Dear Bruce Harper

Thank you for your in-depth and informed comments. Our responses to your comments and the comments of the other two reviewers will greatly improve the manuscript. Below are our responses (in blue) to each comment in turn.

While responding to comments from all reviewers we found a bug in our code in the way we treat asymmetry. As described in the original manuscript, we first remove an estimate of the asymmetry due to forward speed from the input best track V_{max}. The portion removed is a function of the TC translation speed, $V_a=1.173V_t^{0.63}$, following Chavas et al., (2017). We then add back an estimate of the asymmetry to the spatial 10m wind field diagnosed by KW01, again following Chavas et al., (2017). And in addition, we apply a factor that varies with radial distance from the storm center (the factor is equal to 1 at the radius of maximum winds and then decays with increasing radius) following Jakobsen and Madsen (2004). The bug was that the code missed adding back the V_t-dependent factor and only added back the radially-dependent factor. This caused us to add back the full value of V_t at R_{max} which caused too strong asymmetry, particularly for fast moving storms over Japan for example. In response, original figures 2, 3 and 4 will be corrected. The analysis in original figure 7 is for along-track winds and so will not be affected by this bug fix.

Chavas, D. R., Reed, K. A. and Knaff, J. A.: Physical understanding of the tropical cyclone wind-pressure relationship. Nature communications, 8(1), p.1360, https://doi.org/10.1038/s41467-017-01546-9, 2017.

Jakobsen, F. and H. Madsen, H.: Comparison and further development of parametric tropical cyclone models for storm surge modeling. Journal of Wind Engineering, 92, 375-391, https://doi.org/10.1016/j.jweia.2004.01.003, 2004.

Bruce Harper

The authors seek to “to advance our understanding of overland wind risk in regions of complex terrain and support wind risk assessments in regions of sparse historical data” through the application of a combination of analytical, numerical and empirical techniques. The authors make reference to many very successful more simplified approaches that have been developed over the past three decades that emanate from wind engineering, atmospheric science and insurance loss initiatives. They put the case that previous approaches, in their assessment, lack the essential capacity to incorporate complex terrain and “the essential dynamics and physics” of tropical cyclone (TC) behaviour where “the accuracy of wind speeds over urban (sic) is of critical importance”. The use of a diagnostic 3D numerical boundary layer model is central to their thesis.

While the desirability of such an approach can be supported, where practical, it is ironic that the demonstrated model skill is so poor at reproducing recent historical TC winds in areas where there is a significant amount of quality data available and in mostly flat landscapes. In no way does the model “compare favourably” with the displayed data. An 8 to 10 m/s error band in any hindcast event intended to assist, for example, offshore engineering design or sensitive onshore high-rise design, would be regarded as completely unacceptable. For insurance losses where damage is noted by the authors to be additionally highly nonlinear with wind speed, it would be massively unreliable. The stated need to apply an empirical "20% adjustment for urban areas" is not only completely inconsistent with the theoretical high ground being argued but is symptomatic of a modelling system that has some significant problems.

As you state below, the challenge of developing a globally applicable approach is that the accuracy at the individual event level will be poorer than for a model developed for individual events or specific regions. This means that for our model evaluation for the subset of observation-rich storms over the U.S. we do not expect that our model should be able to compete with individual event-level approaches that assimilate far more observational data. A globally applicable approach has a number of unique benefits that derive from its small amount of required input data and physical response of the boundary layer winds to terrain. These benefits include generating events in data sparse-regions, generating synthetic events, and application to downscaling TC tracks from global climate models. The case for a global approach will be made stronger in the revised manuscript.

We note that recently published work that used similar quantities of input data and local wind multiplication factors to account for terrain features (Tan and Fang, 2018) also showed typical errors of 8 to 10m/s. Their Fig. 6 shows comparison with observations for 36 TCs during 1970–2014 for 25 stations in Hainan Island, China. They show 8 to 10m/s errors are typical in the 10-minute sustained winds. Our approach therefore compares favorably with another global modeling approach.

In addition, we also note this magnitude of error are also present in hurricane surface wind vectors utilizing C-band dual-polarization synthetic aperture radar observations, when compared to collocated QuikSCAT-measured wind speeds. For the case of Hurricane Bill Zhang et al. (2014) (Their Fig. 9a and 9b, copied here below) shows scatter (over the ocean) similar in magnitude to our scatter.

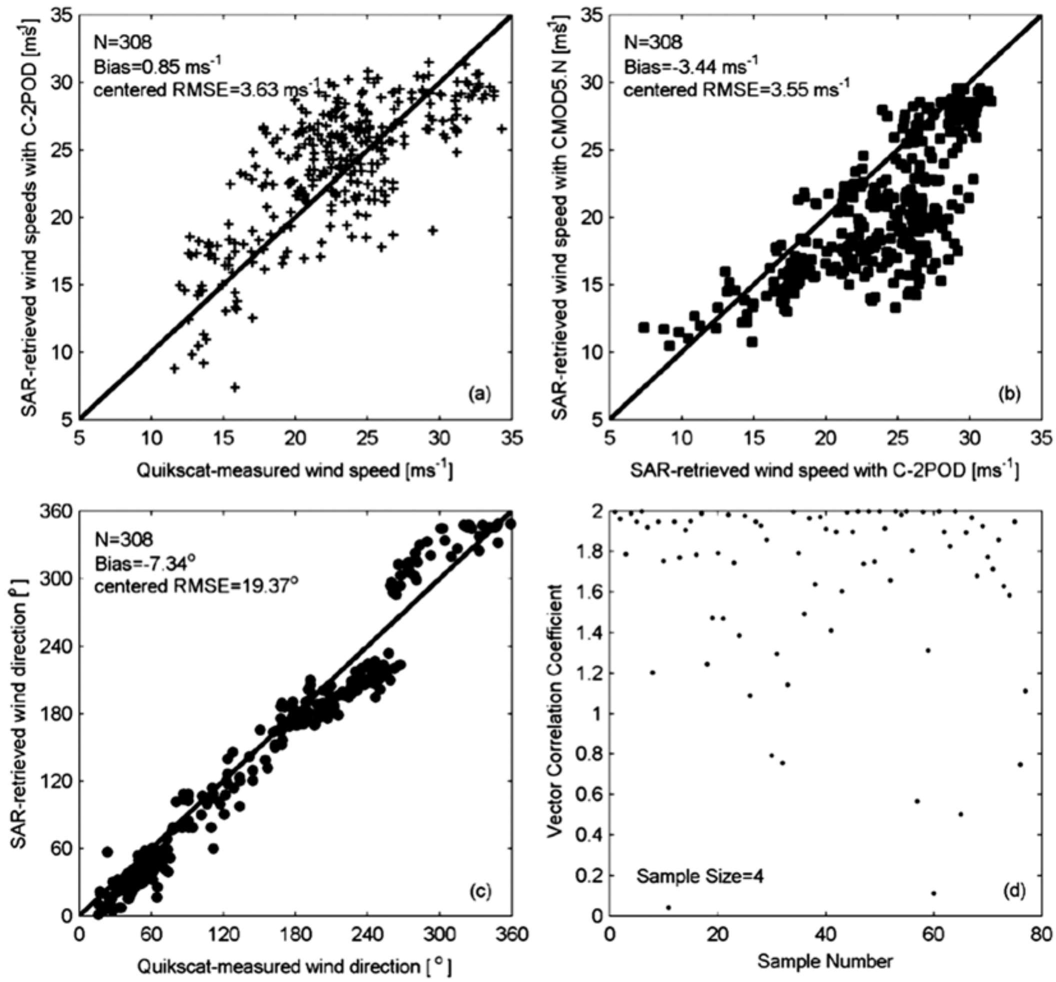

Fig. 9. (a) SAR-retrieved wind speeds from the C-2POD model vs QuikSCAT-measured wind speeds, (b) SAR-retrieved wind speeds from the CMOD5.N model vs from the C-2POD model, (c) SAR-retrieved wind directions vs QuikSCAT-measured wind directions, and (d) vector correlation of wind directions from SAR and QuikSCAT (sample size is 4). Hurricane Bill winds from SAR and QuikSCAT are acquired at 2226 and 2254 UTC 22 Aug 2009, respectively.

The revised manuscript will tone down the assertion that our model compares favorably with observations and will state that it compares favorably with recently published work.

Tan, C. and Fang, W.: Mapping the wind hazard of global tropical cyclones with parametric wind field models by considering the effects of local factors. International Journal of Disaster Risk Science, 9(1), 86-99, https://doi.org/10.1007/s13753-018-0161-1, 2018.

Zhang, B., Perrie, W., Zhang, J.A., Uhlhorn, E.W. and He, Y., 2014. High-resolution hurricane vector winds from C-band dual-polarization SAR observations. Journal of Atmospheric and Oceanic Technology, 31(2), pp.272-286.

We agree that our use of a 20% correction factor for winds over urban areas is weak. On further consideration, and in response to similar concerns from other reviewers, we decided to remove this bias correction factor from this study. The original factor of 20% was determined by comparing our simulations with surface station data in urban areas for a subset of 8 landfalling

U.S. hurricanes. This bias correction step was added to aid application of the dataset but clearly patches over an underlying problem.

Holmes (2007) found that the roughness length for urban areas can vary between 0.1 and 0.5m for suburban regions and rise to between 1 and 5m for densely packed high-rises in urban centers. Our model uses a single roughness length for all urban areas (suburban and city centers) of 0.8m and this was taken from the MODIS land use dataset – the same as used in the Weather Research and Forecasting (WRF) model. This value is too high for suburban areas, where a value closer to 0.2 is typical (Yang et al. 2014). Depending on the specific siting of the wind observing stations, it's probable that the introduction of multiple urban categories with different roughness lengths would improve our low wind speed bias. This detailed investigation is beyond the scope of this paper and we choose to leave this for future work.

Holmes, J. D., 2007. Wind loading of structures. 2nd ed. London and New York, Taylor & Francis

Yang, T., Cechet, R.P. and Nadimpalli, K., 2014. *Local wind assessment in Australia: Computation methodology for wind multipliers*. Geoscience Australia. 2014/33.

The Challenge of a Global Approach and the Expected Benefits

The problem with developing tools for global application unfortunately means that accuracy is inevitably impacted by the need to adopt spatially and/or temporally compromised globally available datasets. This situation limits, and actively dissuades, examination of the myriad of fine scale site-specific influences on the TC surface wind during a specific event that cannot be ignored. These have traditionally been transparently handled by reference to standard exposure and application of statistically based boundary layer turbulence approaches. The authors' more complex and computationally demanding approach needs to demonstrate at least a comparable utility.

Please see our comment above. Our approach shows comparable utility to a recently published globally applicable approach that combines parametric wind profile modeling with local wind multiplication factors.

In any case, it is not clear what practical application there is in producing such a global (deterministic) event set, given that the essential need is for risk management that implicitly requires a probabilistic approach. Cherry-picking of historical events does not yield firm statistical guidance and any such results will most likely be less reliable than any regional wind speed risk assessment that is based on (even sparse) long term data sets, given that aggregation of sites is often justified. In spite of the practical challenges in this space, the various engineering design standards around the world have overtime assembled realistic and likely suitably conservative wind risk frameworks, all of which embody the need to allow for height, terrain and topography effects.

A deterministic event set is of significance importance to the risk management industry for two reasons. Firstly, it allows for the validation of the wind field module of tropical cyclone catastrophe models (probabilistic approach). This module is a critical component of these models, and they do not use a physical model for representing the wind speed and rely on site coefficients to represent friction and topography. The module is used to generate stochastic wind fields and historical wind fields for a selection of events. A global deterministic event set based on a physical model is therefore an excellent tool for validating catastrophe model's wind module via a comparison of historical wind footprints. Secondly, modelling as-if losses from historical storms is a widely used approach to stress test reinsurance structures to ensure companies have adequate protection, and can also be used in submissions to regulators. A global event set is an excellent basis for building scenarios to test as-if losses. These points will be emphasized in the revised manuscript.

The model's development is touted as valuable for insurance-related purposes. However, the principal cause of increasing world-wide losses for insurers is, together with uninformed risk-based planning, the failure to implement known good design, construction and inspection practices for residential development. The importance of globally modelling large scale terrain influences is also overstated given that, compared with typically nearly-flat or undulating conurbations, there is negligible insurance exposure to wind hazard in areas of very high or steep terrain.

We fully agree that one of the principal causes of increasing losses is poor implementation of known good construction practices. Indeed, Simmons et al. (2018) quantified the importance of a strong and well enforced wind building code for reducing insured wind losses in Florida.

There is an industry need for ever more accurate modelling of risk, often for single sites. Representing terrain influences is well known to influence wind speed, and mountainous countries often have both a high risk of tropical cyclones and significant insurance exposure, for example the Philippines and Japan. Therefore we believe an accurate representation of the influence terrain is very important, it is also included in all recent respected proprietary catastrophe models. This point will be made in the revised manuscript.

Simmons, K.M., Czajkowski, J. and Done, J.M., 2018. Economic effectiveness of implementing a statewide building code: the case of Florida. Land Economics, 94(2), pp.155-174.

Comments on Method

(2.) The step that "removes an estimate of the asymmetry due to storm motion" from the surface wind relies on the assumption that historical Vmax do reliably include such an influence. While Dvorak, for example, implies that is the case there is no specific allowance in the methodology. Hence the adopted empirical adjustment likely has little merit in terms of overall accuracy.

This is an important point. We agree that the extent to which asymmetry due to forward speed is included in best track estimates of Vmax may be questionable. But given that these winds are Earth relative measurements we assume that the benefits of removing an uncertain estimate of asymmetry outweighs the cons of not doing so. In this assumption, we follow the approach of others (e.g., Chavas et al. 2017). This point is made in the revised manuscript.

Chavas, D. R., Reed, K. A. and Knaff, J. A.: Physical understanding of the tropical cyclone wind-pressure relationship. Nature communications, 8(1), p.1360, https://doi.org/10.1038/s41467-017-01546-9, 2017.

In quoting Harper, Kepert and Ginger (2010) (aka the WMO wind averaging guidelines - hereafter HKG), the assumption that numerically modelled winds calculated at a small timestep are representative of so-called 1-min sustained winds is incorrect. Section 1.6 of HKG specifically advises on that topic noting that numerical models without explicit eddy representation only estimate mean wind speeds, not gusts such as the so-called 1-min sustained wind. However, correcting for that (e.g. per HKG Table 1.2) will likely have no specific effect on the model performance.

Our numerically modeled winds are calculated on a timestep of a few seconds. This means that the model can only resolve wind variations of about 4 to 7 times the model timestep (depending on the variability of the flow). The instantaneous model outputs are therefore by no means the instantaneous wind. Rather, they are closer to the 1-minute average wind. It is true that our numerical modeling does not explicitly resolve turbulence. It parametrizes eddies. We agree with section 1.6 of HKG that the outputs of numerical models without explicit turbulence should be considered as the mean wind. And we consider the 1-minute wind to be a mean wind speed and not a gust measure.

(2.1) The authors state that the model is "agnostic to the source of the track data" as though that is some advantage, whereas the vast majority of historical datasets consist only of (lat, lon, Vmax) estimates with acknowledged high variability between agencies and also over-time. Rmax is also noted to be an essential parameter but is only recently available in some regions and not transparently derived. While these drawbacks are unavoidable, it further emphasises the challenge of using any (global) historical track data without critical assessment and expecting a high level of accuracy in the estimation of terrain-sensitive surface winds.

The cyclone-scale footprint will only be as good as the input best track data and the wind profile model. We also agree that details of how parameters such as Rmax are derived in historical track datasets are not transparent and are likely to vary wildly through time and between agencies. We will remove the statement that the model is "agnostic to the source of the track data" because while the model will function using different data sources, the accuracy of the cyclone-scale footprint will be very sensitive. This point will be made in the revised manuscript.

(2.2) While the Willoughby profile may be superior to some others for a hands-off global application it still requires an outer scale assumption and to note that TC scale, and its temporal evolution, is a critically important parameter in accurately modelling surface winds. The adopted land use surface roughness, with a scale of about 1 km, seems reasonable enough but should also be verified by example for the set of modelled storms. To note also that the references cited for drag coefficients are very dated and the authors could adopt more recent evidence to better suit their argued approach.

In response to this comment and a similar comment from another reviewer, the revised manuscript will better highlight the assumptions needed to model the outer winds of the Willoughby profile. Firstly, we now detail that the length scale for the transition region across the eyewall is set to 25km when Rmax is greater than 20km and set to 15km otherwise. For the shape of the vortex outside Rmax, we hold the faster decay length scale fixed at 25km, following the recommendation of Willoughby et al. (2006), and allow the second length scale (*X1*) and the contribution of the fast decay rate (*A*) to vary the shape of the profile. For the slower decay length scale, we now include note that Willoughby et al. (2006) show a wide range of the second length scale occurs in nature, with their Fig. 11 showing it can range from about 100 km to over 450 km. We therefore choose to allow *A* and *X1* to vary with the readily available parameters Rmax, Vmax and latitude, following Eqn. 11 in Willoughby et al. (2006).

The effects of the adopted land use surface roughness is included in the evaluation of the subset of U.S. storms shown in Fig. 4.

We will include an updated reference for the variation in surface drag over the ocean with wind speed. And we choose to keep the key early reference to Garratt (1977).

Comments on Results and Evaluation

(3.) The Puerto Rico example illustrates the intention and ability of the model to introduce terrain and topographic variability into its results but, without any verification, is otherwise meaningless and simply an "artist's impression".

We agree with the need to compare with available observations. This Puerto Rico example will be expanded (also in response to comments from other reviewers) to better demonstrate the effects of adding KW01, surface roughness, and topography. This expanded analysis will include comparison with the available surface station data.

(4.) To note that the use of 3-h sampled wind data is typically inappropriate for TC passages within, say, 100 km and may be a principal source of the poor comparisons. Application of HKG Table 1.1 is also dubious given the 3-h sampling but more so because of its limited and nominal exposure classes, which appear inconsistent with the aim of deriving fine scale surface winds. HKG Section 1.2 says "The aim has been to provide a broad-brush guidance that will be most useful to the forecast environment rather than a detailed analytical methodology" and "In particular, post analysis of TC events should seek to use the highest possible site-specific analytical accuracy for estimating local wind speeds. This would include consideration of local surface roughness, exposure and topographic effects when undertaking quantitative assessments of storm impacts." This implies an approach like Powell et al. (1996) is needed in such cases. Again, these oversights in applying HKG will likely have little effect on model performance, but do point to a lack of rigour in matters of wind magnitude adjustment.

Thank you for this comment. We apologize that made a mistake in the text. In the original manuscript we incorrectly stated that "points are shown for all instances of the observed time falling within 3 hours of the model time." This didn't make sense, and will be corrected to "comparisons between observations and model are made using model time within 5 minutes of the observation time". We were able to do this because we output the model fields every 10 minutes.

We agree that the application of HKG Table 1.1 to scale the observations to the 1-minute wind is limited by the nominal exposure classes. But we also don't have the resources to conduct a detailed investigation into the specific exposure of each surface observing station. We therefore choose to leave this as an uncertainty in the evaluation. Over many surface observing sites over many landfalling events this uncertainty should approach zero if the effect is random. Though we agree that there is a clear need to conduct detailed site-specific comparisons in future work. This point will be made in the revised manuscript.

As noted previously, the demonstrated performance of the model is very poor compared with the numerous less-complex examples that are cited. In spite of the authors' tendency to downgrade the utility of H*WIND I would instead encourage pursuing such comparisons in order to locate the model deficiencies and improve its performance for the demonstration storms.

In our previous response we made the case that our approach is at least comparable to a recently published globally applicable model. It was not our intention to downgrade the utility of H*WIND. H*WIND has many benefits over our approach such as including the effects of storm-scale asymmetry unrelated to forward speed. We will rephrase the references to H*WIND accordingly. Perhaps the only major disadvantage of H*WIND is that data are only available for a subset of U.S. storms.

One challenge of evaluating our model using H*WIND is the need for assumptions for how to convert our model wind speeds to the open terrain representation in H*WIND. While a comparison against H*WIND would be useful to explore the scales of variability that are not included in our modeling approach, we choose to restrict our evaluation to surface station data.

Comments on Global Landfalling TC Footprints

(5) The detailed commentary and interpretation of aspects of modelled landfall and inland wind decay characteristics in various localities seems to overlook the fact that the model is only reacting to the imposed "best track" intensity variation and therefore can have no better skill than offered by a simpler parametric approach. Surely fully dynamic modelling (e.g. HWRF or similar) is needed to reliably explore such impacts.

Our modeling approach only allows for one-way forcing by the gradient wind and pressure profiles on the tropical cyclone boundary layer. Possible upscale effects of local terrain back on the entire TC structure is not permitted and is therefore missed. For example, studies have shown that the TC track itself can be impacted by high mountain ranges (see the discussion in Ramsay and Leslie 2008). Our approach contains these effects to the extent they are contained in the input best track data. It is correct that our gradient winds are reacting to the imposed best track, but the winds at 10m above the ground are reacting to the local terrain. We will state this separation of effects more clearly in the revised manuscript.

We agree that NWP models would capture more of the terrain effects and therefore would be better suited to study terrain effects on the inland wind decay. But one of the goals of our study is to explore the effects across a global dataset, and generating this dataset would be computationally impractical for 2km resolution NWP modeling.