Response to reviewer #2

We thank the reviewer for his/her comments, which are reproduced in black hereafter. Our responses are in blue. In the revised version of the manuscript, all the modifications are in red.

1. General comments

Dupuy et al. explored methods for grid-to-grid downscaling of surface wind. Building on the latest contributions in the field, which are well documented throughout the manuscript, they evaluate a variety of approaches. In doing so, they have also included key components of previous works. This contributes to a certain continuity in the literature, which is appreciated and helps focus new efforts. Rather than proposing a new architecture, the authors focus on the different modeling choices in terms of target variables and loss functions. They find that specific approaches perform best for either wind speed or wind direction, but not both at the same time. They show how the approaches can be combined to yield the best results on their evaluation.

The manuscript is well-written and easy to follow. The methodology is solid, although some minor aspects could be improved. Overall, a very valuable contribution to the field. I accepted with minor revisions, see section 2.2.

2. Specific comments

2.1 Discussions/clarifications

- Predictors: in the final results have you used time-related variables, such as cosine and sine components of the hour of the day? If not, have you considered them during your study? In combination with topographical predictors, they might help model the diurnal cycle.

We did not consider adding such predictors since we already used meteorological predictors that are highly influenced by the diurnal cycle: solar radiation, temperature, … Moreover, the relation of the hour of the day with the diurnal cycle varies along the year since the sunset and sunrise times are depending on the date. We added a sentence in the revised manuscript in that sense (section 2.2, lines 97-99).

- Figure 11a and b: I might be wrong, but 11b (CNN) seems to have a chessboard-like pattern, however 11a (WRF LR) does not. Since it is my understanding that CNN results use WRF LR as input, I find it a bit strange. Are you sure that the same interpolation method (bicubic) is used in both cases?

We confirm that we performed the same bicubic interpolation for all the predictors as well as for the LR WRF forecasts before using them.

We also noticed that after downscaling, the MAE for the speed (Fig. 12b of the revised manuscript) features a grid pattern corresponding to the original LR grid data, in general with lower values (i.e. a better forecast) at the center of the LR grid cells. This feature is also slightly visible on the MAE for the speed from WRF LR (Fig. 12a of the revised manuscript), where MAE values are low (cf. the bottom left part of the plots in Fig. 1 below). The pattern is not visible on the more northern areas of the domain, possibly because of higher MAE values. Therefore, our guess is that the better forecast performance at the center of the LR grid cells in WRF LR (even after the bicubic interpolation) causes that pattern in the CNN results.

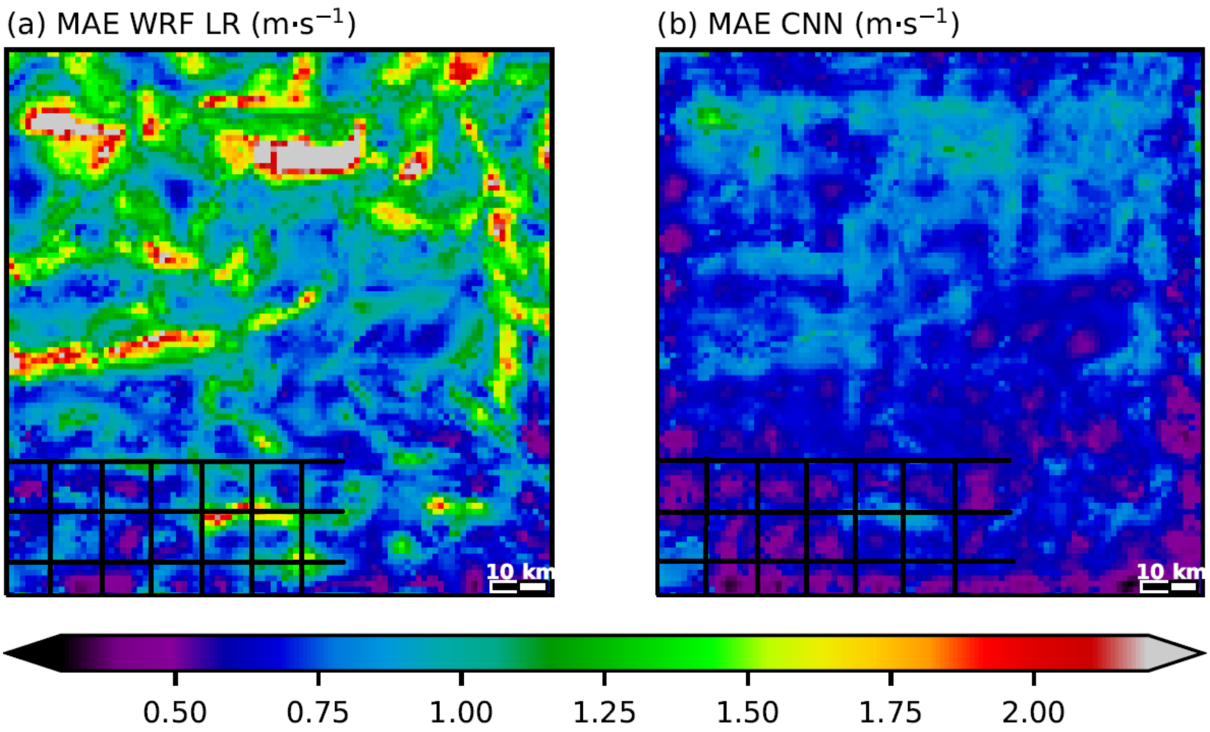

Figure 1: Same as Fig. 12 a and b of the revised manuscript, with the LR grid added on the bottom left part of the plots.

- u^2 + v^2 = 1: could you not strictly enforce this condition directly in the architecture, by having the model predict "u(x)" and "sign(x)" in "v = sign(x) sqrt(1 - u^2)"? If possible, it could be a nice addition to the manuscript.

Thank you for this suggestion. As the other reviewer made a similar comment, a joint response is provided below.

We tried to implement the suggested modification by computing the value of $\widehat{\tilde{v}}$ knowing only the value of $\widehat{\tilde{u}}$ using the formula $\widehat{\tilde{u}}^2 + \widehat{\tilde{v}}^2 = 1$ in order to get couples of $\widehat{\tilde{u}}$ and $\widehat{\tilde{v}}$ values that are consistent. However, it is not possible to derive the sign of $\widehat{\tilde{v}}$ from this formula. Therefore, using the results of $CNN_{\tilde{u},\tilde{v}}$, we used the sign from the output $\tilde{v}$ in addition to the $\tilde{u}$ output values to calculate $\widehat{\tilde{v}}$ as follows:

$$\widehat{\tilde{v}} = sign(\widehat{\tilde{v}}) \times \sqrt{1 - \widehat{\tilde{u}}^2}$$

Similarly, we computed $\widehat{\tilde{u}}$ as:

$$\widehat{\tilde{u}} = sign(\widehat{\tilde{u}}) \times \sqrt{1 - \widehat{\tilde{v}}^2}$$

These two new tests are called $CNN_{\tilde{u}\rightarrow\tilde{v}}$ and $CNN_{\tilde{v}\rightarrow\tilde{u}}$ in the revised version of the article (note that no additional CNNs were trained since we used the results of the $CNN_{\tilde{u},\tilde{v}}$). We got results that were worse than with $CNN_{\tilde{u},\tilde{v}}$ and $CNN_{\tilde{u},\tilde{v},L2}$ on the direction forecast.

This new approach is described in section 2.3.2 (lines 175-180). The results of $CNN_{\tilde{u}\rightarrow\tilde{v}}$ and $CNN_{\tilde{v}\rightarrow\tilde{u}}$ are presented in Fig. 4 and lines 245-248 in the revised manuscript.

- Area of study: given the focus on complex terrain, I question whether the D3 domain is the best choice. Moving it a bit further to the east would have included a more diverse set of topographical features (although would have excluded Mont Ventoux, ironically), possibly giving more insights and highlighting the benefits of the presented methods even more. This does not change the value of the manuscript, but it might be something to consider for future contributions.

We agree with the reviewer that it would have been interesting to have a HR dataset over an ever more complex terrain area. However, these simulations, performed by a team of the CEA (Commissariat à l'Énergie Atomique et aux Énergies Alternatives) located at Cadarache, France (which is located at the center of the D3 domain), are originally motivated by impact studies of potential releases over the site. It is therefore not possible for us to get a large dataset on a shifted domain.

2.2 Minor revisions requested:

- 288x288 domain: this is not the same as the D2 domain, correct? I would make this more clear, and maybe include this domain in Figure 1.

It is right that the 288x288 domain is not exactly the same as the D2 domain. The D2 domain has a side of 297 km long, whereas it is shorter by 9 km for the 288 x 288 domain. But both domains are centered on the same point, and their borders are very close each other (cf. Fig. 2 below). We added these details in the text (lines 123-126) and we changed the Fig. 1 in the new version of the manuscript as requested.

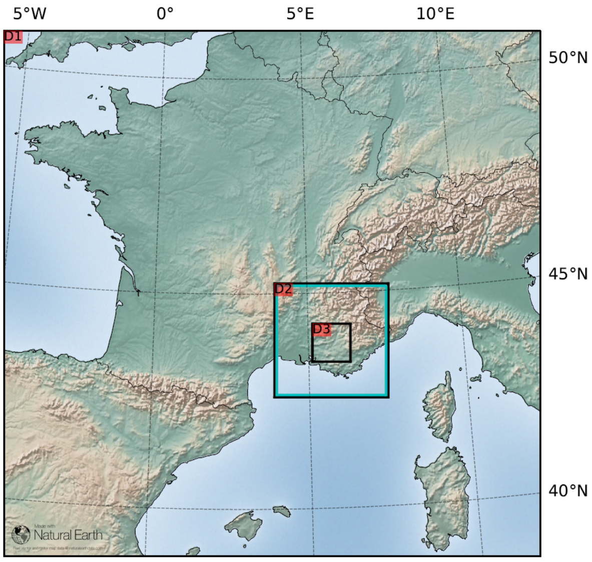

Figure 2: Representation of the 3 nested domains of the WRF model. The blue square represents the area corresponding to the 288x288 domain.

- Please provide more details on the cross-validation strategy.

As the other reviewer made a similar comment, a joint response is provided below.

For each CNN model tested, we performed a k-fold cross validation (k=4, so 4 different trainings for each CNN model tested are performed) so that we have the largest dataset to test the models. For each of the 4 trainings, 25% of the dataset is used for testing while the remaining 75% are used for training. By combining the results of the 4 trainings applied to the 4 test sets, we get a testing dataset for the whole period (that is to say the largest dataset possible in our case).

As this is highly related to the evaluation process, we think it is more appropriate to describe this method in section 2.4. We added more details in that section.

- Figure 4: it would help to have the label of the best-performing model in boldface.

We changed the label of the best-performing model to boldface for each metric on Fig. 4.

- Overall wind speed climatology: while I appreciate the in-depth analysis at the two specific sites, and understand its value, particularly for the qualitative evaluation, I believe more domain-level (or at many randomly selected points, if the size of the data is a constraint) quantitative analysis is needed to complement the verification metrics. For instance, it would be nice to see scatter plots or conditional quantile plots (see Wilks 2011) for wind speed for the entire spatial and temporal domain.

According to this suggestion, we added the comparison of wind speed climatology on the whole domain (comparison of probability density functions for WRF HR, WRF LR and the CNN), cf. Fig. 11 and related comments (lines 328-332) in the new version of the article.

- The sub-optimal generalization capability outside the D3 domain is to be expected for this methodology. If you already have some relevant results, they should be included and discussed. In the future, it will be interesting to see how domain-agnostic models compete with domain-specific models.

As the other reviewer made a similar comment, a joint response is provided below.

As stated in the paragraph mentioned, this is only a preliminary work. Indeed, for now, the dataset on the other sites is really small (a simulation of 72 hours on a single site, cf. Figs. 1 and 2 below), preventing a significant analysis. Therefore, this work must be extended by generating a larger (in time and spatially) HR dataset in order to get significant insight from the results. For this reason, we think these results are not worth publishing at the present state and must remain as a mention to future work to make.

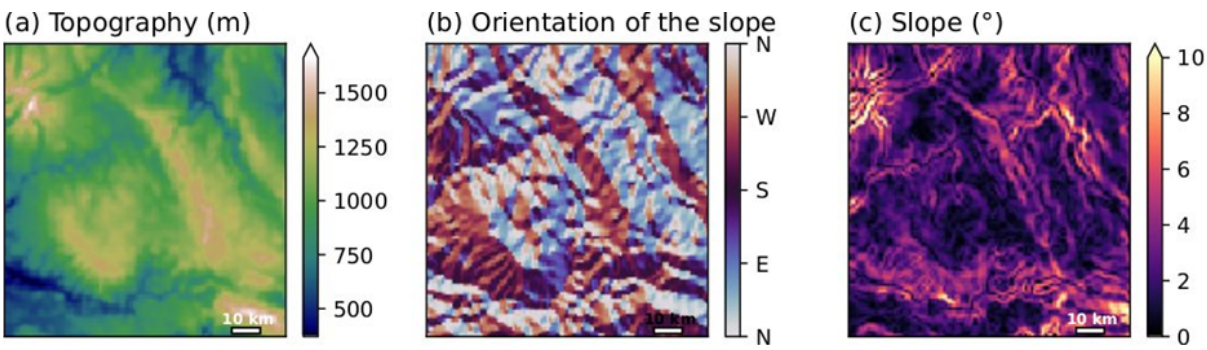

Figure 1: New domain tested (center of France) – (a) topography (in m a.s.l.), (b) orientation of the slope and (c) local slope.

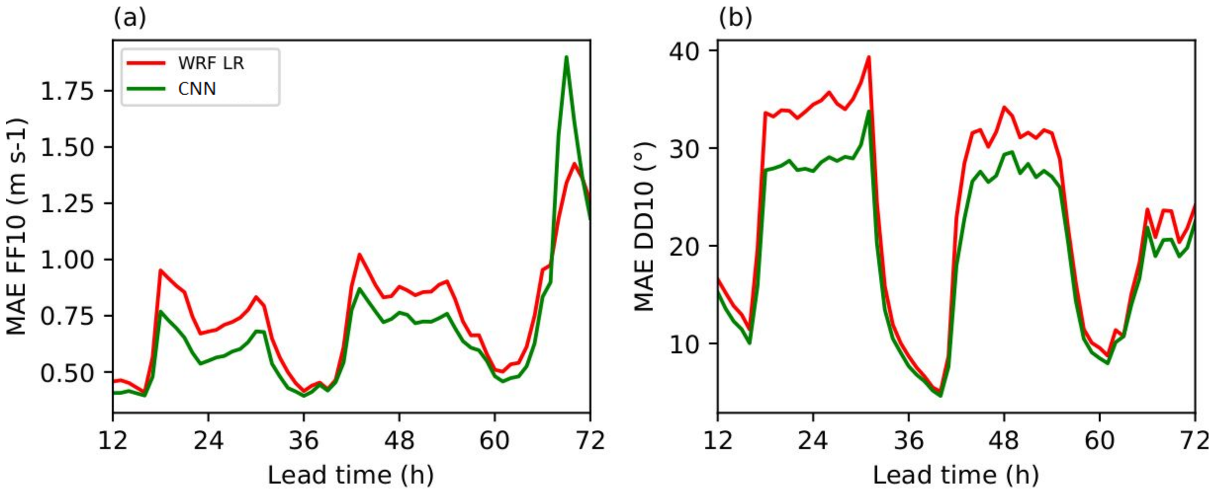

Figure 2: Evolution of the MAE on the wind speed (a) and direction (b) average over the whole domain for a 3-day simulation period.

3. Technical comments

- L62-65: a bit convoluted. I suggest rephrasing it as "31 km to 9 km (ratio close to 3) in Höhlein et al. ((2020)" etc.

We changed the sentence following this suggestion (lines 62-64).

- L326: "The diurnal cycle remains" Is this referring to the "cycle" in the MAE? I find it a bit confusing. Please rephrase.

Yes, it was referring to the diurnal cycle of the MAE. We rephrased as:

"After downscaling, the MAE is reduced for all times and its diurnal cycle remains, …" (lines 362-364)