCLI3369.1, 2005.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102, 16 663–16 682, doi:10.1029/97JD00237, 1997.

MunichRe: Winter storms in Europe (II)- Analysis of 1999 losses and loss potenticals, Munich Reinsurance Company, http://www.planat.ch/fileadmin/PLANAT/planat_pdf/alle_2012/2001-2005/Munich_Re_Group_2002_-_Winter_storms_in_Europe_II.pdf, 2001.

Muskulus, M. and Jacob, D.: Tracking cyclones in regional model data: the future of Mediterranean storms, Adv. Geosci., 2, 13–19, doi:10.5194/adgeo-2-13-2005, 2005.

Pleim, J. E.: A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing, Journal of Applied Meteorology and Climatology, 46, 1383–1395, doi:10.1175/JAM2539.1, 2007a.

Pleim, J. E.: A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part II: Application and Evaluation in a Mesoscale Meteorological Model, Journal of Applied Meteorology and Climatology, 46, 1396–1409, doi:10.1175/JAM2534.1, 2007b.

Rajczak, J., Pall, P., and Schär, C.: Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region, Journal of Geophysical Research: Atmospheres, 118, 3610–3626, doi:10.1002/jgrd.50297, http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50297/abstract, 2013.

Rivière, G., Arbogast, P., Maynard, K., and Joly, A.: The essential ingredients leading to the explosive growth stage of the European wind storm Lothar of Christmas 1999, Quarterly Journal of the Royal Meteorological Society, 136, 638–652, doi:10.1002/qj.585, 2010.

Santos-Alamillos, F. J., Pozo-Vázquez, D., Ruiz-Arias, J. A., Lara-Fanego, V., and Tovar-Pescador, J.: Analysis of WRF Model Wind Estimate Sensitivity to Physics Parameterization Choice and Terrain Representation in Andalusia (Southern Spain), Journal of Applied Meteorology and Climatology, 52, 1592–1609, doi:10.1175/JAMC-D-12-0204.1, 2013.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, Nature, 427, 332–336, doi:10.1038/nature02300, 2004.

Schiesser, H. H., Pfister, C., and Bader, J.: Winter storms in Switzerland North of the Alps 1864/1865–1993/1994, Theoretical and Applied Climatology, 58, 1–19, doi:10.1007/BF00867428, 1997.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, Tech. Rep. TN–475+STR, National Center for Atmospheric Research Boulder, Colorado, USA, 2008.

Staniforth, A.: Regional modeling: A theoretical discussion, Meteorology and Atmospheric Physics, 63, 15–29, doi:10.1007/BF01025361, 1997.

Stensrud, D. J.: Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models, Cambridge University Press, Cambridge, United Kingdom, first edn., 2007.

Stucki, P., Brönnimann, S., Martius, O., Welker, C., Rickli, R., Dierer, S., Bresch, D. N., Compo, G. P., and Sardeshmukh, P. D.: Dynamical downscaling and loss modeling for the reconstruction of historical weather extremes and their impacts – A severe foehn storm in 1925, Bulletin of the American Meteorological Society, doi:10.1175/BAMS-D-14-00041.1, 2015.

van der Linden P. and Mitchell, J., eds.: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project., Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK., 2009.

von Storch, H. and Weisse, R.: Regional storm climate and related marine hazards in the Northeast Atlantic, in: Climate extremes an society, edited by Diaz, H. F. and Murnane, R. J., pp. 54–73, Cambridge Unviersity Press, Cambridge, 2008.

von Storch, H., Langenberg, H., and Feser, F.: A Spectral Nudging Technique for Dynamical Downscaling Purposes, Monthly Weather Review, 128, 3664–3673, doi:10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2, 2000.

Whiteman, C. D.: Mountain Meteorology: Fundamentals and Applications, Oxford University Press, New York, 2000.