Reply to Report #1: Review of nhess-2020-299

We thank the reviewer for their comments. Our replies are inline below:

Specific comments

I have a comment on the response and paper edits from Reviewer 2 comment 3 about wind speed discrepancy.

Manuscript extract:

"In general, median peak gust speeds from the RAL2 model ensemble are found to be 22 to 43 m s-1 faster compared to ERA5, but it is known that extreme gusts associated with vigorous convection in ERA5 are generally under-estimated, sometimes by a factor of two (Owens and Hewson, 2018)."

My two questions are:

1. Can you add another comment about the implication of this bias? Are you suggesting that the bias is only in ERA5 and the RAL2 winds are correct, or might they also have errors associated with them.

> Generally, only grid spacings on the order of 1 km are comparable to the size of particularly energetic eddies in the planetary boundary layer (e.g. Leutwyler et al., 2017), so the turbulent processes as well as the dominant turbulent length scale will be under resolved in both our downscaled model and ERA5. We have added text to Section 2 to clarify this.

2. Can you comment on the difference between IBTrACS and RAL2? If they are more similar than RAL2 and ERA5, this would help to justify ERA5 errors in wind gusts. Originally the statement was that this 22-43 m/s error was in ERA5 and IBTrACS, but was this incorrect?

> These numbers refer to the gust speed difference against ERA5 only. It was incorrect to reference IBTrACS in association with these numbers, as IBTrACS data does not estimate gust speed. The median difference for wind speed (across all events) is 18 m s^{-1} for ERA5. Comparing IBTrACS wind speeds, median difference against the US forecast is −3 m s^{-1} and against the New Delhi forecast is 5m s^{-1}. We have added a sentence in Section 2 to clarify this.

Line 82: Is the model called RA2 or RAL2?

> Thank you for highlighting this. For consistency we use 'RAL2'.

Line 197: Missing 'to'

> Corrected.

Reply to Report #2: Review of nhess-2020-299

We thank the reviewer for their comments. Our replies are inline below:

Please explain in greater detail what value the hazard maps here would add to the short-term forecast problem in the situation where the same model could be run in ensemble NWP mode in real time using appropriate initial conditions. Why is that by itself not the best way to produce forecast warnings?

> The focus of this paper is not about short-term forecasting. As you describe, ensemble NWP (and implicitly processes such as data assimilation) are the best solution for forecasting a specific event at lead-times of hours to days. The process we describe in Section 3.1 could just as well be applied to ensemble NWP data, as our hazard maps. We have added words in Section 3.1 to try and clarify this.

The paper really needs to have some more explanation of key aspects of the underlying simulations: domain size, duration of simulations and lead time before landfall, a little information about physical parameterizations, definition of a gust, how the simulations were validated, etc. This information is really essential and the reader shouldn't be made to look up a different paper for it (particularly one that isn't published yet, and might not be, as far as we know). The present paper is short and could easily include a new section on this.

> We have added the following to Section 2:

Each ensemble member requires a 24-hour spin-up period as the RAL2 model adjusts from weak initial conditions taken from the ERA5 driving global model (of Hersbach et al., 2020). This initial 24 hours of model data are discarded in subsequent analysis and data files. Thereafter, each ensemble member is free running for a further 48 hours, with hourly boundary conditions provided by ERA5. Collectively, the ensembles members sample a range of lead times before landfall from 12-36 hours.

The parameterised RAL2 gust diagnostic represents a prediction of the 3-second average windspeed at every timestep. The maximum of this 3-second average speed over an hour is then taken to give the hourly maximum 3-second gust speed. While not truly resolving deep convection, RAL2 is able to explicitly represent deep convective processes within the resolved dynamics. At these kilometre-scale resolutions the lower horizontal size limit of convective cells is still set by the effective resolution of 5 to 10 times the grid length (Boutle et al., 2014; Skamarock, 2004). Generally, only grid spacings on the order of 1 km are comparable to the size of particularly energetic eddies in the planetary boundary layer (Leutwyler et al., 2017), so the turbulent processes as well as the dominant turbulent length scale will be under resolved in our downscaled model (and also ERA5). The RAL2 model uses a gust parametrisation based on 10 m wind speed with scaling proportional to the standard deviation of the horizontal wind that also accounts for friction velocity, atmospheric stability and roughness length (Lock et al., 2019).

> Section 2 already includes details of the domain size. The paper describing the modelling set-up and data validation has now been accepted for publication in Scientific Data (Steptoe et al., 2021).

<u>References</u>

Boutle, I. A., Eyre, J. E. J., & Lock, A. P. (2014). Seamless Stratocumulus Simulation across the Turbulent Gray Zone. *Monthly Weather Review*, *142*(4), 1655–1668.

https://doi.org/10.1175/MWR-D-13-00229.1

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/https://doi.org/10.1002/qj.3803

Leutwyler, D., Lüthi, D., Ban, N., Fuhrer, O., & Schär, C. (2017). Evaluation of the convection-resolving climate modeling approach on continental scales. *Journal of Geophysical Research: Atmospheres*, *122*(10), 5237–5258. https://doi.org/https://doi.org/10.1002/2016JD026013

Lock, A., Edwards, J., & Boutle, I. (2019). *Unified Model Documentation Paper 024: The Parametrization of Boundary Layer Processes*.

Skamarock, W. C. (2004). Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. *Monthly Weather Review*, *132*(12), 3019–3032. https://doi.org/10.1175/MWR2830.1

Steptoe, H., Savage, N., Sadri, S., Salmon, K., Maalick, Z., & Webster, S. (2021). Tropical cyclone simulations over Bangladesh at convection permiting 4.4km & 1.5km resolution. *Scientific Data*, (accepted). https://doi.org/10.1038/s41597-021-00847-5