P2.62

AN INTERCOMPARISON OF WRF-ARW AND JMA-NHM PERFORMANCE IN PREDICTION OF TROPICAL CYCLONES OVER THE SOUTH CHINA SEA IN 2008

S.T. Chan *, T.F. Chan and W.K. Wong
Hong Kong Observatory, Hong Kong, China

1. INTRODUCTION

Hong Kong is located in the sub-tropics region on the southeastern coast of China. During July to September every year, Hong Kong is most likely to be affected by tropical cyclones (TCs). On average, about 30 TCs form in the western North Pacific including the South China Sea every year, and about half of them reach typhoon strength (maximum winds of 118 kilometers per hour or more). Climatologically, Hong Kong is normally affected by around six TCs per year.

The Hong Kong Observatory (HKO) has been operating a mesoscale NWP modeling system called the Operational Regional Spectral Model (ORSM) since 1999. ORSM was adapted from the Japan Meteorological Agency (JMA). The model, configured at a finest resolution of 20 km, is formulated using the hydrostatic governing equations to provide numerical guidance for short-term weather prediction up to 3 days ahead. With a view to further enhancing its capability of severe weather prediction, HKO commenced experimental trials of the JMA Non-hydrostatic Model (JMA-NHM) (Saito *et al.* 2007) a few years ago. With the experience gained over the past few years, HKO is going to operate a new NWP system based on JMA-NHM starting from the rain season in 2010. The new system will consist of two forecast domains with horizontal resolution at 10 km and 2 km respectively (Wong 2010).

Meanwhile, HKO has also been applying the WRF-ARW (Skamarock *et al.* 2005) to provide ad-hoc support on high resolution site-specific forecasts. Both JMA-NHM and WRF-ARW possess sophisticated physical schemes to represent the convective and mesoscale processes in the prediction of high-impact weather such as TCs. In order to investigate the performance of JMA-NHM and WRF-ARW in the prediction of TCs over the South China Sea, the 10-km version of the two models were run to simulate six TCs which affected Hong Kong in 2008. The model predicted TC track, TC intensity and precipitation were used to assess the model performance. In the next section, the design of experiments is described. The model intercomparison results are provided in Section 3. A summary of the study and concluding remarks are given in Section 4.

2. DESIGN OF EXPERIMENTS

Six TC cases, namely Neoguri (0801), Fengshen (0806), Kammuri (0809), Nuri (0812), Hagupit (0814) and Higos (0817) which affected Hong Kong in 2008 were chosen in this paper. 72-hour integrations of both JMA-NHM and WRF-ARW were performed and the initial time of the model run for each of the six cases are as shown in Table 1.

Table 1. The initial time of the model run for the six TC cases.

TC Case	Initial time
Neoguri (0801)	12 UTC, 16 April 2008
Fengshen (0806)	00 UTC, 23 June 2008
Kammuri (0809)	12 UTC, 4 August 2008
Nuri (0812)	00 UTC, 20 August 2008
Hagupit (0814)	00 UTC, 22 September 2008
Higos (0817)	12 UTC, 2 October 2008

To facilitate a fair comparison as much as possible, the configurations and domains of JMA-NHM and WRF-ARW were set to be the same as each other. Fig. 1 depicts the coverage of the model domain. The initial and boundary conditions of both models were taken from the forecasts of JMA Global Spectral Model (GSM) (JMA 2007) with horizontal resolution at 0.5 degree in latitude and longitude. The sea-surface temperature (SST) was specified by the NCEP high-resolution daily SST analysis at 0.083 degree resolution. The circulations of the TCs in the model initial condition were interpolated from GSM analysis and no TC bogus was introduced in the model runs. Neither any specific tuning on the parameters of the physical processes was made. More details on the model configurations are given in Table 2.

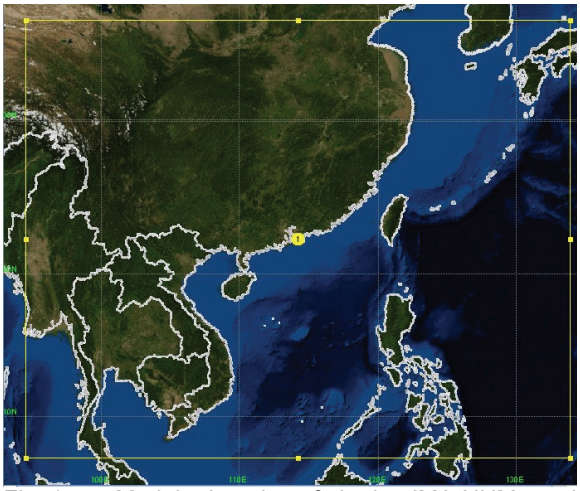

Fig. 1. Model domain of both JMA-NHM and WRF-ARW.

* *Corresponding author address:* S.T. Chan, Hong Kong Observatory, 134A Nathan Road, Hong Kong, China; email: stchan@hko.gov.hk

The TC track and intensity predictions were determined by the minimum sea level pressure every six hours. Apart from the JMA-NHM and WRF-ARW runs, predictions from GSM of JMA based on the 0.5-degree data were included for comparison. All predictions were verified against the best track data of HKO.

Table 2. Model configurations of WRF-ARW and JMA-NHM

Model version	**WRF-ARW Version 3.1**	**NHM Version 0902**
Horizontal resolution	10 km	
No. of grid points	401 x 321	
Vertical coordinates	Terrain-following	
No. of vertical levels	50	
Time step	40 seconds	
Forecast range	72 hours	
Initial/Boundary condition (BC)	JMA GSM (0.5 degree, BC interval: 6 hours) + NCEP SST (0.083 degree)	
Width of BC deployed (km)	50	240
Projection	Mercator	
Land Use	USGS Global Land Cover Characterization (GLCC) 30 second data and 24 land-use types	
Topography	USGS GTOPO 30 second data	
Microphysics	WSM 5	6-class bulk microphysics
Longwave Radiation	RRTM scheme	Based on GSM radiation scheme
Shortwave Radiation	Dudhia scheme	
Land Surface	5-layer Thermal diffusion scheme	4-layer Thermal diffusion scheme
Planetary Boundary layer	Mellor-Yamada-Nakanishi-Niino Level 3	
Cumulus Parameterization	Kain-Fritsch (KF) scheme	modified KF

For precipitation, a quantitative verification was conducted based on the NOAA CMORPH (CPC MORPHing) data which was taken as the ground truth. CMORPH is a technique which produces global precipitation analyses at a very high spatial and temporal resolution (Joyce *et al.* 2004). This estimated precipitation is derived from low orbiter satellite microwave observations including SSM/I (DMSP), AMSU (NOAA), AMSR-E (Aqua) and TMI (TRMM) exclusively, and whose features are transported via spatial propagation information obtained entirely from geostationary satellite IR data. The resolution of the CMORPH dataset is 0.25x0.25 degree and data for every three hours was available. Past verification showed that CMORPH yielded a high correlation of as much as 0.7 against rain gauge data. However, there was a tendency for CMORPH to underestimate rainfall over the tropical Pacific Ocean (Sapiano and Arkin 2009).

To align the differences in data resolution, both WRF-ARW and NHM outputs were regridded using bilinear interpolation to the same 0.25-degree grid of CMORPH for precipitation verification. Regions near the lateral boundaries were excluded from the verification to eliminate the influence of the boundaries, as depicted in Fig. 8(b) and (c). Using CMORPH and the regridded outputs of JMA-NHM and WRF-ARW, the bias score (BS), equitable threat score (ETS), and root mean squared error (RMSE) were calculated for day-1 (T+0 to T+24), day-2 (T+24 to T+48) and day-3 (T+48 to T+72) predictions against various precipitation thresholds.

3. RESULTS

3.1 *Track Prediction*

The mean track errors of JMA-NHM, WRF-ARW as well as GSM against the HKO best track data in the six TC cases are plotted in Fig. 2. The number of samples started to drop beyond T+48 due to TCs having made landfall and dissipated over land in the latter part of forecast period. Only three TCs remained active at T+72, namely Neoguri, Nuri and Hagupit.

Over the first 36 hours of integration, all three models exhibit similar skills in track prediction with JMA-NHM registering the smallest errors in general. Thereafter, both JMA-NHM and WRF-ARW perform apparently better than the global model. For example, the track error of GSM climbs above 200 km and almost reaches 300 km in the day-3 forecasts, whereas those of JMA-NHM and WRF-ARW stay below 180 km during the same period. The error of WRF-ARW grows at a comparatively slower rate and achieves the lowest errors among all three models in the later part of the forecast period. However, due to the small sample size, the observations above are not statistically significant.

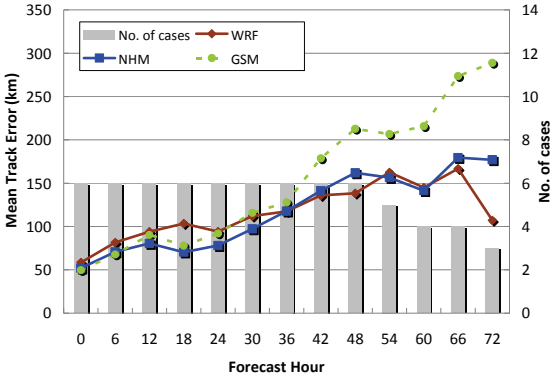

Fig. 2. Mean track errors of six TC cases.

The results suggest that both the non-hydrostatic models are quite competitive in track prediction at longer time scales when compared with GSM. As pointed out by Wang *et al.* (2006), the relatively poor initial conditions obtained from the coarse grid of the global model may have limited the benefits gained in the forecast skill from the high-resolution WRF-ARW model (and similarly from JMA-NHM). The deployment of a proper data assimilation system would likely improve the shorter term forecasts substantially.

3.2 *Intensity Prediction*

The intensity of the model forecasts were measured by the minimum mean sea level pressure (MSLP) at TC centre. It is noted that despite the use of TC bogussing in the analysis of GSM, the initial intensity of the TCs studied are in general substantially weaker than the best track intensity (Note: instead of using the full-resolution 0.1875-degree GSM data, 0.5-degree data were only available and used in the experiments). The average magnitude of intensity underestimation amounts to 12-13 hPa. This magnitude is also reflected in the mean absolute error plot at T+0 given in Fig. 3.

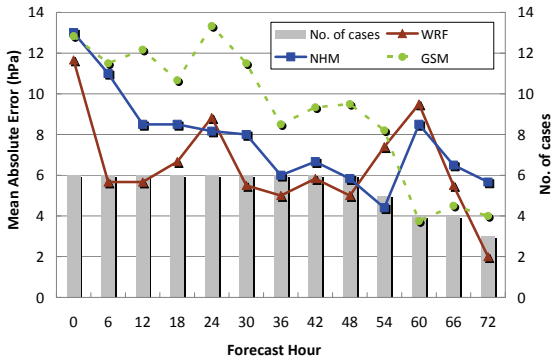

Fig. 3. Mean absolute errors of TC intensity forecasts.

As the model spun up, the model TC vortices gradually adjusted themselves, resulting in the gradual drop of intensity errors as the forecast hour progresses. Benefited from a higher resolution (10 km for JMA-NHM and WRF-ARW versus around 20 km for GSM) to facilitate a more realistic simulation of the TC core, the two non-hydrostatic models outperform GSM again in the first two days of forecasts. It is further noted that errors of WRF-ARW are lowest on average over almost the whole forecast period. The model spin-up of WRF-ARW is faster than JMA-NHM or GSM, especially during the first 6 hours of simulation.

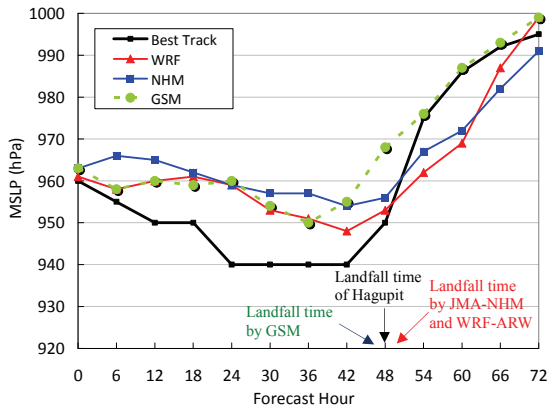

Fig. 4. MSLP forecasts by different models against HKO best track in the case of Hagupit (0814).

On the other hand, between T+60 and T+72, GSM outperformed JMA-NHM and WRF-ARW. Indeed, much of the errors of JMA-NHM and WRF-ARW can be attributed to the case of Hagupit in which the storm was forecast to weaken too slowly against actual, in part due to the delayed landfalling of the storm as predicted by the models (Fig. 4). On the other hand, GSM forecasted the storm to make landfall earlier than actual and generated exceptionally accurate intensity forecast in the last 12 hours of projection.

3.3 *Precipitation Prediction*

Figs. 5, 6 and 7 give the BS, ETS and RMSE against various precipitation thresholds for day-1, day-2 and day-3 forecasts from JMA-NHM and WRF-ARW respectively. BS equal to 1, less than 1, and larger than 1 indicate an unbiased forecast, under-forecast, and over-forecast, respectively. It is noted that BS of JMA-NHM forecasts for precipitation thresholds exceeding 100 mm/day increases rapidly above 1, suggesting the model tends to over-forecast the occurrence of heavy precipitation. This situation is particular obvious in the day-3 forecasts. Comparing with JMA-NHM, BS of WRF-ARW is closer to unity across nearly the whole range of thresholds except for light precipitation (5 mm/day or lower). A gross over-estimation of the light precipitation is apparent for WRF-ARW over the whole forecast range and an example is shown in Fig. 8(b) and (c) during Nuri (0812).

Similar observations of high BS for light precipitation were also noted by Hayashi *et al.* (2008) in the prediction of tropical precipitation by WRF-ARW in the Southeast Asia. This deficiency has resulted in the lower ETS at small thresholds achieved by WRF-ARW when compared with JMA-NHM. Ignoring those small precipitation thresholds at 5mm/day or below, WRF-ARW gets higher average ETS than JMA-NHM on day 1. The advantage over JMA-NHM narrows on day 2, and the trend seems to have reversed on day 3. By definition, ETS ranges from -1/3 to 1, with 0 indicating no skill and 1 being a perfect score. The verification results suggest that both JMA-NHM and WRF-ARW are skilful over different precipitation thresholds, though the skills have decreased considerably from day 1 to day 3.

Considering the daily precipitation amounts, WRF-ARW performs better across all 3 days with smaller RMSE than JMA-NHM. This might seem to contradict the previous observation of higher ETS achieved by JMA-NHM on day 3 but actually, high ETS can co-exist with high RMSE. A closer examination of the data reveals that Nuri has dominated the verification results on day 3. Fig. 8(c) shows the intense precipitation associated with Nuri predicted by JMA-NHM on day 3. The simulation by JMA-NHM matches closer against the CMORPH data [Fig. 8(a)] than that of WRF-ARW [Fig. 8(b)], despite the fact that JMA-NHM over-estimates the precipitation intensity by producing an extensive area of precipitation in excess of 300 mm/day not verified by CMORPH. Consequently, JMA-NHM has registered higher hit counts for a wide range of precipitation thresholds and hence the higher ETS, but at the same time the RMSE of the precipitation amounts would also be higher due to an over-estimation of the intensity of the heavy precipitation areas.

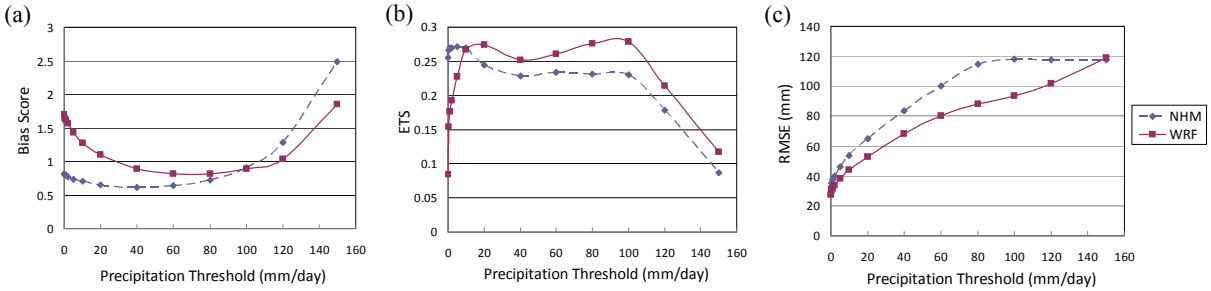

Fig. 5. (a) BS, (b) ETS and (c) RMSE of the day-1 precipitation forecasts from JMA-NHM and WRF-ARW.

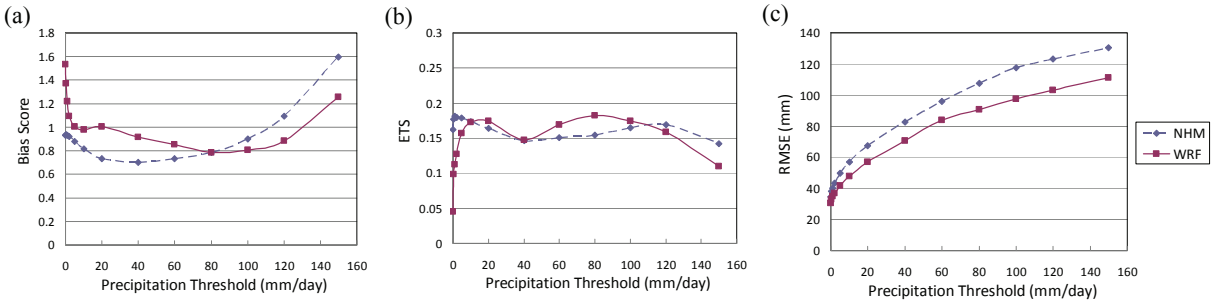

Fig. 6. Same as Fig. 5 except for day-2 forecasts.

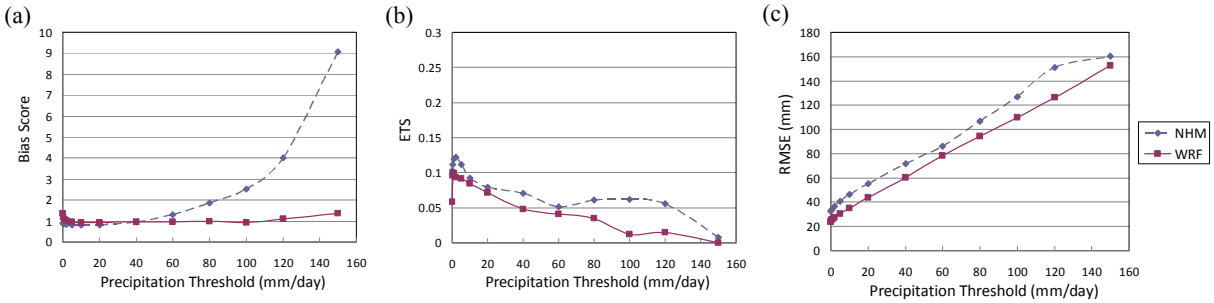

Fig. 7. Same as Fig. 5 except for day-3 forecasts.

Apart from the general quantitative verification of precipitation, the aggregated rain volume within the verification domain as well as the rain peaks, i.e. the maximum daily precipitation amount of the grids in the domain, were also studied as a function of lead time for each case. For the rain volume, the WRF-ARW generally predicted more rain especially during Neoguri, Fengshen and Hagupit [Fig. 9(a)], whereas JMA-NHM predicted less rain in all cases on day 1. As shown in Table 3, the differences against CMORPH amount to around +20% and -20% respectively. The greater rain volume predicted by WRF-ARW than JMA-NHM may partially be attributed to the rapid and faster spin-up of the former so that more precipitation can be generated in the early simulation stage.

Table 3. Average percentage error of model-predicted rain volume against CMORPH.

Model	Day 1	Day 2	Day 3	Overall
WRF-ARW	+20.6%	+2.6%	-2.3%	+7.3%
JMA-NHM	-19.5%	-10.9%	+3.9%	-9.3%

On day 2 and day 3, the total rain volume predicted by WRF-ARW match closely to CMORPH. Meanwhile, the negative bias in JMA-NHM forecasts decreases on day 2 and a slight over-estimation of precipitation even emerges on day 3.

For the rain peaks, both models produced generally higher daily precipitation amounts than CMOPRH on day 1 and day 2 [Fig. 9(a) and (b)]. Although WRF-ARW predicted more rain than JMA-NHM on these two days according to Table 3, the rain peaks predicted by WRF-ARW are mostly lower than those of JMA-NHM across all 3 days. WRF-ARW on average generated rain peaks closer to CMORPH, and it is especially so on day 3 [Fig. 9(c)].

The above observations concerning the better rain volume and rain peaks prediction by WRF-ARW than JMA-NHM are consistent with the verification results in terms of RMSE presented in previous paragraphs. A point to note here is that the simulation in the latter part of the forecast periods involve TCs in the weakening or dissipation stage when they came close to the land mass or after making landfall, making the verification results for precipitation not entirely objective but track dependent. Due to the small sample size, the verification results and associated conclusions for day 2 and day 3 should be treated with caution.

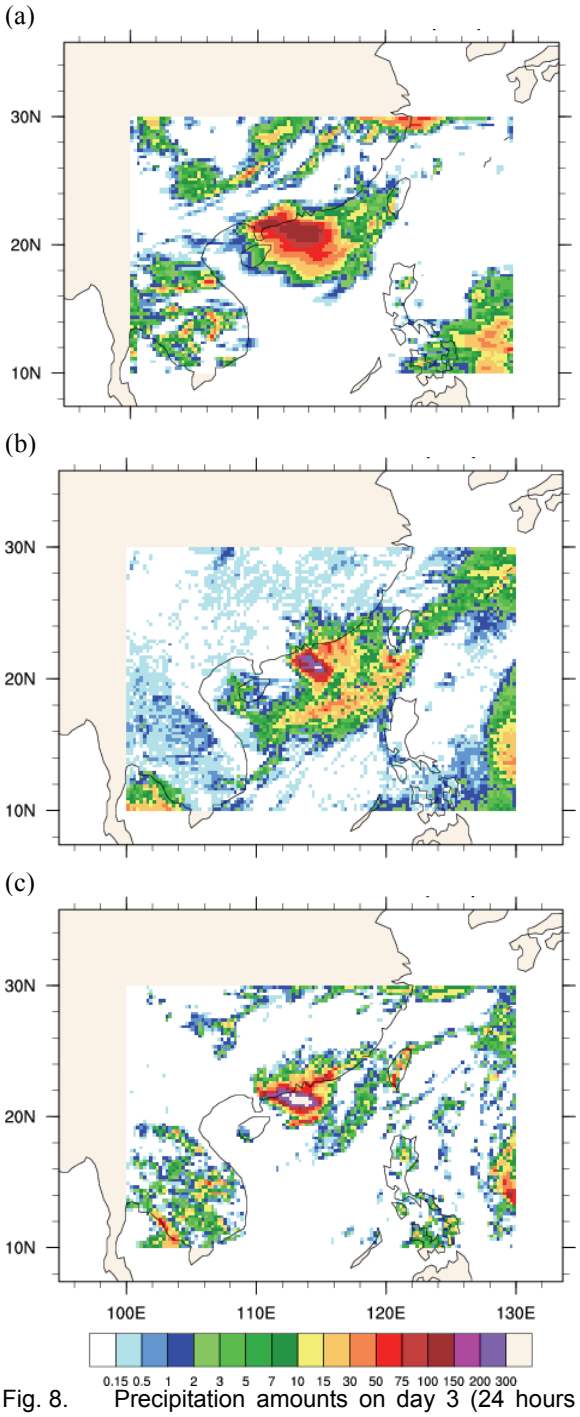

Fig. 8. Precipitation amounts on day 3 (24 hours ending 00 UTC on 22 August 2008) in the simulation of Nuri: (a) CMORPH, (b) WRF-ARW and (c) JMA-NHM.

3.4 *Computational Cost*

The modeling experiments were conducted on a Linux cluster, using 13 nodes with two single-core Xeon 2.8 GHz processors and 4 GB of memory each.. The average execution time for the 72-hour integrations of JMA-NHM and WRF-ARW of the TC cases are around 16 hours and 8 hours respectively. Apparently, WRF-ARW displays an advantage over JMA-NHM in terms of computation efficiency.

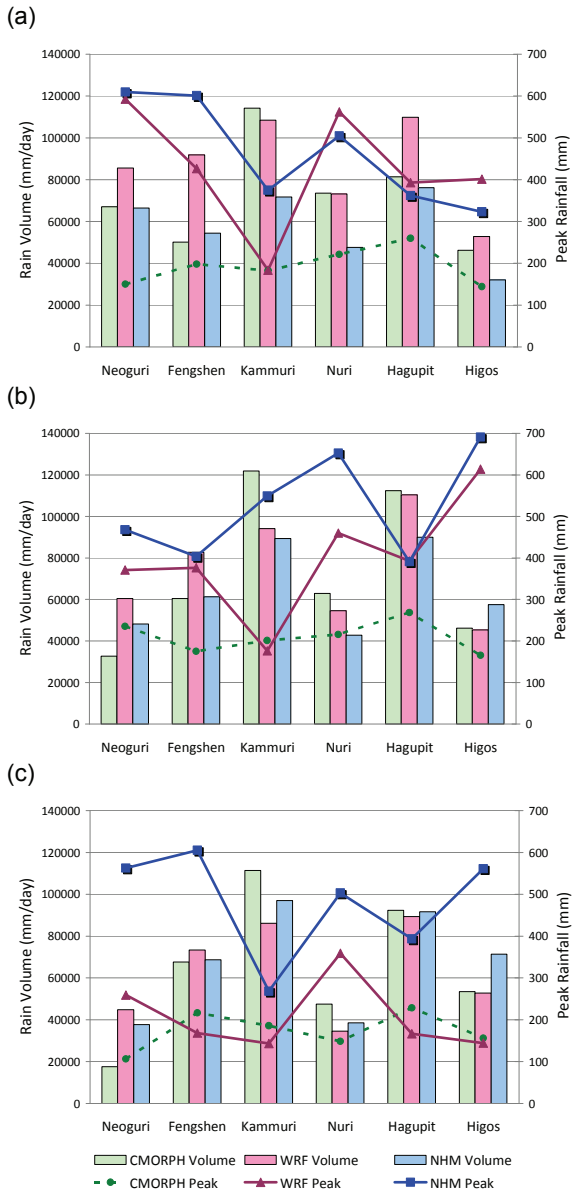

Fig. 9. The rain volume and rain peaks for the six TC cases on (a) day 1, (b) day 2 and (c) day 3.

Some factors may explain the differences in computational efficiency between the models: (a) the compiler is different – the WRF-ARW was compiled using PGI (version 6.2.5) while JMA-NHM was built using Intel compiler (version 10.1.011). The default compilation of WRF-ARW uses single precision floating point, more aggressive compilation options and optimized mathematics library in PGI to speed up the calculation. On the other hand, JMA-NHM was compiled with the standard optimized option in the absence of an on-par mathematics library such as the Intel Math Kernel Library; and (b) the single precision in NHM is not a very exhaustive implementation – double precision is still being used in certain parts of the model physics. It is noted that a smaller difference in the computational efficiency was reported by Hayashi *et al.* (2008) (with WRF-ARW around 40% more efficient than JMA-NHM).

4. SUMMARY AND CONCLUDING REMARKS

An intercomparison of JMA-NHM and WRF-ARW has been conducted to examine their performance in the prediction of TCs over the South China Sea. With a horizontal grid spacing of 10 km and 50 vertical levels using the same domain configurations and very similar physics schemes, 72-hour integrations of the models were run to simulate six tropical cyclones which affected Hong Kong in 2008.

Both models exhibited similar skills in track prediction, with JMA-NHM exhibiting a smaller average track error over the first 36 hours of prediction. On the other hand, WRF-ARW has advantage in intensity forecasts due to a more efficient spin-up of tropical cyclone circulation than JMA-NHM. In the absence of separate TC bogussing, both models still achieved very promising results by displaying appreciable improvement over the predictions from GSM. The deployment of a proper data assimilation system would likely improve the accuracy further especially in the shorter term forecasts.

The quantitative precipitation forecasts from both models have also been verified against the precipitation analysis data of CMORPH. Judging from BS, ETS and RMSE, WRF-ARW generally performed better than JMA-NHM in the quantitative precipitation forecasts of daily rainfall amounts, though the former tended to over-forecast the occurrence of light precipitation at 5 mm/day or below. On the other hand, JMA-NHM suffered from the tendency to over-forecast the occurrence of heavy precipitation. In terms of rain volume, the WRF-ARW generally predicted more rain while JMA-NHM predicted less in the first two days of forecast. The markedly differences in the behaviour of both models in precipitation prediction suggest that certain schemes of the model physics, such as the cloud microphysics and the cumulus parameterization, require tuning for optimal prediction of precipitation associated with TCs.

The computational cost of the WRF-ARW model tested is only half of that of JMA-NHM under the same domain configurations and almost the same physics schemes. One reason is the use of a less capable compiler in building the JMA-NHM used in this study. But if JMA-NHM is to be deployed and run in single precision, a thorough implementation of the program codes to single precision would certainly help boost the execution speed of the model.

References

Hayashi, S., K. Aranami, and K. Saito, 2008: Statistical Verification of Short Term NWP by NHM and WRF-ARW with 20 km Horizontal Resolution around Japan and Southeast Asia. SOLA, 4, 133-136.

JMA, 2007: Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecast System and Numerical Weather Prediction, March 2007, 194 pp.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydromet., 5, 487-503.

Saito, K., J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita and Y. Honda, 2007: Nonhydrostatic atmospheric models and operational development at JMA. J. Meteor. Soc. Japan, 85B, 271-304.

Sapiano, M.R.P. and P.A. Arkin, 2009: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data. J. Hydromet., 10, 149-166.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF Version 3. NCAR Technical Note NCAR/TN-475+ STR.

Wang, W., C. Davis, J. Klemp, G. Holland, and M. DeMaria, 2006: Evaluation of WRF-ARW High-Resolution Tropical Storm Forecasts in 2005 Season.

Wong, W.K., 2010: Development of Operational Rapid Update Non-hydrostatic NWP and Data Assimilation Systems in the Hong Kong Observatory. *3rd International Workshop on Prevention and Mitigation of Meteorological Disasters in Southeast Asia*, 1-4 March 2010, Beppu, Japan.