17D.10

Convective-Scale Simulations of Tropical Cyclones with the Met Office Unified Model

Chris Short*
Met Office, Exeter, UK

1 Introduction

The Met Office (MO) has recently entered into a partnership with the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) to develop a complete modelling, forecast, guidance and impacts service for the Philippines.

Much of the hazardous weather affecting the Philippines is related to tropical cyclone (TC) activity in the area. The operational global configuration of the MO Unified Model (UM) is currently producing better TC predictions than ever before (Heming, 2016). However, the spatial grid is relatively coarse (approximately 17km at mid-latitudes) which prevents it from being able to resolve the sharp gradients in the inner regions of storms, necessary for an accurate representation of TC intensity and structure (e.g. Fierro et al. 2009 and references therein).

The objective of this work is to investigate the added value of using a higher resolution (4.4 km grid length) limited-area configuration of the UM for TC forecasts in the Philippines region.

2 Numerical details

2.1 Model configurations

The 4.4 km limited-area configuration of the UM used here employs the latest MO dynamical core, ENDGame (Wood et al., 2014), the blended turbulence parameterisation scheme, a 'grey zone' convection parameterisation (in which shallow convection is parameterised but mid and deep convection are treated explicitly), and scale-aware warm rain microphysics (all of which are described in Boutle et al. 2014). The vertical level set used has increased resolution in the upper troposphere relative to the level set used in MO mid-latitude regional models, so that tropical deep convection is better resolved. This model is given the label P4.4.

The P4.4 model is a so-called down-scaler (i.e. no data assimilation), where initial and boundary conditions are derived from a driving model. The driving model in this case is the MO global model, using the current operational configuration, Global Atmosphere (GA) 6.1 (Walters et al., 2016). The global model is referred to as G26 hereafter.

The key features of the P4.4 and G26 model configurations are summarised in Table 1.

2.2 Trial process

Rigorous testing of model TC predictions requires a large number of storm cases. For this purpose, the MO has recently developed a regional modelling system for TCs that utilises an automatically re-locatable domain, capable of simulating storms anywhere in the World, on demand.

The re-locatable system has been used to re-run all storms that developed in the Northwest Pacific basin between 1st August and 30th November 2013 with the P4.4 model configuration. The Northwest Pacific basin was particularly active during this time with 22 TCs in total, 6 of which made landfall in the Philippines, including the devastating Typhoon Haiyan.

The system is cycled every 24 hours (at 00Z), yielding multiple forecasts for each storm. The resulting number of TC forecasts is shown as a function of lead time in Figure 1.

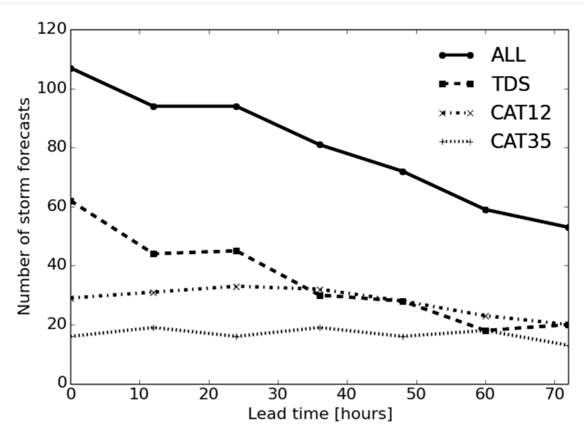

Figure 1: Total number of storm forecasts as a function of lead time. The various line types correspond to the different intensity categories described in the text.

*Corresponding author address: Chris Short, Met Office, FitzRoy Road, Exeter, Devon, EX1 3PB, UK; email: `christopher.short@metoffice.gov.uk`

Model name	Type	Convection	Boundary-layer scheme	Time step	Grid spacing	Number of levels	Model top
P4.4	Regional	Explicit	Blended	100 s	4.4 km	80	38.5 km
G26	Global	Parameterised	1D non-local	600 s	26 km	70	80 km

Table 1: Basic details of the the regional and global configurations of the Unified Model used in this study.

The various different lines in Figure 1 show the number of forecasts when storms are grouped into three intensity categories: tropical depressions and storms (TDS), category 1-2 (CAT12) and category 3-5 (CAT35) systems. At a given lead time, the category a storm is assigned to is determined by its *observed* peak wind speed at that time [1]. If the wind speed is less than 64 knots, between 64 and 95 knots, or greater than 95 knots, the storm is assigned to the TDS, CAT12 or CAT35 category, respectively. A storm will typically move between these different intensity categories during a forecast.

Verification statistics for all runs were produced in post-processing using the MO TC tracking software (Heming, 2015). To construct a homogeneous sample from the P4.4 and G26 model runs, statistics are matched on storm, forecast initialisation time and lead time.

3 Evaluation of model performance

3.1 Intensity forecasts

Figure 2 shows the mean error in TC central pressure as a function of forecast lead time for the P4.4 and G26 models, relative to IBTrACS observations.

The initial central pressure is higher than observed (∼60 hPa error for CAT35 systems), signifying that TCs are too weak in the global model analysis. Recent work at the MO has sought to address this by including central pressure estimates from TC warning centres in the global model data assimilation (DA) cycle; see Section 4.

Considering the entire sample, the 4.4 km model predicts a steady decrease in central pressure error with lead time. The bias actually becomes negative beyond $T + 48$, implying lower central pressures than observed, on average. A likely cause of this over-deepening is the lack of ocean feedback on the atmospheric circulation in the model. By contrast, the global model error remains approximately constant throughout the forecast. The magnitude of the pressure bias is always smaller in the P4.4 model than the G26 model, demonstrating the added value of the higher resolution model.

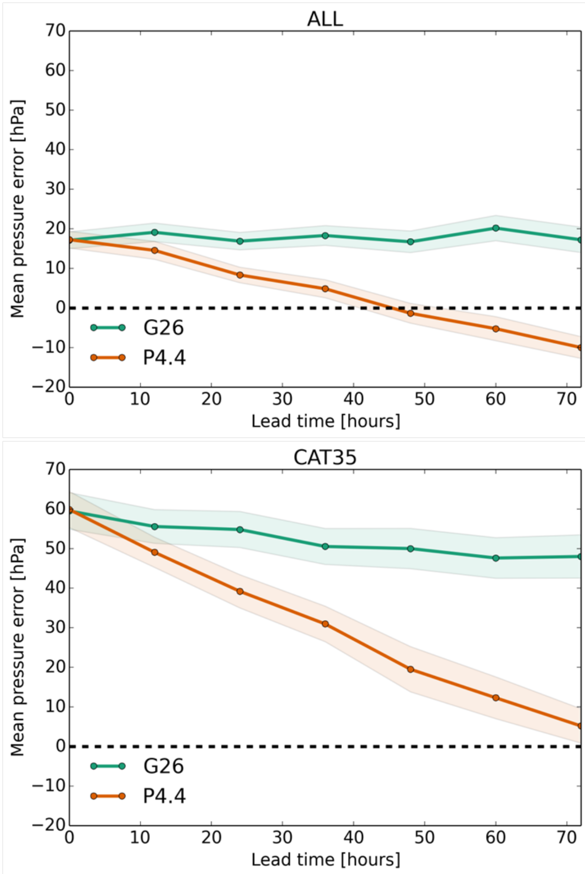

Figure 2: Mean error in forecast central pressure as a function of lead time for the P4.4 and G26 models (solid lines). The shaded regions correspond to one standard error of the mean. The top panel is for the full sample of tropical cyclones and the bottom panel is for category 3-5 systems only.

From an operational perspective, being able to accurately predict the intensities of the most severe (CAT35) storms is most important since these have the greatest destructive potential. For these systems, the P4.4 model is clearly far superior to the G26 model.

Switching focus to TC winds, Figure 3 shows the wind-pressure relation for the P4.4 and G26 model configurations. The observed relation is also shown, constructed from IBTrACS data for the storms in the sample.

The 4.4 km model offers a considerable improve-

[1] The observed peak wind speeds are 1-minute mean surface wind speeds from the International Best Track Archive for Climate Stewardship, IBTrACS (Knapp et al., 2010).

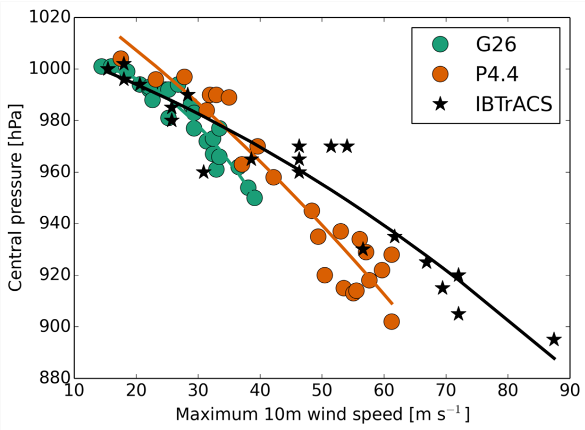

Figure 3: Wind-pressure relation for all storms in the sample. Data from the P4.4 and G26 model configurations are shown, along with observations from IBTrACS. The solid lines are second-order polynomial fits to the data points.

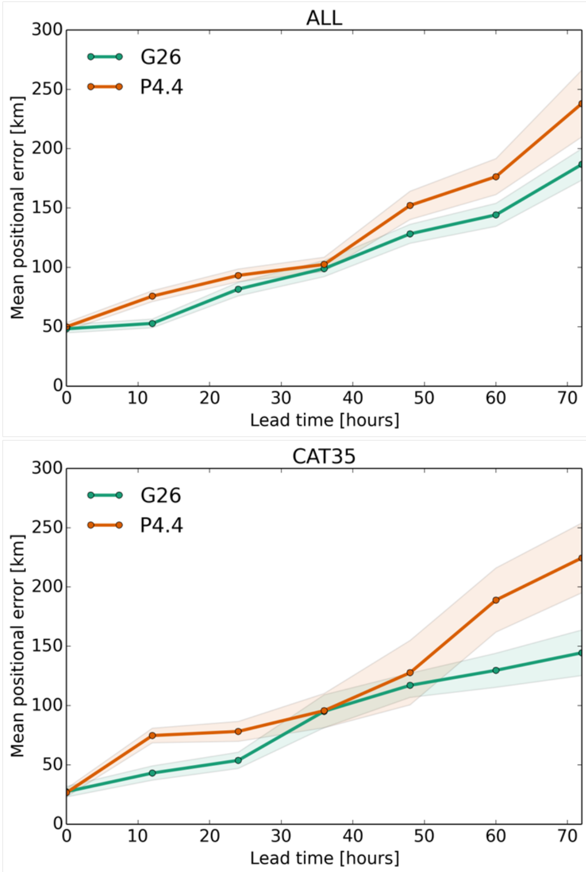

Figure 4: Mean error in forecast storm position as a function of lead time. The layout of the panels is the same as in Figure 2.

ment over the global model, because it begins to populate the lower-right corner of the plot, where the most intense systems lie. However, like the global model, it predicts a wind-pressure relation that is too steep compared to observations. In other words, the maximum surface wind speed for a given central pressure is too low. One possible reason for this is that UM surface fluxes are in error at high wind speeds (the impact of surface fluxes on storm intensity has been investigated by Green and Zhang 2014, for example).

3.2 Track forecasts

Figure 4 displays the mean error in the forecast position of TCs given by the P4.4 and G26 models, relative to IBTrACS observations.

It is clear that, on average, the P4.4 model yields larger errors in TC position than the global model, especially for the most intense storms. However, the null hypothesis that the mean difference in track error between the P4.4 and G26 models is zero cannot be rejected at the 0.05 significance level, when serial correlation between the successive forecasts for each storm is accounted for. A larger sample is required to establish whether track predictions are indeed systematically worse in the regional model.

Nonetheless, it is still interesting to ask why the P4.4 model gives less accurate track forecasts than the G26 model for the current sample. There are many possible reasons. For example, differences in the distribution of diabatic heating between the models, which could easily arise due to their different treatments of convection. Asymmetries in convective heating are well known to modulate TC motion (Wu and Wang, 2001; Fovell et al., 2010; Cao et al., 2011). There are also differences in the cloud microphysics and radiation parameterisations used in the models, both of which also influence TC tracks (Fovell and Su, 2007; Fovell et al., 2008, 2010). Investigative work to determine the source of the larger track errors in the P4.4 model is in progress.

3.3 Precipitation forecasts

Figure 5 shows radial profiles of the azimuthally-averaged precipitation rate in TCs for the P4.4 and G26 models, 2 days into the forecast and averaged over all storms in the sample. For reference, a corresponding profile derived from the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B42 (Huffman et al., 2007) product is also shown. Before averaging, all model data has been up-scaled to the coarser TMPA grid using an area-weighted re-gridding scheme.

The main point to note is that, relative to TMPA, both the P4.4 and G26 models significantly

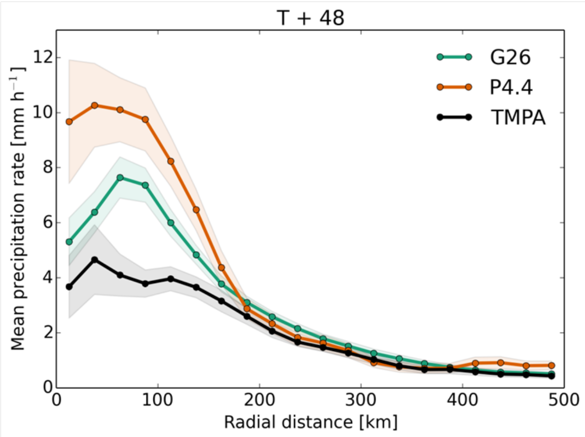

Figure 5: Mean precipitation rate in tropical cyclones as a function of radial distance from the storm centre. Predictions from the P4.4 and G26 models are shown, along with the corresponding observed mean rain-rate profile from TMPA. Shaded regions are one standard error of the mean. All $T+48$ storm forecasts have been included.

over-estimate the rain rate in the central regions of TCs ($r \lesssim 150$ km). This bias is larger in the 4.4 km model than the global model.

The statistical distribution of precipitation rates is illustrated in Figure 6, which shows the relative contribution of different rain rates to the total rainfall in TCs at a lead time of $T + 48$.

It is evident that, relative to TMPA, the P4.4 and G26 models predict a distribution of rain-rates that is skewed towards high values, significantly under-estimating the contribution from moderate rain and over-estimating the heavy rain, particularly the P4.4 model. The global model is able to reproduce the observed contribution from light rain, but the P4.4 model under-estimates it. Both models give very little extreme rain, consistent with the satellite data.

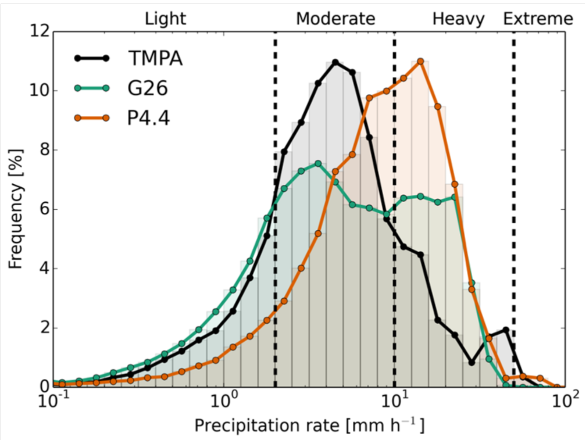

Figure 6: Histograms showing P4.4 and G26 model predictions for the relative contribution of different rain rates to the total rainfall in tropical cyclones. All $T+48$ storm forecasts have been included. A corresponding histogram derived from observational data (TMPA) is shown for comparison. The vertical dashed black lines demarcate the boundaries between different rainfall regimes: light (<2 mm h^{-1}), moderate (2-10 mm h^{-1}), heavy (10-50 mm h^{-1}) and extreme (>50 mm h^{-1}).

As a final comment, it has been known for some time that limited-area configurations of the UM generate excessive precipitation amounts in tropical regions. Recent work has revealed that this is due to spurious moisture production by the semi-Lagrangian dynamical core of the model. A potential solution based on the mass conservation scheme described in Aranami et al. (2015) is currently being tested, which has been shown to improve moisture conservation in regional models. It is expected that this would reduce the overall amount of rainfall in TCs and shift the peak of the histogram shown in Figure 6 towards lower rain-rates. This will be investigated in future work.

4 Improving the initialisation of tropical cyclone forecasts

In the MO global analyses used to initialise the P4.4 model, TCs are too weak and there can be significant positional errors (recall Figures 2 and 4). In fact, the largest errors in intensities occur at the initial time. Regional down-scalers can thus be handicapped from the very start of a forecast.

In a bid to improve the representation of TCs in global analyses, central pressure estimates from TC warning centres are now ingested as part of the MO global model DA cycle. Since entering operations, this method - known as TCCP - has proved beneficial to global model TC predictions, improving both track and intensity (Heming, 2016).

To test how this improved TC initialisation in the global model affects regional model forecasts, another trial has been run that is a subset of the four-month period considered previously in this report. This shorter trial spans the period from 26th September to 12th November 2013, which includes 14 TCs in the Northwest Pacific basin. For each case there are two P4.4 model runs, starting from global analyses generated with (trial) and without (control) TCCP, respectively. Despite the shorter trial period, the number of cases is comparable to that shown in Figure 1 because the cycling frequency of the re-locatable system was increased to every 12 hours (at 00Z and 12Z).

Note that, for the TCCP trial, the resolution of the global model was increased to match that now used operationally at the MO. The horizontal grid spacing is approximately 17 km at mid-latitudes, so this model is thus given the label G17. The G17 model is otherwise identical to the G26 model.

4.1 Impact on track forecasts

Figure 7 displays the mean error in the forecast position of TCs (the direct positional error, DPE) as a function of lead time for the control and TCCP runs with the P4.4 and G17 models.

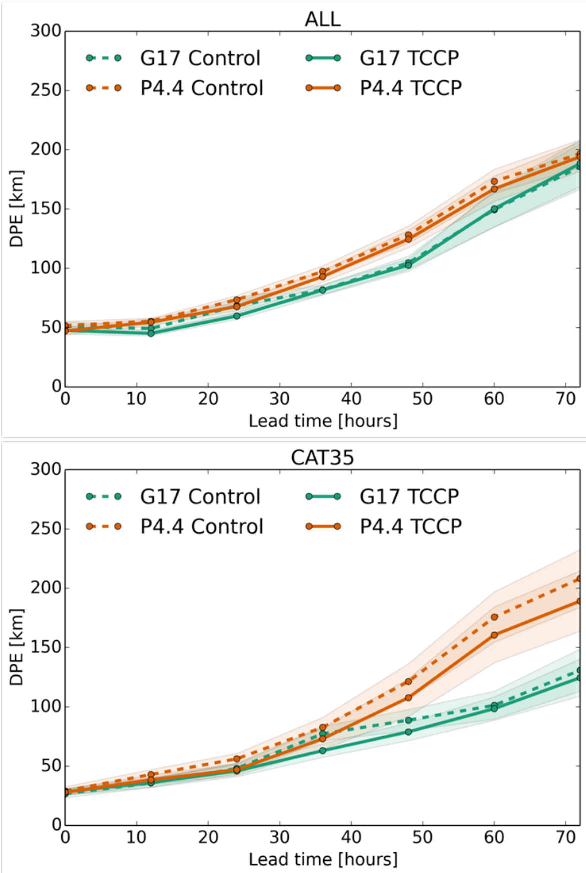

Figure 7: Mean error in forecast storm position as a function of lead time for the P4.4 and G17 models, starting from global analyses with (solid lines) and without (dashed lines) the assimilation of central pressure estimates from TC warning centres. The shaded regions are one standard error of the mean. The top panel is for the full sample of tropical cyclones and the bottom panel is for category 3-5 systems only.

It is evident that employing the TCCP scheme in the global model has a positive impact on P4.4 model track forecasts. Errors are reduced most for CAT35 storms (by 10% when averaged over all lead times). This is an encouraging step in the right direction. However, more work is clearly required on both the DA and convective-scale modelling fronts to further narrow the gap between the regional and global models.

4.2 Impact on intensity forecasts

Figure 8 shows the mean bias in TC central pressure as a function of forecast lead time in the control and TCCP runs, for both the P4.4 and G17 models.

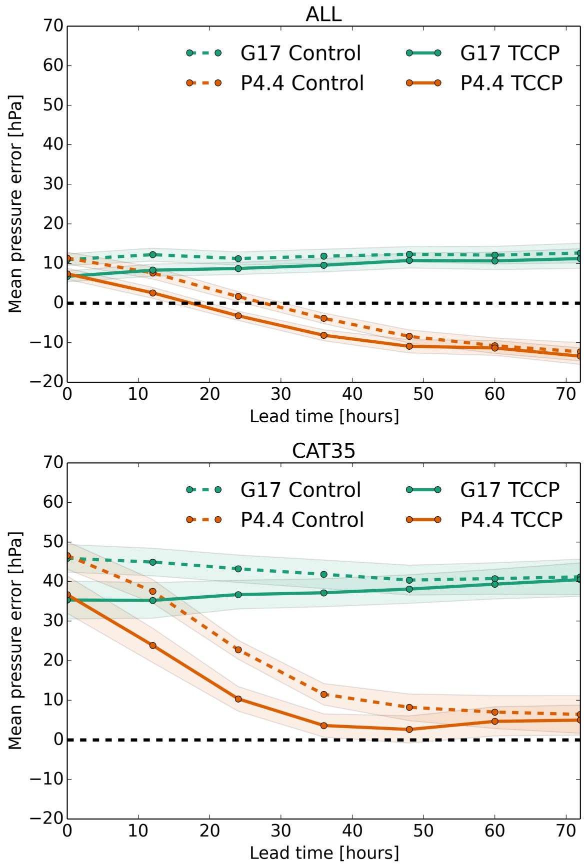

Figure 8: Mean bias in forecast central pressure as a function of lead time. The layout of the panels is the same as in Figure 7.

As one might expect, incorporating observational estimates for TC central pressures into the global analysis reduces the initial positive central pressure bias in the P4.4 model. Note that the mean bias in analysed intensity is still over 35 hPa for CAT35 systems though. The main reason for this is that central pressure observations were flagged and ignored by the global model DA for the most intense cases because the implied increments were too large. Clearly, future work should be concentrated on improving the initialisation of the most severe storms.

Focussing on the whole sample, the performance of the P4.4 model is improved in the first day or so when TCCP is included in the global model. However, as mentioned previously, the P4.4 model can over-deepen storms relative to observations. On average this happens earlier in the TCCP runs than in the control runs, leading to a larger bias in the middle stages of the forecast. At long lead times, the TCCP and control runs give very similar results. When averaged over all lead times, there is some cancellation of errors in the initial and middle phases of the forecast, with the result that TCCP offers a small improvement over control.

However, for the most intense storms, the P4.4 TCCP runs yield far better intensity predictions than the control runs, with a 19% reduction in the mean absolute error for central pressure when averaged over all lead times. Indeed, the mean bias in central pressure is less than 5 hPa for all lead times beyond $T + 36$, and does not go negative. Again, it is for these storms that the real value of the regional model over the global model is most apparent.

5 Summary

The aim of this work has been to examine how well a high-resolution (4.4 km grid length) limited-area configuration of the Met Office (MO) Unified Model can represent tropical cyclones (TCs) in the Philippines region.

Relative to the MO global model, the regional model adds value for TC forecasting, providing much better intensity predictions (both central pressures and surface wind speeds). As a direct consequence, the wind-pressure relation is also closer to that observed. The biggest improvement is seen in the most intense systems, which is important since it is these storms that are the most destructive. Storm structure is more realistic too, although this has not been discussed here.

Several key issues have been identified, which will be addressed over the coming years as collaborative work between the MO and the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) continues. These are:

(i) In the global analyses used to initialise high-resolution forecasts, the intensities of TCs are too weak and there are significant positional errors. Recent global model development has strived to address this by implementing a new data assimilation procedure - called TCCP - whereby central pressure estimates from TC warning centres are incorporated into the analysis. The TCCP method has proved beneficial to MO global model performance, yielding both better track and intensities. Here, it has been shown that the overall impact of the TCCP method on regional model TC forecasts is also positive. The improvement is most striking for intense storms, with a drop in mean track and central pressure errors of 10% and 19%, respectively. However, work remains to be done to further reduce errors at the initialisation time, especially for rapidly intensifying systems (see Heming 2016 for more discussion).

(ii) The 4.4 km model tends to over-deepen storms. This is probably because the model does not account for the reduction of sea-surface temperature caused by wind-driven oceanic mixing. There are plans to investigate this in future work with coupled atmosphere-ocean runs.

(iii) Regional model track forecasts are less accurate than the global model, with an overall increase of 19% in track error. Implementing the TCCP method helped reduce the gap between the 4.4 km and global models, so errors at the initialisation time may be part of the problem. The trials conducted in this study are not long enough to establish whether the larger track errors seen in the regional model are statistically significant. Nonetheless, it is an interesting result and work to understand their origin is currently underway.

(v) Compared to TRMM satellite observations, both 4.4 km and global models predict excessive rainfall in the inner regions of TCs ($r \lesssim 150$ km), with a distribution of rain-rates that is skewed towards higher values than observed. The regional model suffers from larger biases than the global model. However, it is expected that the imminent improvements to moisture conservation in the UM will benefit model predictions of TC rainfall.

References

K. Aranami, T. Davies, and N. Wood. A Mass Restoration Scheme for Limited-Area Models with Semi-Lagrangian Advection. *Q.J.R. Meteorol. Soc.*, 141:1795–1803, 2015.

I. A. Boutle, J. E. J. Eyre, and A. P. Lock. Seamless stratocumulus simulation across the turbulent gray zone. *Mon. Wea. Rev.*, 142:1655–1668, 2014.

Y. Cao, R. G. Fovell, and K. L. Corbosiero. Tropical cyclone track and structure sensitivity to initialisation in idealised simulations: A preliminary study. *Terr. Atmos. Sci.*, 22:559–578, 2011.

A. O. Fierro, R. F. Rogers, F. D. Marks, and D. S. Nolan. The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. *Mon. Wea. Rev.*, 137:3717–3743, 2009.

R. G. Fovell and H. Su. Impact of cloud microphysics on hurricane track forecasts. *Geophys. Res. Lett.*, 34:L24810, 2007.

R. G. Fovell, K. L. Corbosiero, and H. C. Kuo. Cloud microphysics impact on hurricane track as revealed in idealized experiments. *J. Atmos. Sci.*, 66:1764–1778, 2008.

R. G. Fovell, K. L. Corbosiero, A. Seifert, and K. N. Liou. Impact of cloud-radiative processes on hurricane track. *Geophys. Res. Lett.*, 37:L07808, 2010.

B. W. Green and F. Zhang. Sensitivity of tropical cyclone simulations to parametric uncertainties in air-sea fluxes and implications for parameter estimation. *Mon. Wea. Rev.*, 142:2290–2308, 2014.

J. Heming. Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. 2015. URL `http://www.metoffice.gov.uk/media/pdf/b/5/MetO_TC_tracking_and_verification.pdf`.

J. Heming. The impact on tropical cyclone predictions of a major upgrade to the Met Office global model. American Meteorological Society 32nd Conference on Hurricanes and Tropical Meteorology, 2016. URL `https://ams.confex.com/ams/32Hurr/webprogram/Manuscript/Paper292655/AMS32HURR17D.9.pdf`.

G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, 8:38–55, 2007.

K. R. Knapp, M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann. The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Amer. Meteor. Soc.*, 91:363–376, 2010.

D. N. Walters et al. The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *In preparation*, 2016.

N. Wood, A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat, and J. Thuburn. An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Q.J.R. Meteorol. Soc.*, 140:1505–1520, 2014.

L. Wu and B. Wang. Effects of convective heating on movement and vertical coupling of tropical cyclones: A numerical study. *J. Atmos. Sci.*, 58: 3639–3649, 2001.