REVIEW OF RECENT PROGRESS IN TROPICAL CYCLONE TRACK FORECASTING AND EXPRESSION OF UNCERTAINTIES

Julian T. Heming, Fernando Prates, Morris A. Bender, Rebecca Bowyer, John Cangialosi, Phillippe Caroff, Thomas Coleman, James D. Doyle, Anumeha Dube, Ghislain Faure, Jim Fraser, Brian C. Howell, Yohko Igarashi, Ron McTaggart-Cowan, M. Mohapatra, Jonathan R. Moskaitis, Jim Murtha, Rabi Rivett, M. Sharma, Chris J. Short, Amit A. Singh, Vijay Tallapragada, Helen A. Titley, Yi Xiao.

ABSTRACT

The Ninth International Workshop on Tropical Cyclones (IWTC-9) took place in Hawaii, USA in December 2018. This review paper was presented at the Workshop under the Tropical Cyclone Track topic.

The forecasting of tropical cyclone (TC) track has seen significant improvements in recent decades both by numerical weather prediction models and by regional warning centres who issue forecasts having made use of these models and other forecasting techniques. Heming and Goerss (2010) gave an overview of forecasting techniques and models available for TC forecasting, including evidence of the improvement in performance over the years. However, the models and techniques used for TC forecasting have continued to develop in the last decade. This presentation gives an updated overview of many of the numerical weather prediction models and other techniques used for TC track prediction. It includes recent performance statistics both by the models and the regional warning centres.

Keywords: Tropical Cyclone, World Meteorological Organization Workshop, Numerical Models, Track, Forecasting.

1. Introduction

In 2010 the World Scientific book Global Perspectives on Tropical Cyclones: From Science to Mitigation was published (Chan and Kepert, 2010). This included a chapter on track and structure forecasts of TCs (Heming and Goerss, 2010). The aim of this subtopic report is to provide an update on some aspects of this chapter. Many of the Numerical Weather Prediction (NWP) models used for tropical cyclone (TC) track forecasting are described (deterministic and ensemble) together with recent performance statistics. Contributions are also included from Regional Specialized Meteorological Centres (RSMCs) and other operational TC forecasting centres, including usage of NWP models and other forecast aids such as consensus forecasts.

An outline of the structure of the report is as follows:

- Global Models
  - European Centre for Medium-Range Weather Forecasts (ECMWF)
  - Met Office (UK)
  - National Centers for Environmental Prediction (USA)
  - Japan Meteorological Agency
  - USA Navy
  - Canadian Meteorological Centre
  - National Centre for Medium-Range Weather Forecast (NCMRWF, India)
  - WGNE Intercomparison of Global Models

- Regional Models
  - Met Office (UK)
  - National Centers for Environmental Prediction (USA)
  - USA Navy
  - Météo-France
  - Bureau of Meteorology (Australia)

- Ensemble Prediction Models
  - European Centre for Medium-Range Weather Forecasts (ECMWF)
  - Met Office (UK)

Corresponding author: Julian T. Heming, julian.heming@metoffice.gov.uk

DOI: 10.6057/2019TCRR04.01

- National Centers for Environmental Prediction (USA)
- Japan Meteorological Agency
- Canadian Meteorological Centre
- Météo-France

• Operational Forecasting Centres
  - RSMC Miami (National Hurricane Center, USA)
  - Joint Typhoon Warning Center (USA)
  - RSMC Tokyo (Japan Meteorological Agency)
  - Bureau of Meteorology (Australia)
  - RSMC La Réunion (Météo-France)
  - RSMC Nadi (Fiji Meteorological Service)
  - RSMC New Delhi (India Meteorological Department)
  - Canadian Hurricane Centre

2. Global Models

a) *European Centre for Medium-Range Weather Forecasts (ECMWF)*

Model Formulations

Several ECMWF Integrated Forecasting System (IFS) model cycles have been implemented since 2014 with many technical and scientific upgrades, including an increase of horizontal resolution in March 2016 of both the High-RESolution (HRES) and ENSemble (ENS). In the same year the ocean model was upgraded from one degree to a quarter-of-degree resolution, and from 42 vertical levels to 75. With the implementation of IFS cycle 45r1 on the 5 June 2018, HRES became a coupled ocean-atmosphere system, as was then already the case for the ENS for the medium/extended and long-range forecasts. The coupling to the ocean had a significant impact in improving the intensity forecast errors of TCs (Mogensen et al., 2017). Table 2a.1 contains the main model features relevant to forecasting TCs.

Recent Performance

The annual average of HRES TC position forecast errors (all TC basins) over the past decade are shown in Figure 2a.1 (blue lines). The performance of the 10-day TC position forecasts from the ERA5 reanalysis system (run in forecast mode) for 0000 and 1200 UTC has been computed using the same TC tracking software as in operations and the result is included in Figure 2a.1 (red lines). The ERA5 reanalysis is based on IFS cycle 41r2, which was operational between March and November 2016 (see table 2a.1 for details). This dataset provides a useful benchmark against which to compare the HRES operational run since the model configuration used by ERA5 remains consistent throughout the whole period it is compared with the operational run. This helps us to eliminate the year-to-year variability in TC predictability in evaluations of the operational model. Overall there is a good correlation over the year of the annual mean position errors between HRES and ERA5 forecasts, at all lead times. Further details of the production and utility of ERA5 are available in Hersbach et al. (2018).

Mean TC position errors for HRES at 120-hour and 168-hour forecast lead times have decreased by 25% and 40% respectively since 2008. Successive model upgrades have been responsible for the decrease of the mean position errors over recent years. The differences between HRES and ERA5 forecasts became smaller following the model resolution upgrade in 2010, following continual improvements in model physics and following the addition of new observation systems alongside improvement to the assimi-

TABLE 2a.1. Configuration of the ECMWF HRES and ENS systems

	HRES[1]/ENS[2 & 3]	**ERA5**
Spatial resolution	$T_{co}1279$ (0-10 days)[1]/$T_{co}639$ (0-15 days)[2]/$T_{co}319$ (16-46 days)[3] Equivalent to 9 km[1]/18 km[2]/32 km[3]	31 km (TL639)
Vertical resolution	137 levels[1]/91 levels[2] to 1 Pa	137 levels to 1 Pa
Atmospheric Data Assimilation Window for 0000 and 1200 UTC	12-hourly 4D-Var for the HRES, and with the 25 low-resolution members of the Ensemble Data Assimilation (EDA) providing perturbed analyses for the ENS	12-hourly 4D-Var, 10 perturbed ensemble members[4]
Ocean analysis system (OCEAN5)	Provide ocean and sea-ice initial conditions for HRES and ENS (including 5 members with perturbed initial conditions)	
3-D Ocean model (NEMO)	0.25° horizontal resolution with 75 vertical layers. Coupled with the HRES and ENS. Provides SST to HRES and ENS.	
Ocean wave model (ECWAM)	0.125°[1]/0.250°[2]/0.5°[3] resolution. Provides surface stress, Stokes drift and turbulent energy flux at the ocean surface. Coupled with HRES and ENS.	0.36° resolution
Model error (only ENS)	Stochastic Perturbed Parametrization Scheme (SPPT) in the ENS and EDA.	

[3]run Monday & Thursday only.

[4]random perturbations to observations and to model physical tendencies. Current configuration allows an additional 6 hours of observations to be used in the assimilation window than ECWMF operational.

lation that include using the EDA background information in 4DVAR HRES analysis. At shorter forecast lead times the improvement in mean error has been less pronounced in recent years.

Future Developments

The plan for the next model cycle (46r1) is to allow more 'late' observations to enter the system during the assimilation process, without changing the delivery time of the forecasts.

There are also longer term plans to investigate how information from TC near real-time reports (storm position) or the '2-D minimum divergence' method, can help better resolve the wind direction ambiguities, intrinsic to scatterometer wind data, near the centre of TCs. This will reduce errors in the analysed TC position (Figure 2a.1), and, one would expect, in the forecasts also.

Web Links to model information and further verification

Model upgrades in chronological order, with links to extra information:

https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model

The latest confluence page for model cycle 45r1 includes relevant information on the scientific and technical upgrades. It also contains the score card summarizing the positive/negative impacts of that model cycle, including a short summary of the forecast performance for TCs:

https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+cycle+45r1

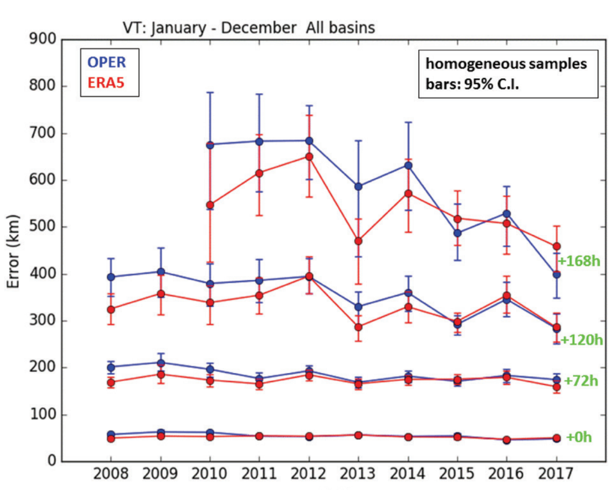


Fig. 2a.1. Annual average of the TC position errors (km) at analysis time, and in three-, five- and seven-day forecasts from the operational HRES (blue), and with 41r2 forecasts initialised with the ERA5 reanalysis (red) shown as a reference. Verification is against the estimates of observed position reported in near real time. All TC basins are included. Bars indicate 95% confidence intervals based on the bootstrap method.

b) Met Office (UK)

Model Configuration

The Met Office Unified Model (MetUM) is a numerical model of the atmosphere used for both weather and climate applications. The MetUM is suitable for NWP, seasonal forecasting and climate modelling with forecast times ranging from a few days to hundreds of years. Furthermore, the MetUM can be used both as a global and a regional model.

The MetUM's dynamical core solves the compressible non-hydrostatic equations of motion with semi-lagrangian advection and semi-implicit time stepping. Sub-grid scale processes such as convection, boundary layer turbulence, radiation, cloud, microphysics and orographic drag are represented by parameterizations. The global NWP configuration has a spherical latitude-longitude grid with spacing 0.140625° x 0.09375° (about 10 km at mid-latitudes) and 70 levels in the vertical. More details of the latest model configuration can be found in Walters et al. (2017). For data assimilation, the model uses a hybrid incremental 4D-Var scheme with 44 short forecasts from the Global Ensemble (Clayton et al., 2013). TCs are initialised by assimilating central pressure estimates from regional TC warning centres issued in real time (Heming, 2016).

The global NWP version of the MetUM is run four times per day out to 168 hours (0000 and 1200 UTC) and 69 hours (0600 and 1800 UTC). TCs are tracked after the 0000 and 1200 UTC runs of the model (Heming, 2017) and forecast guidance messages issued both on the Global Telecommunications System and the Met Office web site for use by TC warning centres.

Forecast Performance

Figure 2b.1 and 2b.2 show the 5-year running mean of TC track forecast errors from the global MetUM for the northern and southern hemispheres. This shows a long term downwards trend in forecast errors which has accelerated since the introduction of a major model change in 2014 which included the introduction of a new dynamical core, a physics upgrade and increase in horizontal resolution (Walters et al., 2017; Heming, 2016). 5-day forecast errors are now lower than 2-day errors were 25 years ago.

Future Plans

The next major change to the global NWP MetUM is coupling to the ocean which is expected to be operational by 2020. This is not expected to improve TC forecast track significantly, but in trials produces much better predictions of TC intensity in cases where ocean feedback is important, such as slow-moving TCs.

In the longer term a completely new modelling framework known as LFRic is being developed to overcome the challenge of weather and climate prediction on the next generation of supercomputers:

https://www.metoffice.gov.uk/research/modelling-systems/lfric

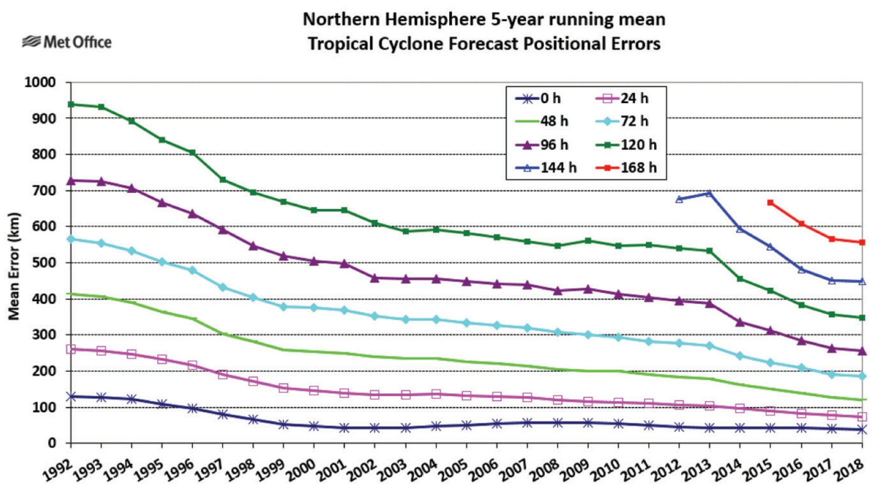


FIG. 2b.1. Met Office global model 5-year running mean TC track forecast error for the northern hemisphere.

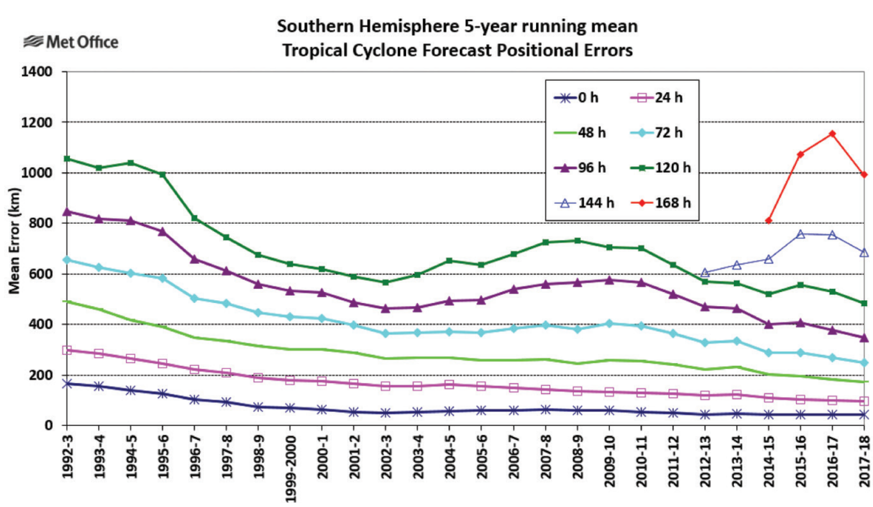


FIG. 2b.2. Met Office global model 5-year running mean TC track forecast error for the southern hemisphere.

c) *National Centers for Environmental Prediction (USA)*

NCEP Operational Models

NOAA's National Centers for Environmental Prediction (NCEP) provides real-time deterministic and ensemble based TC forecast guidance across the globe, primarily to the forecasters at the National Hurricane Center (NHC), the Central Pacific Hurricane Center (CPHC), US Navy Joint Typhoon Warning Center (JTWC) and various public and private forecast agencies across the world. The deterministic models include operational Global Forecast System (GFS) and two exclusive high-resolution TC specific models – the Hurricane Weather Research and Forecast System (HWRF) and the Hurricanes in Multi-scale Ocean-coupled Non-hydrostatic (HMON) models. The ensemble based forecast guidance comes from NCEP Global Ensemble Forecast System (GEFS) and the North American Ensemble Forecast System (NAEFS) that includes multi-model ensembles using Canadian and US Navy ensembles.

Forecast Performance from Operational GFS

NCEP operational GFS is the cornerstone of NCEP Production Suite. GFS is based on Global Spectral Model (GSM) dynamic core and employs a sophisticated 4D Hybrid Ensemble-Variational (4D EnVar) Global Data Assimilation System (GDAS). GFS uses a special technique for relocating TC position in the model background. Apart from providing medium range forecast guidance for global TCs, it also provides initial and boundary conditions for various downstream models including HWRF and HMON.

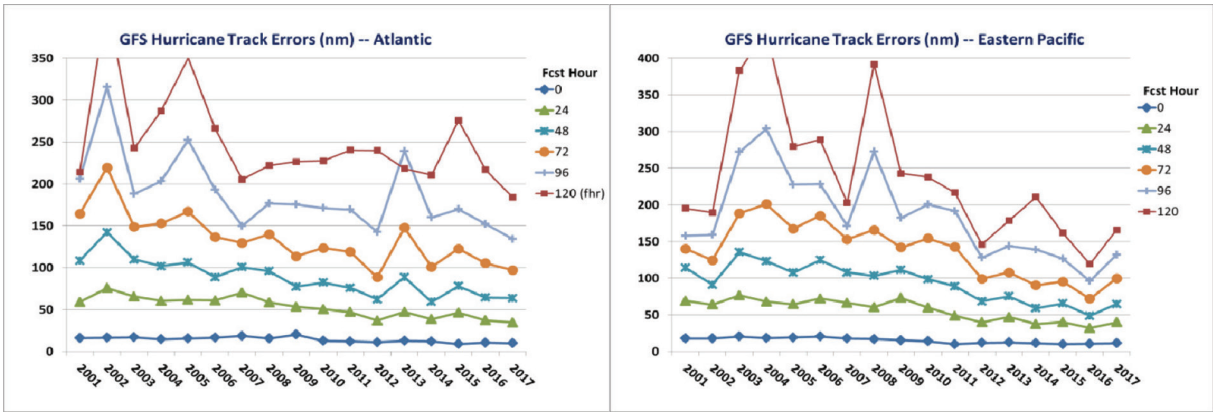


Fig. 2c.1. Annual mean GFS TC track errors for Atlantic and east Pacific for 2001-2017.

GFS is one of the main model used in NHC's track forecast consensus, and has shown significant improvements in the forecast skill over the past several years in the North Atlantic and North Eastern Pacific basins. Figure 2c.1 shows track forecast errors since 2001.

FV3-Based Model Development at GFDL

After serving US National Weather Service for 38 years, the spectral dynamic core based GFS is being replaced with Next Generation Global Prediction System (NGGPS) selected Finite Volume Cubed Sphere (FV3) dynamical core in early 2019.

The GFDL FV3 dynamical core (Putman and Lin, 2007) is the dynamical core (Figure 2c.2) of choice for many global models in the United States, having been implemented in the GFDL suite of global weather and climate models, the NASA GEOS modelling system, and other systems around the world. FV3 was selected as the dynamical core for the US National Weather Service's Next-Generation Global Prediction System (NGGPS), on the basis of its superior accuracy, forecast skill, computational performance, conservation properties, and numerical stability. The FV3 dynamical core has been transferred to the National Weather Service's (NWS) National Center for Environmental Prediction (NCEP) as the replacements for the NWS's Global Forecast System (GFS) and is scheduled to become operational at NCEP in mid-2019.

FV3 uses the forward-in-time scheme and Lagrangian vertical coordinate of Lin (2004), based on the Lagrangian dynamics of Lin and Rood (1997) and the finite-volume pressure gradient force of Lin (1997). FV3 is distinguished from its predecessor, FV (Lin 2004), through the discretization on a quasi-uniform cubed-sphere grid, which avoids the singularity at the poles of the earlier latitude-longitude grid. FV3 has the capability to locally-refine its global cubed-sphere grid by either two-way interactive nesting (Harris and Lin, 2013) or stretching (Harris et al., 2016) using a Schmidt transformation.

The version of the FV3 that is run in near real time by scientists at GFDL is referred to as fvGFS. It uses the operational GFS initial condition (cold start), compared to the parallel version now being run and evaluated extensively at NCEP using a cycled data assimilation that has been developed for FV3-GFS. In order to test the robustness and skill of of their version of fvGFS, scientists at GFDL have performed retrospective forecasts for most of 2015, 2016 and 2017, cold started from the GFS initial condition. The TC track performance were evaluated in all ocean basins (Figure 2c.3) for this three year sample, which provided a robust sample size of slightly over 2000 cases. Overall the track performance was found to be as good as the operational GFS, with some reduction in track error, particularly in the Northwest Pacific where the 48-hour and 72-hour track error was reduced 8%.

An improved version of fvGFS has been developed at GFDL and has been run daily in real time since July 2018. The new 2018 GFDL version of fvGFS introduces the Yonsei University (YSU) planetary boundary layer scheme and a dynamically-active mixed-layer ocean model. The

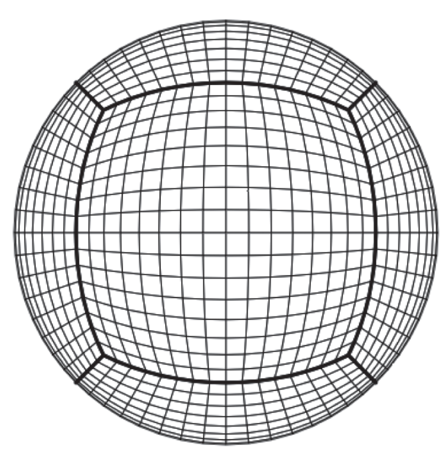


Fig. 2c.2. Schematic picture of the FV3 cubed sphere grid.

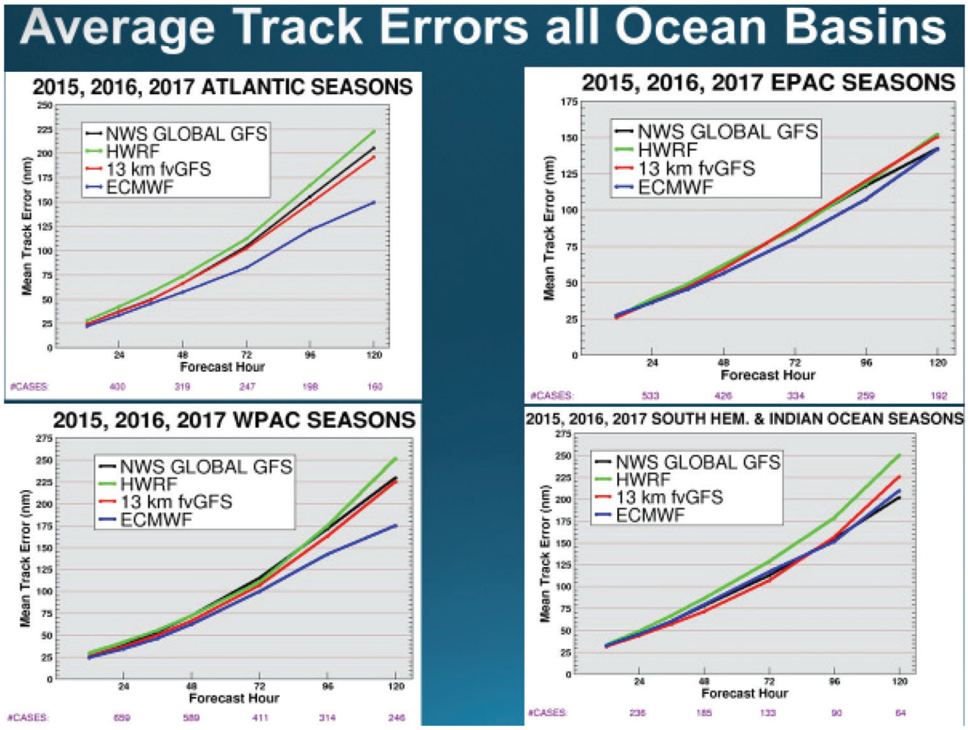


FIG. 2c.3. Average track errors (nm) for the combined 2015, 2016 and 2017 seasons, for the Atlantic, Northeast Pacific, Northwest Pacific and combined Indian Ocean and South Pacific basins, comparing the operational GFS (black), HWRF (green) and ECMWF (blue), with the experimental version of the new 13 km fvGFS model (red), for forecasts starting at 0000 and 1200 UTC.

2018 GFDL fvGFS also includes a revised positive-definite tracer advection scheme and the number of vertical levels is increased from 64 to 91. Also in this new version, the GFDL microphysics has the capability to be called in-line directly from the dynamics, allowing a much faster calling frequency to the microphysics.

So far, the results for the 2018 TC seasons are showing superior track performance with the 2018 GFDL fvGFS compared with both the current operational GFS, the version run in parallel at the Environmental Modelling Center (FV3-GFS) scheduled for operational implementation in mid-2019 and most of the other operational numerical guidance.

Forecast plots from GFDL's fvGFS model are available in real time in the following web site: https://data1.gfdl.noaa.gov/fvGFS/?MODEL=fvGFS

More details on FV3-GFS development and evaluation can be found at these web sites: https://vlab.ncep.noaa.gov/web/fv3gfs http://www.emc.ncep.noaa.gov/users/Alicia.Bentley/fv3gfs/

TABLE 2d.1. Specifications of GSM and GEPS

Model	Resolution (Grid Spacing)	Vertical Levels	Forecast Range (Initial Time)	Initial Condition	Number of Ensemble Members
GSM	TL959 (0.1875°, 20 km)	100 layers (top: 0.01 hPa)	132 hours (0000, 0600, 1800 UTC) 264 hours (1200 UTC)	4D-Var Analysis	-
GEPS	TL479 (0.375°, 40 km)	100 layers (top: 0.01 hPa)	132 hours (0600, 1800 UTC) 264 hours (0000, 1200 UTC)	Global analysis with ensemble perturbations	27

d) Japan Meteorological Agency

Model Formulations

JMA runs several NWP models for various purposes; the Global Spectral Model (GSM), the Meso-Scale Model (MSM), the Local Forecast Model (LFM), the Global Ensemble Prediction System (GEPS) based on a low-resolution version of GSM, an ensemble prediction system based on an atmosphere-ocean coupled model and other NWP models for specific targets such as ocean waves and sea ice extents. For TC information, GSM and GEPS are mainly used. The specifications of the GSM and GEPS for the JMA supercomputer system, which was upgraded in June 2018, are as follows. More detailed information for the models is available at JMA's website:

https://www.jma.go.jp/jma/en/Activities/nwp.html.

Forecast Performance

Figure 2d.1 shows GSM annual mean position errors since 1997. The annual mean errors for 30-, 54- and 78-hour predictions in 2017 were 106, 182 and 300 km, respectively. All were better than those in 2016. 30-, 54- and 78-hour GSM predictions are verified as these are used as primary information by forecasters creating 24-, 48- and 72-hour operational forecasts, respectively.

Future Plans

JMA will increase the horizontal resolution of GSM from the current 20 km to 13 km within the next five years and enhance its physical processes optimized for the increased resolution, aiming at enhancing the representation of smaller scale features of meteorological phenomena including TCs. This will lead to the higher forecast accuracy.

JMA will start all-sky microwave radiance assimilation. In addition, JMA will also improve the assimilation method of high-resolution atmospheric motion vectors derived from Himawari-8. Increase in accuracy of the initial fields of cloud/precipitation areas and initial wind fields is expected with these improvements, which will also result in the higher prediction accuracy of meteorological phenomena including TCs.

e) US Navy

Model Configuration

The Navy Global Environmental Model (NAVGEM) is the U. S. Navy's global weather prediction system, developed by the Naval Research Laboratory (NRL) and run operationally at Fleet Numerical Meteorology and Oceanography Center (FNMOC). NAVGEM deterministic forecasts are made four times per day out to 180 hours and there is also a 20-member NAVGEM ensemble that produces 384-hour forecasts twice per day.

NAVGEM is a spectral model utilizing a Semi-Lagrangian/Semi-Implicit dynamical core. The operational model, as of September 2018, has a horizontal resolution of 31 km (spectral triangular truncation of T425) and 60 vertical levels. The model uses a sea-surface temperature analysis performed by FNMOC, and is held fixed throughout the forecast. Deep convection is parameterized using the Simplified Arakawa-Schubert scheme (Moorthi et al., 2001). Turbulent mixing in the boundary layer and shallow convection are parameterized using an eddy-diffusivity/mass flux approach (Suselj et al., 2014). Further details about the model can be found in Hogan et al. (2014).

The data assimilation system used to define the initial state for NAVGEM is the NRL Atmospheric Variation Data Assimilation System – Accelerated Representer (NAVDAS-AR). NAVDAS-AR uses a hybrid background error covariance matrix combining ensemble-based and static com-

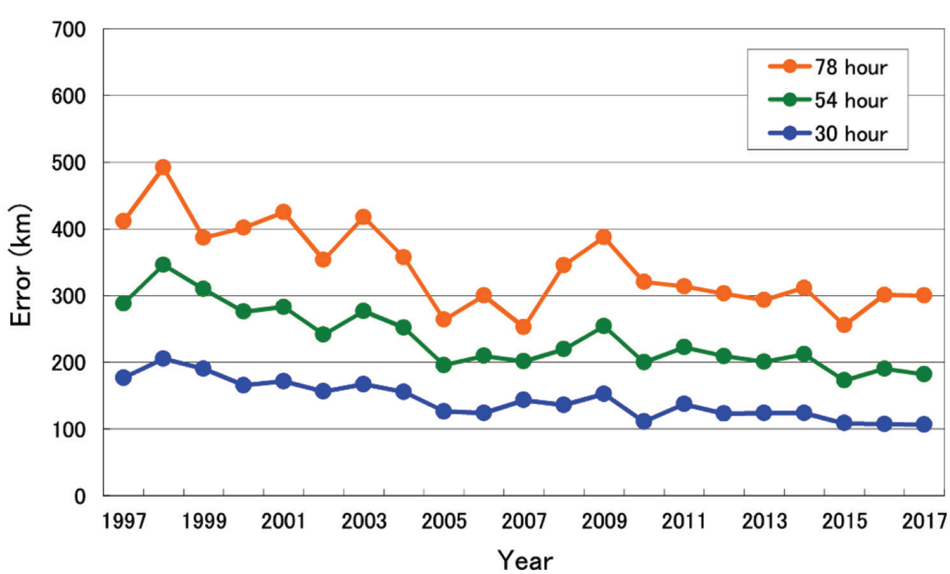


Fig. 2d.1. GSM annual mean position errors in 30-, 54- and 78-hour forecasts in the western North Pacific basin

ponents. For TCs, synthetic wind profiles are assimilated in to represent the TC. The synthetic wind profiles are based on the real-time analysis of position, intensity, and surface wind radii performed by JTWC, NHC or CPHC.

Forecast Performance

Figure 2e.1 shows track mean absolute error for a homogeneous comparison of NAVGEM, GFS, and UKMO global model (UKMET) forecasts of 2017 (left panel) and 2018 (right panel) Atlantic, Northeast Pacific, Central North Pacific, and Northwest Pacific TCs. For both 2017 and 2018, NAVGEM track forecast errors are clearly larger than those of GFS and UKMET.

In recent years, NAVGEM track predictions tend to perform best with respect to other global models in the Northwest Pacific basin (as compared with Atlantic and Eastern North Pacific basins), particularly at later lead times. This can be seen in Fig. 2e.2, which shows track mean absolute error statistics for the 2017 and 2018 Northwest Pacific seasons. The manner in which NAVGEM represents TCs at the initial time and the resolution of the forecast model are best suited for large TCs. The preponderance of large TCs in the Northwest Pacific relative to other Northern Hemisphere TC basin may explain NAVGEM's relatively good performance in that region.

Future Plans

Two major upgrades are planned for NAVGEM to greatly increase the horizontal resolution of the model. The first upgrade, planned for 2019, increases the horizontal resolution to 19 km, implements a two-time level semi-Lagrangian scheme to replace the current 3-time level scheme, and includes refinements to the model physics and how the physics are coupled with the dynamics. A follow-on upgrade, planned for 2020, will feature another increase to horizontal resolution (from 19 to 13 km), an increase in the number of vertical levels from 60 to 100, and additional physics improvements.

Some of deficiencies in NAVGEM track predictions are thought to be related to the TC synthetic wind profile assimilation technique discussed earlier, which was designed for a much lower resolution model and less sophisticated data assimilation system than contemporary NAVGEM and NAVDAS-AR. Work is underway to develop new techniques for TC initialization that depend more heavily on satellite observations and the model background state rather than TC synthetic observations. With improved initialization of TCs alongside the planned increases in NAVGEM resolution, we expect that the model will make substantial progress in TC track prediction in the coming years.

Finally, an initial operational capability at FNMOC for the new Navy Earth System Model (NESM) is planned for 2019. NESM is a coupled atmosphere-ocean-ice-wave model designed for multi-week deterministic prediction and multi-month probabilistic prediction. The deterministic short-term forecast is planned to run once a day out to 16 days lead time, using a 19 km horizontal resolution NAVGEM atmospheric model, the 4.5 km Hybrid Coordinate Ocean Model (HYCOM), the 4.5 km CICE ice model and 14 km WaveWatch III (WW3) wave model. NESM will provide an exciting new capability for air/ocean/wave coupled predictions of TCs.

f) Canadian Meteorological Centre (CMC)

Model Configuration

The CMC's Global Environmental Model (GEM) is a

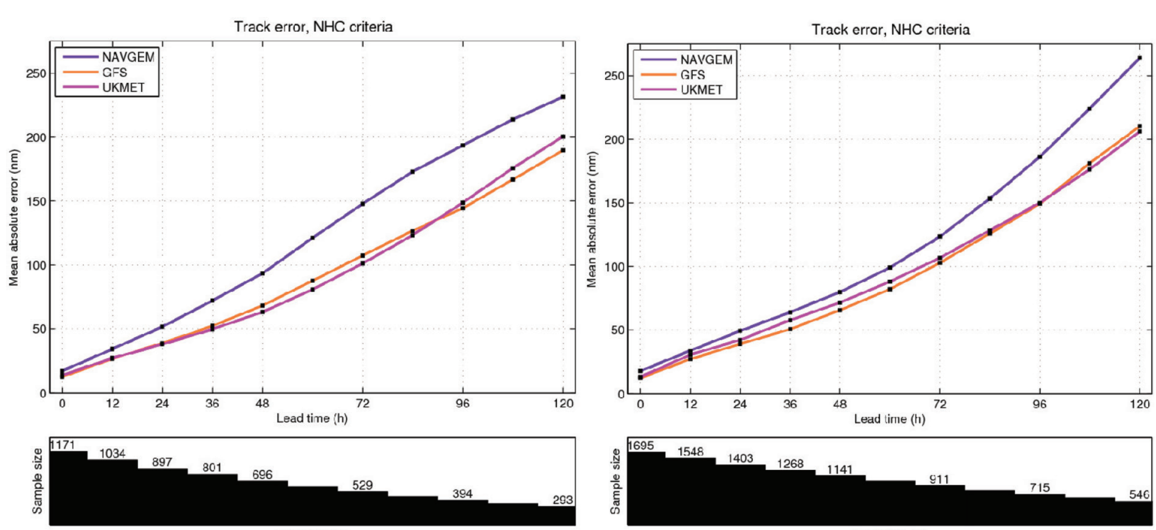


FIG 2e.1. Upper panels show track mean absolute error for NAVGEM, GFS, and UKMET global models. The sample consists of TCs in the Atlantic, Northeast Pacific, and Northwest Pacific basins in 2017 (left) and 2018 (right). Sample size as a function of lead time is shown in the lower panel.

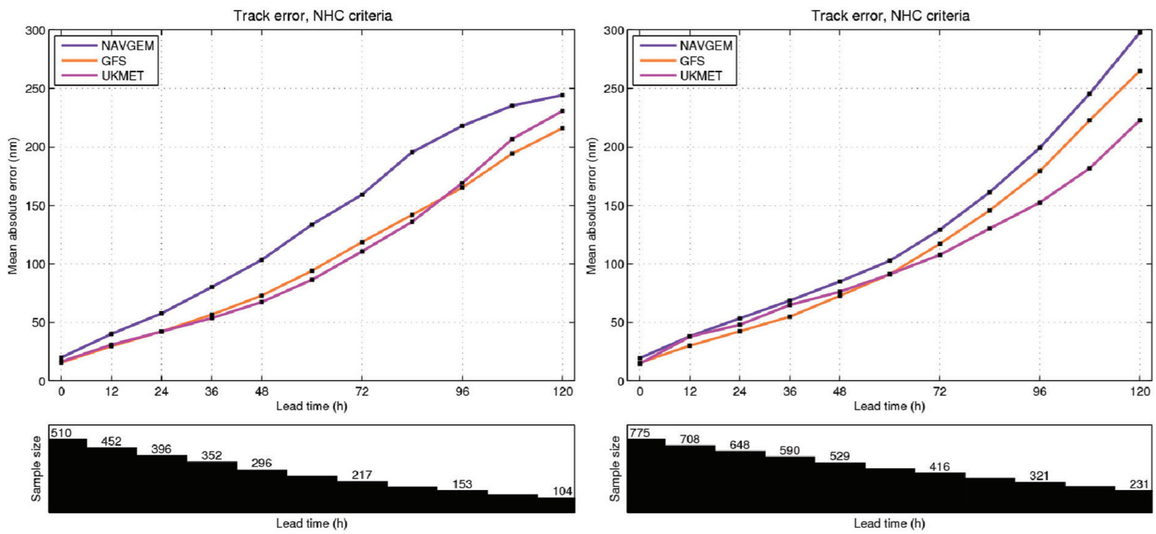


FIG 2e.2. As in Figure 2e.1, but for a sample consisting of TCs in the Northwest Pacific basin in 2017 (left) and 2018 (right).

global grid point model that solves the hydrostatic governing equations on a pair of latitude-longitude Yin-Yang limited area grids (Qaddouri and Lee, 2011). The global deterministic prediction system has a horizontal grid spacing of approximately 25 km with 80 staggered log-hydrostatic-pressure levels in the vertical with a top at 0.1 hPa (Girard et al., 2014). The initial conditions for global forecasts are obtained from an ensemble-variational analysis as described by Buehner et al. (2015). No vortex initialization scheme or synthetic observations are used in the analysis cycle. Details about the model configuration and physical parameterizations used in the global system are documented by Zadra et al. (2014a). Since November 2017 the CMC global deterministic system has been coupled to the NEMO ocean model for the full length of its integration (Smith et al., 2018).

Forecast Performance

The GEM model has a well-known over-prediction bias for TCs, most notable in the model's elevated false alarm rate (Zadra et al. 2014b). A modernization and rebalancing effort for the model physics is currently under way at the CMC, with a focus on improving the representation of the tropical atmosphere. TC tracking results from the new physics package are promising, with a highly significant reduction in the frequency bias, increase in threat score and reduction in track error as shown in Figure 2f.1. These changes are scheduled to be implemented in the global modelling system in mid-2019.

g) National Centre for Medium Range Weather Forecasts (India)

NCMRWF Operational Models

NCMRWF NWP models are based on the Met Office's Unified Model (MetUM) and abbreviated as NCUM.

1. NCUM Global Deterministic Model: In the latest upgrade of NCUM, the horizontal resolution of the model was increased from ~17 km (N768L70) to ~12 km (N1024L70).

2. NCUM Regional Model: The regional configuration of the model has resolution of 4 km and includes explicit convection.

3. NCMRWF Ensemble Prediction System (NEPS): NCMRWF global Ensemble Prediction System (NEPS) was upgraded from ~33 km and 44 members (N400L70) to ~12 km (N1024L70) resolution with 23 members.

TC Tracker Implemented at NCMRWF

The Met Office bi-variate approach to tracking TCs is used in the real-time to track TCs in the North Indian Ocean. This method is in contrast to the earlier NCEP method which used any or all of MSLP, 850 hPa and 700 hPa relative vorticity and geopotential height to track TCs (Marchok, 2002). The bi-variate method identifies TCs by examination of the 850 hPa relative vorticity field but then fixes the TC centre to the nearest local MSLP minimum (Heming, 2017). The key advantage of the method is that it gives a strong signal of the approximate centre of the TC even for weak systems and does not depend on the current position information for tracking.

Forecast Products and Performance for Recent Cyclonic Storm Daye

Cyclonic Storm Daye evolved from a depression which developed over the Bay of Bengal in September 2018. Daye made landfall near southern Odisha, also impacting the adjoining north Andhra Pradesh coast, resulting in

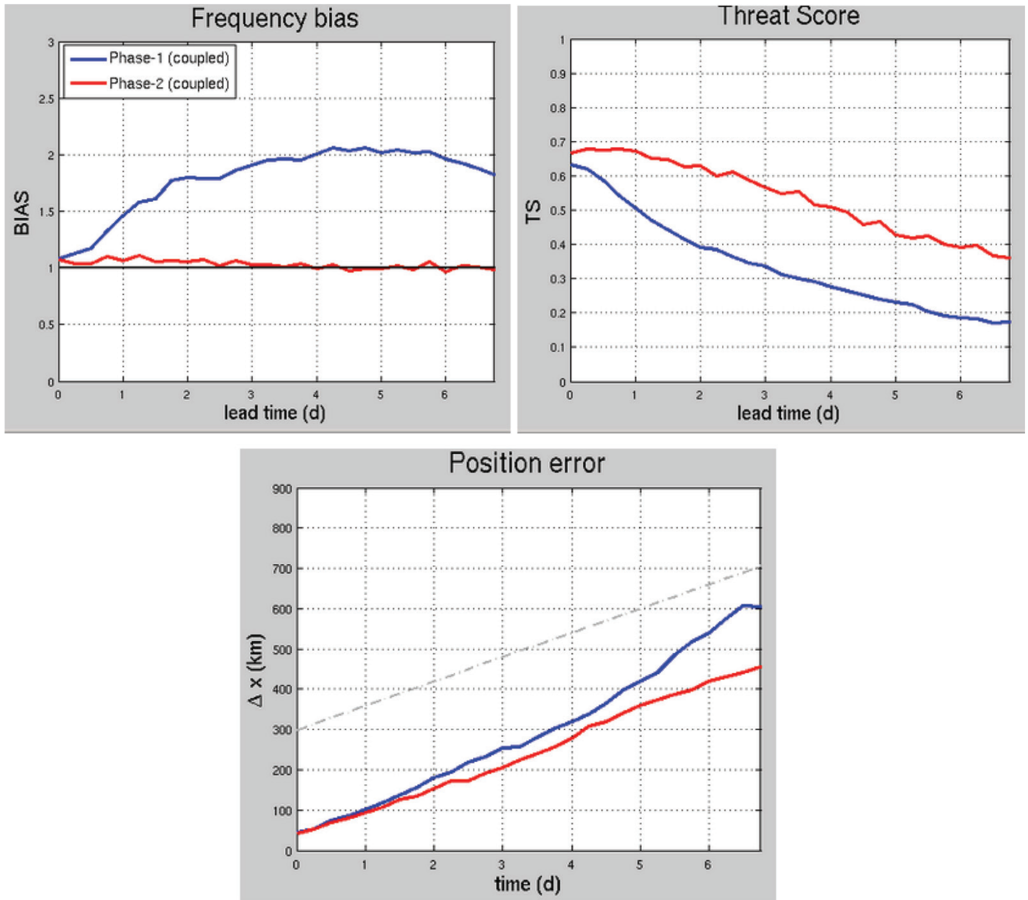


FIG. 2f.1. The frequency bias (upper left), threat score (upper right) and mean track error (bottom) for TCs in the current operational global model at CMC (blue) and a version of the system with updated atmospheric physics scheduled for implementation in mid-2019 (red) is shown for a two-month period in summer 2016. In all panels, results were obtained using a TC tracking algorithm and comparison with the best track (IBTrACS) dataset for systems of tropical depression strength and greater.

heavy rains and strong wind. Figure 2g.1 shows example track forecast products for the NCUM global and NEPS models.

Figure 2g.2 shows mean track forecast errors for the NCUM global and regional models and NEPS ensemble mean for both 0000 UTC and 1200 UTC tracks. All verification for the track forecasts uses the IMD track data. This shows that particularly at longer lead times the NEPS ensemble mean shows the lowest track forecast errors.

h) WGNE Intercomparison of Global Models

History

The Working Group on Numerical Experimentation (WGNE) was established by the World Climate Research Programme Joint Scientific Committee and the World Meteorological Organization Commission for Atmospheric Sciences. The group works to foster the advancement of NWP models, with a membership consisting of representatives from operational NWP centres and research organizations. One of WGNE's many activities is its intercomparison of TC track forecasts, which it has conducted using operational global NWP models since 1991 (Tsuyuki et al., 2002; Yamaguchi et al., 2017). This work is undertaken annually by JMA and includes a large number of both global and regional models used for TC prediction.

Forecast Performance

Due to the large amount of work involved to collect NWP model data from agencies around the world and calculate and publish statistics, results for each calendar year are usually published about 22 months after the end of the year. Thus currently results up to 2016 are the latest available. Large amounts of verification data are available from the WGNE TC Intercomparison web site: http://nwp-verif.kishou.go.jp/wgne_tc. A few examples have been selected for this report.

Figure 2h.1 shows track forecast errors for various NWP models in 2016 for three regions: the Northwest Pacific,

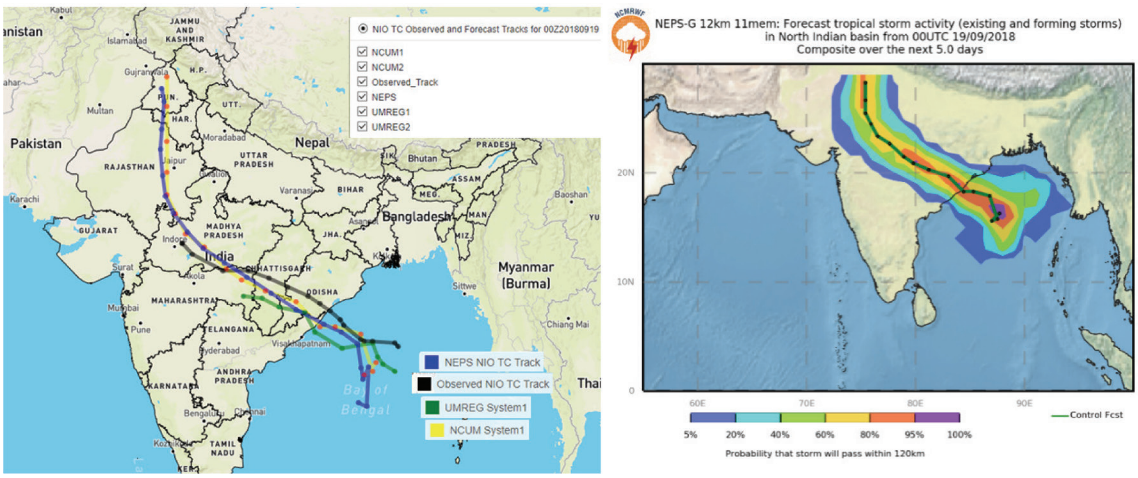


Fig. 2g.1. Forecast tracks for Cyclonic Storm Daye. Right panel also shows ensemble forecast (NEPS) strike probabilities.

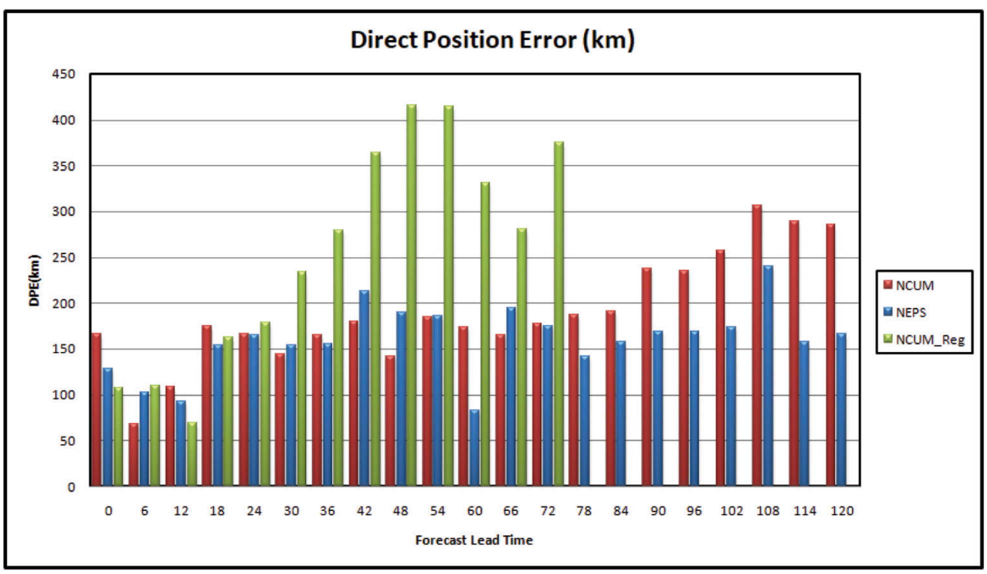


Fig. 2g.2. Direct positional errors in forecasts for Cyclonic Storm Daye (September 2018).

North Atlantic and Australian region. These show that for the Northwest Pacific the ECMWF model was the best performer followed by the KMA and UKMO models. In the North Atlantic the UKMO model was mostly the best performer with KMA and ECMWF also performing well. In the Australian region the NCEP (GFS) model was the best performer with ECMWF being the second best.

Figure 2h.2 shows a long time series of 72-hour forecasts for the same three regions. In the Northwest Pacific the ECMWF model has mostly been the best performer since the start of the intercomparison. There is a clear downwards trend in track forecast errors for many of the models. In the North Atlantic the ECMWF and NCEP (GFS) models have shown to be good performers, particularly in the last decade. The Australian region has only been part of the intercomparison for the last decade and in this region the ECMWF and NCEP (GFS) models perform well.

3. Regional Models

a) Met Office (UK)

Model Configuration

The Unified Model (MetUM) is the Met Office's weather and climate prediction model. It solves the full, deep-atmosphere, non-hydrostatic, Navier-Stokes equations using a semi-implicit, semi-Lagrangian numerical scheme (see Wood et al., 2014 for details). Model prognostic fields are

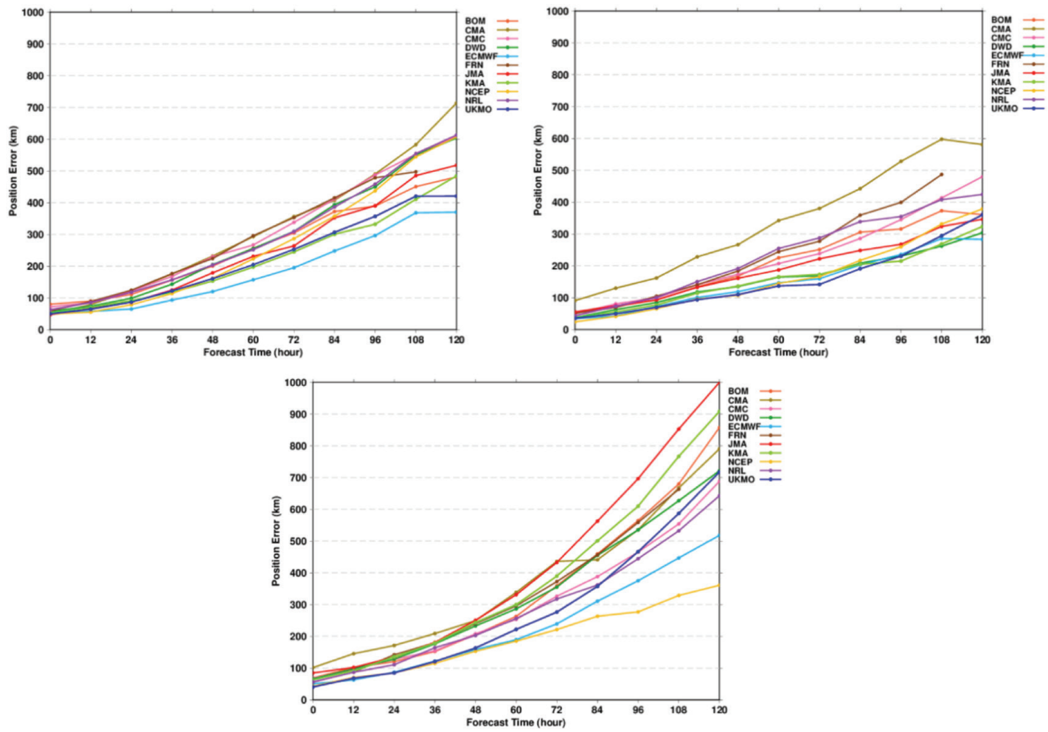


FIG. 2h.1. Track forecast errors from NWP models in 2016. Top left: Northwest Pacific. Top right: North Atlantic. Bottom: Australian region.

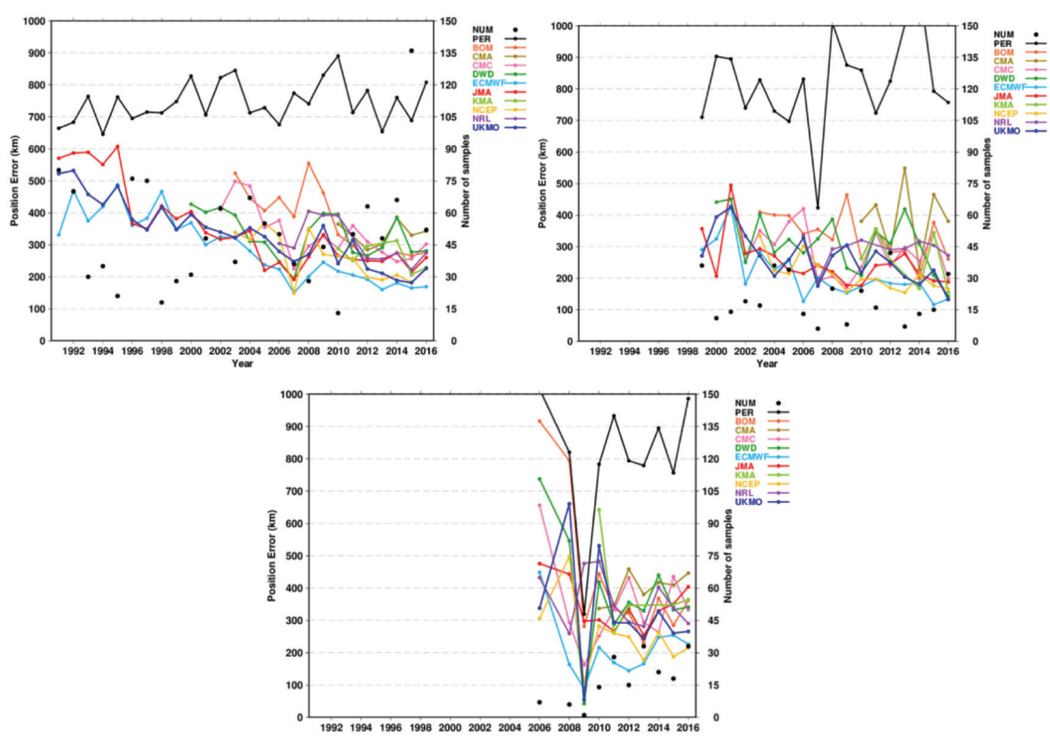


FIG. 2h.2. Time series of 72-hour track forecast errors from NWP models (coloured lines) and persistence (black line). Top left: Northwest Pacific. Top right: North Atlantic. Bottom: Australian region.

discretized on to a regular latitude/longitude grid with Arakawa C-grid staggering (Arakawa and Lamb, 1977), whilst the vertical discretization utilizes a Charney-Phillips staggering (Charney and Phillips, 1953) and a terrain-following hybrid-height vertical coordinate.

The MetUM is used at the Met Office to produce global and regional, deterministic and ensemble forecasts for TCs. Convection-permitting tropical regional models currently being run at the Met Office include:

1. A deterministic 4.4 km grid length model spanning Southeast Asia, with a 1.5 km model for the Philippines nested inside this. This system is run twice a day (0000/1200 UTC) out to 120 hours.

2. An 18-member, 4.5 km ensemble system for a domain covering the Philippines, also run out to 120 hours twice a day.

3. A re-locatable 18-member, 4.4 km ensemble system used to produce on-demand forecasts for major Atlantic TCs.

All of these models have 80 vertical levels, the spacing of which increases quadratically with height up to a fixed lid 38.5 km above sea level. Initial and boundary conditions are supplied by the Met Office operational global model (either deterministic or ensemble system, as appropriate), which uses the Global Atmosphere (GA) 6.1 science configuration (Walters et al., 2017). There is no data assimilation or vortex specification in the regional models; initial conditions are derived by simple interpolation of global model fields. Regional models are one-way nested inside the driving global model and there is no atmosphere-ocean coupling (the sea-surface temperature is held fixed throughout a forecast).

The science configuration of the MetUM used in tropical regional models is Regional Atmosphere and Land – Version 1 (RAL1-T, where the T denotes the tropical version of the configuration) See Bush et al. (2019) for details. The most important difference between RAL1-T and GA6.1 is that the convection parametrization is switched off in the former. From a TC modelling perspective, another key difference is that frictional heating due to the dissipation of turbulence in the surface layer is not included in RAL1-T.

Forecast Performance

As part of RAL1-T testing, a 4.4 km regional model of the Philippines (domain shown in Figure 3a.1) was used to re-run all 2015 TC cases (a particularly active El Niño year). Figure 3a.2 shows the mean error in storm position relative to observations (as measured by the direct positional error, DPE) as a function of lead time. Corresponding mean track errors derived from Met Office global model (GA6.1) and HWRF operational forecasts are also displayed. Note that the storm sample has been homogenized across the models.

Overall, the RAL1-T and GA6.1 models give similar mean track errors (Figure 3a.2a). The DPE increases by

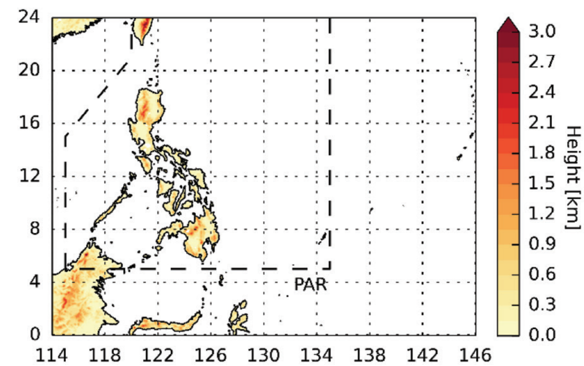


FIG. 3a.1. Philippines regional model domain and orography. The dashed black line shows the portion of the Philippines Area of Responsibility (PAR) inside the domain.

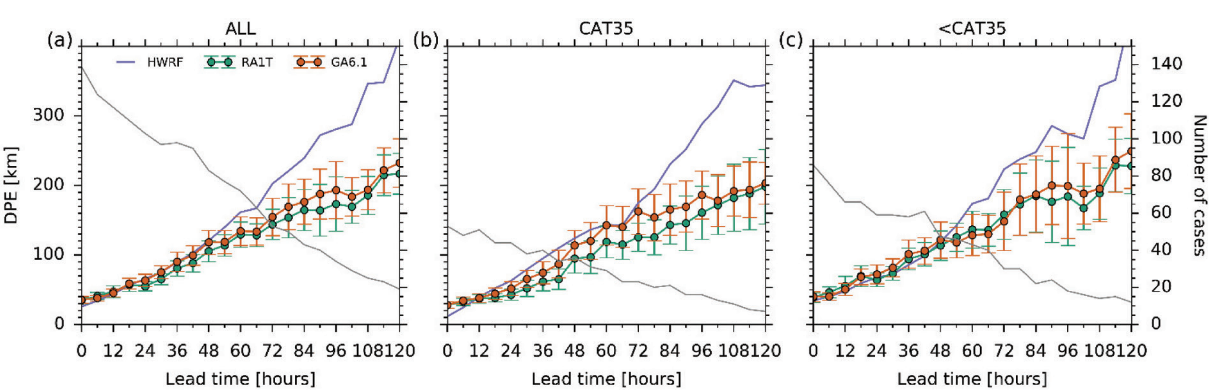


FIG. 3a.2. Error in forecast storm position relative to observations (direct positional error, DPE) as a function of lead time for the RAL1-T regional and GA6.1 global models, for (a) all storms in the sample, (b) storms of category 3 and above, and (c) storms below category 3. The solid lines with error bars represent the mean error and 95% confidence intervals on the mean, respectively. The mean track error from the HWRF model is also displayed. The solid grey lines indicate the number of storm cases (see the right-hand axis of each plot).

approximately 36 km per day of forecast, reaching a maximum of around 200 km at 120 hours. Although track errors relative to observations are comparable, storm positions are typically different in the two models (not shown, but see Short and Petch, 2018), implying that the large-scale steering flow inherited from the driving global model is modified by the regional model. This is likely due to differences in the model physics, in particular the treatment of convection (not parametrized in the regional model), boundary layer and microphysics. All of these are known to affect TC tracks in numerical models (e.g. Li and Pu, 2009; Fovell et al., 2009; Nasrollahi et al., 2012; Biswas et al., 2014; Shepherd and Walsh, 2017).

When the storm sample is stratified by intensity (Figures 3a.2b and 3a.2c), there is an indication that RAL1-T may give improved track forecasts for the most intense storms (category 3 and above, CAT35) at longer lead times, possibly because of the different representation of key physical processes discussed above. However, the limited sample size means this is not a statistically significant result. For weaker storms (below category 3, <CAT35) there is little difference between model track predictions. On average, both models are able to forecast the position of strong storms better than weaker storms.

HWRF gives similar mean track errors to the two MetUM configurations out to 48 hours or so, but the rate of error growth increases beyond this, leading to a larger mean track error in the latter stages of the forecast.

Future Plans

The future development of regional configurations of the MetUM for TC forecasting will be targeted at improving intensity, rather than track predictions. A change that has recently gone into the models described above is a cap on the air-sea drag coefficient at high wind speeds, as motivated by observational data. Trials have demonstrated this considerably reduces the weak bias in TC surface winds seen in regional and (to a lesser extent) global configurations of the MetUM. Looking further ahead, there are plans to improve the initialisation of TC forecasts (at present, the regional model is somewhat handicapped by having to spin-up strong storms from relatively weak global analyses) and to implement some form of air-sea coupling in the model (along with frictional heating in the surface layer).

b) *National Centers for Environmental Prediction (USA)*

Forecast Performance from High-Resolution HWRF

Lately, NCEP's HWRF model has been one of the top performing operational track prediction models. Improvements to model resolution (3 km in 2012, 2 km in 2015 and 1.5 km implemented in 2018), physics and initial conditions enhanced with aircraft observations, have led to steep-step progress in improved numerical guidance. Figure 3b.1

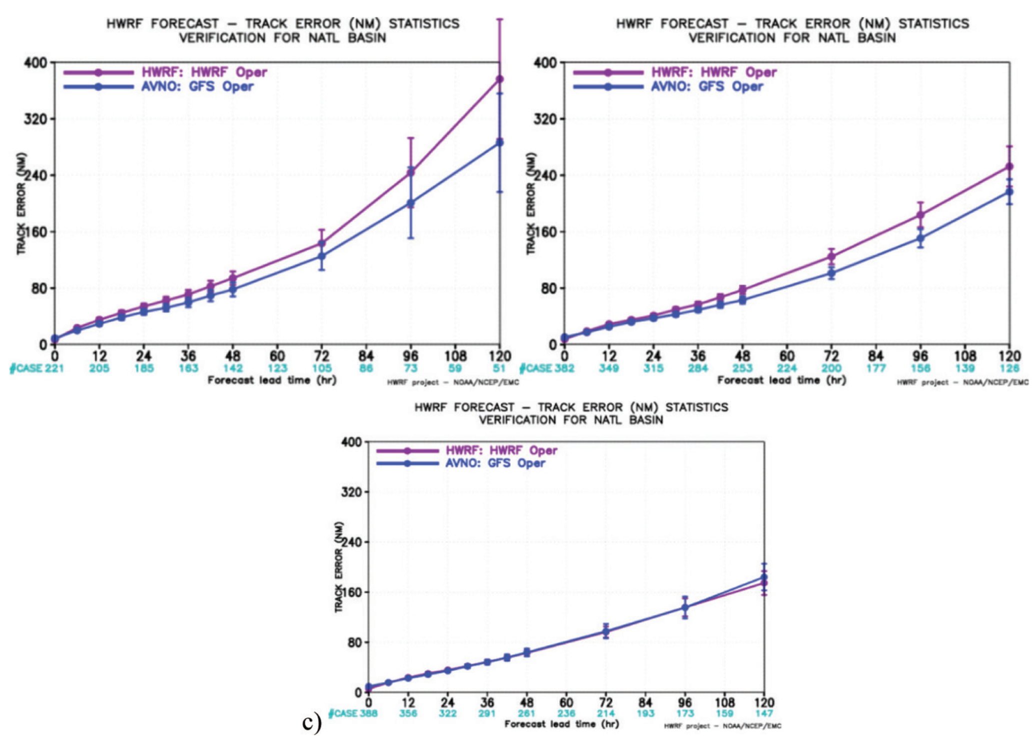


FIG. 3b.1. Operational HWRF track errors for the Atlantic Basin from 2015 (top left), 2016 (top right) and 2017 (bottom).

below illustrates the progress of operational HWRF in forecasting track with significantly reduced errors for the year 2017 as compared with 2015. These improvements can be attributed to increased vertical resolution (from 61 to 75), advanced data assimilation including use of high density aircraft based observations, and implementation of advanced scale-aware convective parameterization scheme. More details on these upgrades are documented in Gopalakrishnan et al., 2018. During the 2017 season, track skills of HWRF were comparable to operational GFS for most lead times.

c) US Navy

Model Configuration

The COAMPS-TC system (Doyle et al., 2014 and 2012) is a high-resolution regional dynamical model designed for prediction of TC track, intensity, and structure and run by Fleet Numerical Meteorological and Oceanography Center (FNMOC). The COAMPS-TC atmospheric model features a non-hydrostatic dynamical core and physical parameterizations for cloud microphysics, boundary layer and free-atmospheric turbulent mixing, surface fluxes, radiation, and deep and shallow convection. The atmospheric model is fully coupled to the Navy Coastal Ocean Model (NCOM; Martin, 2000; Martin et al., 2006) in order to represent the interaction of a TC with the underlying ocean (Chen et al., 2010).

For the 2018 operational version of COAMPS-TC, the atmospheric model consists of a fixed outer grid mesh at 36 km resolution and two storm-following inner grid meshes at 12 km and 4 km resolution. The atmospheric model uses 40 vertical levels, with a top at 10 hPa. The NCOM ocean model is run on a single 7.5 km fixed mesh with 40 levels in the vertical.

The COAMPS-TC atmospheric model is cold started from a global model analysis. The version of COAMPS-TC (run in real-time by NRL) with the initialization from the NOAA GFS system is known as CTCX whilst the version of COAMPS-TC initialized from the Navy NAVGEM system is known as COTC.

A balanced synthetic vortex is inserted in the storm-following 12 km and 4 km grid meshes and replaces the global model analysis in these regions.

Forecast Performance

Figure 3c.1 shows a homogeneous comparison of COTC and CTCX configurations. The retrospective forecast sample consists of 381 cases from TCs that occurred in 2015, 2016 and 2017 mostly in northern hemisphere basins. The track mean absolute error (MAE) for CTCX is far lower than that of the COTC tracks, especially between 24 and 96 hours, when the MAE improvement is 20-30%. This illustrates the strong dependence of regional model TC track forecast performance on the parent global model.

2017 real-time track MAE statistics for the western Atlantic, eastern Pacific, and western Pacific are shown in Figure 3c.2 for a number of regional and global models including COTC and CTCX. The close relationship between the global model track forecast performance and the track performance for COAMPS-TC can also be seen here, as the COTC track MAE closely follows that of NAVGEM and likewise, CTCX closely follows GFS.

Future Plans

Future upgrades are being planned for the COAMPS-TC modeling system. A cycling 4D-Var or ensemble-based

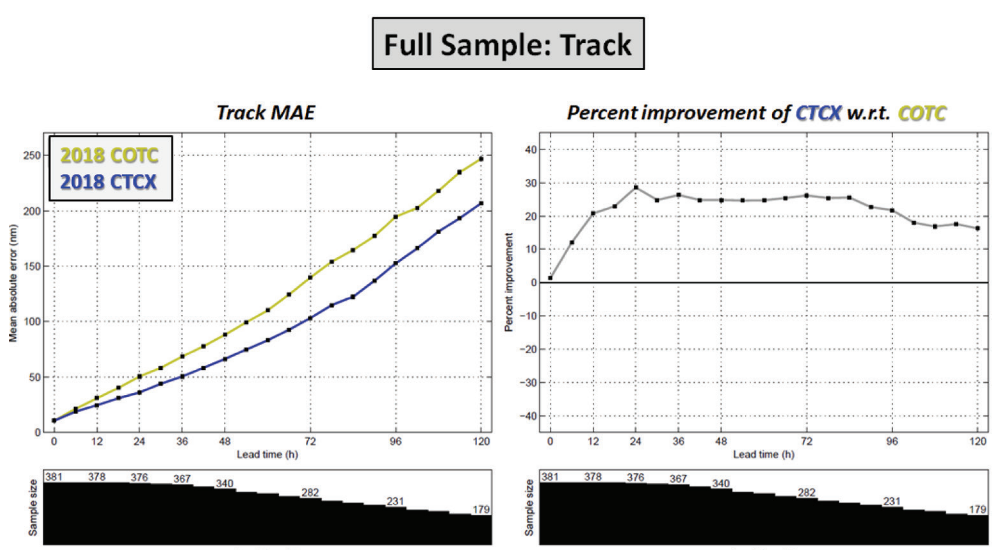
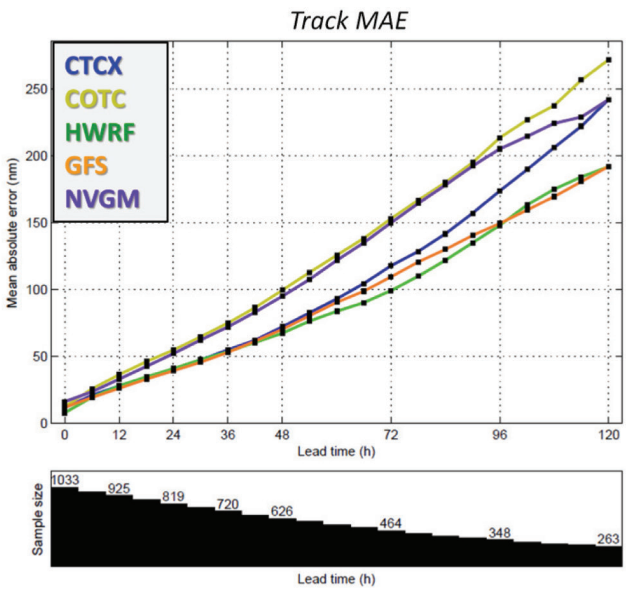


Fig. 3c.1. Full retrospective sample track MAE for COTC and CTCX (left panel) and track MAE percent improvement for CTCX w.r.t. COTC (right panel).



CTCX sample

	0 h	24 h	48 h	72 h	96 h	120 h
Atlantic	403	325	261	212	168	147
EastPac	342	258	195	155	121	81
WestPac	537	434	317	192	136	101
	1282	1017	773	559	425	329

FIG. 3c.2. Track MAE for 2017 operational models including the CTCX (GFS-based) and COTC (NAVGEM-based) COAMPS-TC versions. Other models shown include HWRF, GFS, and NAVGEM. The CTCX sample size is shown in the right panel.

data assimilation system for model initialization will be introduced in the next several years. Upgrades to the physical parameterizations (especially boundary layer and microphysics) and an increase in the vertical resolution from 40 to 60 levels are anticipated within 3 years. An increase in inner-most nest horizontal resolution from 4 km to 2 km is planned for 2021 followed by a further increase to 1 km as operational computing resources permit. The air-ocean coupled model will utilize a coupled data assimilation system, the Navy Coupled Ocean Data Assimilation (NCODA; Cummings, 2005) and the WaveWatch model (WW3) will be added to the coupled system to better represent the air-sea interface. A high-resolution COAMPS-TC ensemble system is being developed to characterize state-dependent track and intensity forecast uncertainty. An 11 member ensemble with 4 km horizontal resolution (and not coupled to an ocean model) is being transitioned to operations at FNMOC later this year.

d) Météo-France

Model Configuration

Météo-France operates five convection permitting models, centred on main French overseas Territories, which have been in operations since February 2016. These versatile systems, used both for daily weather and TC forecast, have static domains (Figure 3d.1).

• Horizontal resolution of 2.5 km, with 90 vertical levels (from 5m); 60s time step; explicit deep convection.

• No data assimilation scheme; their initial and lateral conditions are derived from ECMWF HRES model. For the surface, use of data from Arpège (for continents) and Mercator-Ocean global model PSY4 (1/12°) for the ocean.

• Coupled with the surface model SURFEX, along with a 1D ocean model.

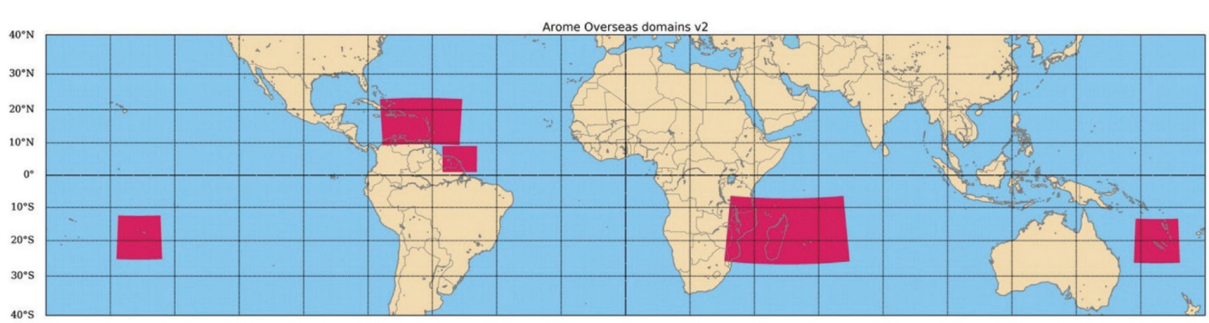


FIG. 3d.1. Arome Overseas model domains

• Four runs per day, up to 42 hours (78 hours on demand, when there is a TC threat for instance).

Further details of the Arome model configuration can be found in Seity et al. (2011), Brousseau et al. (2016) and Termonia et al. (2018).

Forecast Performance

At Météo-France, the main verification of TC track predictions is focused on Arome Overseas models. Since they don't have movable domains, the verification sample is small, even when gathering all domains. Nevertheless, these models have shown good skill in TC track forecasts, at least on par with their driving model (ECMWF HRES model), as shown on Figure 3d.2.

Occasionally Arome Overseas models perform much better that their driving model for TC track prediction. This ability appears as a positive feedback of better intensity on track. The better representation of the TC intensity and structure during the forecast leads to a more accurate steering flow and then a better track prediction.

e) Bureau of Meteorology (Australia)

Model Configuration

The Australian Bureau of Meteorology's ACCESS models are based upon the Met Office Unified Model (MetUM) system. ACCESS-TC2 is the current tropical prediction model configured at a resolution of 0.11°x0.11° in the horizontal and 70 levels in the vertical. Forecasts out to 72 hours are produced from 0000 UTC and 1200 UTC base times and are triggered by the existence of one or more TC within the Asian Tropical domain (covering the South Pacific, East Indian and Northwest Pacific ocean basins), with up to three domains available (for three concurrent TCs). For each domain, five high-resolution analyses are performed every six hours from T-24 (cycle-1) through to T-0 (cycle-5), with a final 72-hour forecast run in cycle-5. First guess fields for cycle-1 are derived from the global ACCESS-G2's initial conditions, and reconfigured to the TC domain; the boundary conditions are obtained from ACCESS-G2 forecasts. For cycles 2 to 5, the system is warm run, i.e., the first-guess input files for these cycles are obtained from forecast data from the previous cycle of ACCESS-TC2.

Unlike the other ACCESS NWP systems, ACCESS-TC2 uses a synthetic observation scheme (Davidson et al., 2014) in addition to the normal in-situ and satellite observations to help define the TC vortex. Based on estimates of present and past locations, central pressure and storm size, vortex specification is used to filter the analysed circulation from the original analysis, construct the inner-core of the storm, impose motion asymmetries consistent with the past motion of the storm, merge the synthetic vortex with the large-scale analysis at outer radii, and relocate the vortex to its observed position. Synthetic observations are extracted from the idealized vortex at a resolution sufficient to resolve the maximum wind at the radius of maximum wind. In ACCESS-TC2, synthetic observations for surface pressure only are then merged with the conventional observations in the ACCESS Observation Processing System (OPS) module.

A comparison of the system specifications of current and upcoming versions of ACCESS-TC2 are listed in Table 3e.1 below.

Forecast Performance

Long term time series plots showing the annual mean track and central pressure errors (verified against JTWC best track data) for the Bureau's dedicated TC NWP sys-

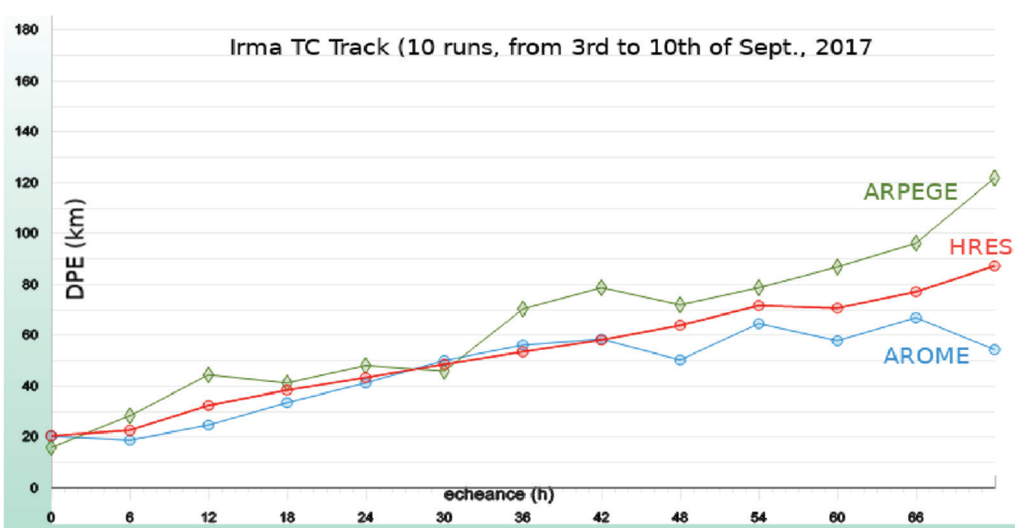


Fig. 3d.2. Track forecast errors for Arome, compared to ECMWF HRES and Arpège for Hurricane Irma (10 model runs). Credit : LACy

TABLE 3e.1. ACCESS-TC System Specifications

	ACCESS-TC2	ACCESS-TC3
Domain	33.0°x33.0°, relocatable anywhere within 3° of the ACCESS Tropical domain, i.e. boundary extremes within 42.0°S to 52.875°N, 63.0°E to 145.875°W.	Same
Geographical limits for initial vortex location	Minimum 3° from ACCESS-TC boundary, i.e. 39.0°S to 49.875°N, 66.0°E to 148.875°W	Same
Geographical limits for vortex tracking program	Minimum 2.25° from ACCESS-TC boundary, i.e. 39.75°S to 50.625°N, 65.25°E to 148.875°W	Same
UM horizontal resolution (lat x lon)	300x300 (0.11°x0.11°)	920x920 (0.036°x0.036°)
Analysis horizontal resolution (lat x lon)	100x100 (0.331°x0.331°)	320x320 (0.102°x0.102°)
Vertical resolution	L70, top level at 80km	L80, top level at 38.5km
Observational data used (6hr window)	AIRS, ATOVS, CrIS, ATMS, IASI, ASCAT, AMV, SYNOP, SHIP, WINDPROFIL, BUOY, AMDARS, AIREPS, TEMP, PILOT, GPSRO,	
Synthetic Surface Pressure Observations	As before plus also SSMIS, ScatSat, increased AMV (possibly 10 minute)	
Sea surface temperature analysis	Daily global 0.25° SST analysis	Same
Soil moisture analysis	N512 soil moisture field. SURF interpolates to the targeted resolution.	N1024 soil moisture field. SURF interpolates to the targeted resolution.
Internal model time step	5 minutes (288 time steps per day)	2 minutes (720 time steps per day)
Analysis time step	15 minutes	6 minutes (240 time steps per day)
Nesting	ACCESS-G2 (N512L70, 25 km)	ACCESS-G3 (N1024L70, ~12 km)

tems since 2005 are shown in Figure 3e.1 below. Between July 2001 and August 2010, the model was the locally-developed TCLAPS (Davidson and Weber, 2000), initially with a 15 km model horizontal resolution but increased to 10 km in June 2008. The MetUM-based ACCESS-TC0 was introduced in Nov 2011, with subsequent minor upgrades in December 2013 (ACCESS-TC1) and December 2016 (ACCESS-TC2). All versions of ACCESS-TC to present have been configured with a model horizontal resolution of 12 km.

Track errors showed a dramatic improvement following the introduction of ACCESS-TC in 2011. Since then, performance has been fairly static, although the 2017 ACCESS-TC2 results were the best yet at almost all forecast ranges.

ACCESS-TC has shown a consistent positive bias in the central pressures, i.e. the modelled pressure is not as deep as the best track estimate. This is expected to improve significantly when we move to a higher horizontal resolution in the next model upgrade.

Future Developments: ACCESS-TC3

An upgraded, higher resolution version of ACCESS-TC is currently undergoing development trials, with a target implementation date of mid-2019. The most significant changes involve an increase in resolution to 0.036°x0.036° (approximately 4 km) in the horizontal and 80 levels in the vertical and use of explicit convection (i.e. not param-

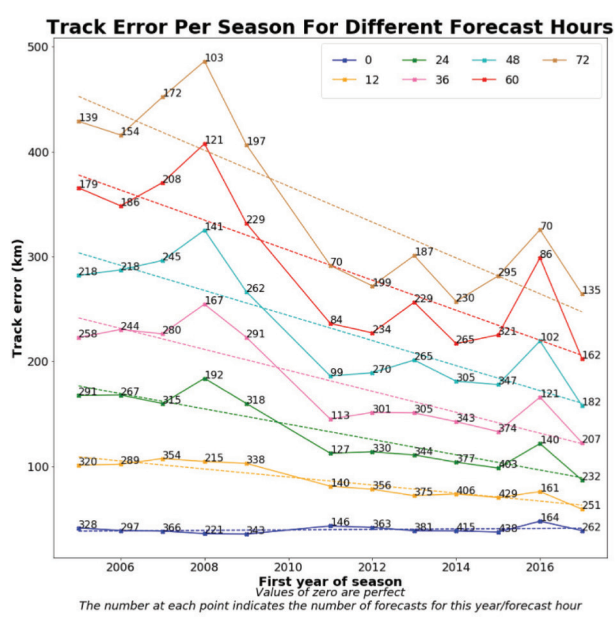


FIG. 3e.1. Long term annual mean track errors (km) for Bureau of Meteorology operational TC NWP models.

eterised convection). It uses TC specific background error covariances derived from a training set of paired forecasts for TC instances between July 2015 and February 2016, generated by the control variable transforms (CVT) method (Bannister, 2008). Software upgrades include the use of more recent versions of the UKMO software components and the UKMO 'RA1T' physics configuration.

ACCESS-TC3 will nest within initial and lateral boundary conditions from the upcoming 12 km ACCESS-G3 system. ACCESS-G3 will also commence assimilating hourly-interpolated TC central pressures derived from international TC advisory bulletins, as is done in the UKMO global model (Heming, 2016).

4. Ensemble Prediction Models

a) European Centre for Medium-Range Weather Forecasts (ECMWF)

Details of the Formulation of the ECMWF ensemble system (ENS) are given under Section 2 (Global Models) of this report.

To assess one aspect of the performance of the ENS TC position forecasts, for the medium range, the annual average of the forecast error of the ENS mean position is compared with the corresponding ensemble spread. If the sample size is large, both measures should match in a well-tuned ensemble forecast system. Figure 4a.1 shows the evolution of the ensemble mean and spread in five-day forecasts. Gradually the differences between the two have become smaller, especially in recent years, attesting to a well-calibrated ensemble. Similar results can be seen for the six and seven-day forecasts (not shown).

Based on the strike probability product available on the ECMWF website (the probability at a given location that a reported TC will pass within a 120 km radius in the next 10 days) a probability verification is routinely computed. Figure 4a.2 shows reliability curves for the last three 12-month periods ending on 30 June. These results indicate the strike probability forecasts are somewhat over-confident at this lead time.

b) Met Office (UK)

Current formulation of the UK Met Office Global Ensemble (MOGREPS-G)

MOGREPS-G is the global component of the Met Office Global and Regional Ensemble Prediction System (MOGREPS). It runs four times a day, at 0000/0600/1200/1800 UTC, at a resolution of N640L70 (c.20 km), out to a forecast lead time of 8 days (192 hours), with output every 6 hours. Each run has 18 members (a control and 17 perturbed members), but the last two runs are time-lagged so that products from each run use a 36-member time-lagged ensemble. MOGREPS currently uses an Ensemble Transform Kalman Filter (ETKF) to perturb the initial conditions in the perturbed members, and includes two schemes, Stochastic Kinetic Energy Backscatter (SKEB) and Stochastic Perturbation of Tendencies (SPT), to account for model error. Soil-moisture, deep-soil temperature and sea-surface temperatures are perturbed to improve near-surface ensemble spread.

Recent verification results

At the Met Office, tropical cyclone tracking (Heming, 2017) is run in real time on the Met Office MOGREPS-G ensemble, the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble (ENS) and National Centres for Environmental Prediction (NCEP) Global En-

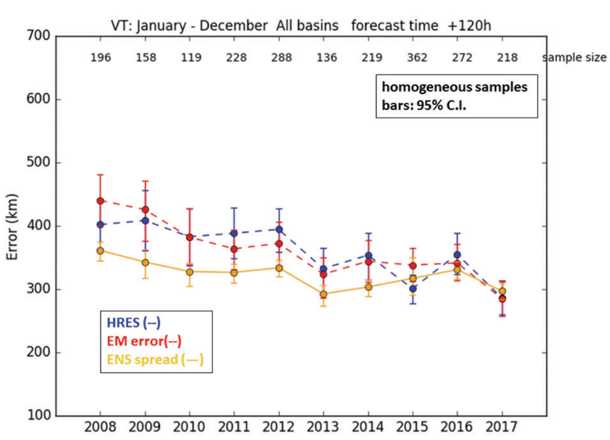


Fig. 4a.1. Annual average of the position forecast error of the ensemble mean (dashed red) and the ensemble spread (solid orange) of the five-day forecasts. Position errors for HRES are shown as well. Verification is against the estimates of observed position reported in near real time. All TC basins are included. Bars indicate 95% confidence intervals based on the bootstrap method.

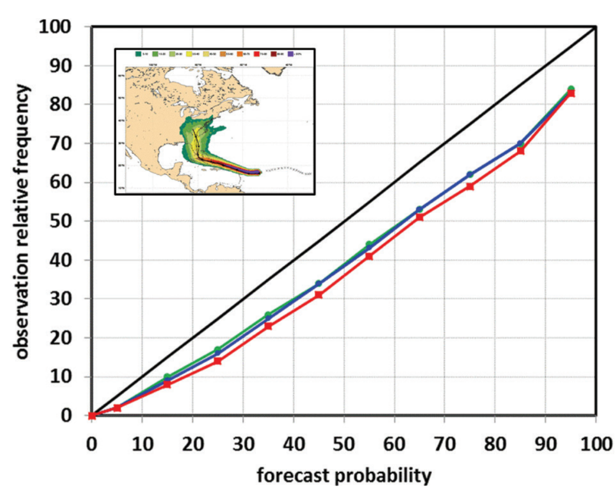


Fig. 4a.2. Probability verification of the 10-day strike probability forecast product (inset box shows an example). Reliability diagram for three 12-month periods; Jul 2015-Jun 2016 (green), Jul 2016- Jun 2017 (blue) and Jul 2017- Jun 2018 (red). Forecast probabilities and observed frequencies are shown as percentages.

semble Forecast System (GEFS). The three ensembles are also combined into a 108-member multi-model ensemble. A range of products, including track and intensity forecasts for both named and forming storms, are produced and distributed to several operational tropical cyclone forecasting centres. The probabilistic forecasts from each global ensemble, and the various multi-model combinations, are evaluated using a probabilistic verification framework. A range of probabilistic verification statistics are calculated to assess the skill, reliability and value of the forecasts, including the Relative Operating Characteristic (ROC), reliability diagrams, relative economic value and Brier Skill Score.

When evaluating ensemble track forecasts, verifying the strike probability forecasts (in this case the probability that a storm will pass within 120 km within the next 7 days) allows for a full analysis of the probabilistic skill and value. Results for the strike probabilities for all named storms in every basin, for the 12-month period January to December 2018 are shown in Figure 4b.1, for the three ensembles and the multi-model ensemble combination. The reliability diagram shows good reliability for all models, with ECMWF ENS showing excellent reliability for all probabilities. MOGREPS-G and NCEP GEFS both show over-forecasting for probabilities 50% and greater. In the relative economic value plot, the multi-model ensemble value curve fully encompasses the three individual models showing the multi-model ensemble combination gives the greatest economic value for all cost-loss ratios. All the models display the greatest relative economic value for very small cost loss ratios (0 to 0.1). For tropical cyclones, user's cost-loss ratios vary significantly but are often very low due to high potential losses.

Planned developments to the MOGREPS-G ensemble

A major change to the MOGREPS-G ensemble is scheduled to go live in Autumn 2019, as the ensemble perturbation system used in MOGREPS-G is changed from Ensemble Transform Kalman Filter (ETKF) to an ensemble of data assimilations (En-4DEnVar; Bowler et al., 2017). In the new system, data assimilation is performed for each member, creating increments relative to its own background trajectory. Figure 4b.2 shows that for 850 hPa winds in the tropics, although in the current ensemble system ETKF gives good spread at initial time, this spread grows too slowly compared to the root mean square error. Comparative trials of the new En-4DEnVar have shown much faster spread growth, with a much better match to errors, which are also reduced compared to the current system. A partial re-centring around the deterministic analysis gives an additional increase in skill and reduces jumpiness. The effect on tropical cyclone track and intensity is currently being evaluated using trial data, but it is hoped that it will lead to a significant improvement in the ensemble spread.

c) *National Centers for Environmental Prediction (USA)*

Forecast Performance

NCEP operates a 21-member Global Ensemble Forecast System (GEFS) to provide probabilistic guidance based on ensemble perturbations from GDAS Ensemble Kalman Filter (EnKF) initial conditions and Stochastic Total Tendency Perturbations (STTP) in the forecast model. In addition, multi-model ensembles based from NOAA (GEFS), Navy (FNMOC), ECMWF, and Canada are routinely produced in operations. Figure 4c.1 shows comparison of track forecast errors from various ensemble means for 2017 Atlantic hur-

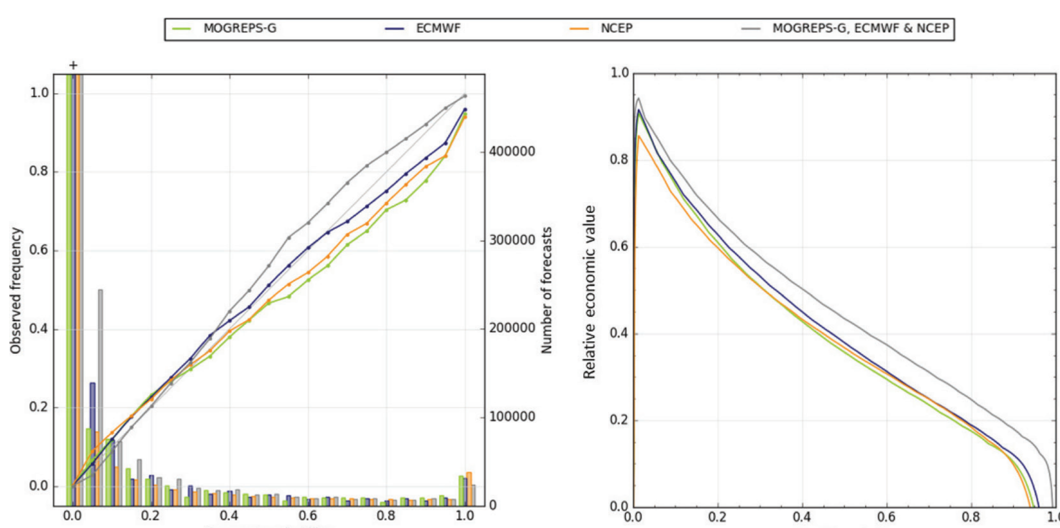


FIG. 4b.1. Verification plots comparing MOGREPS-G, ECMWF ENS, NCEP GEFS and multi-model ensemble forecasts of named storm strike probability for January to December 2018. Reliability diagram including a sharpness diagram on the x-axis (left) and relative economic value plot (right).

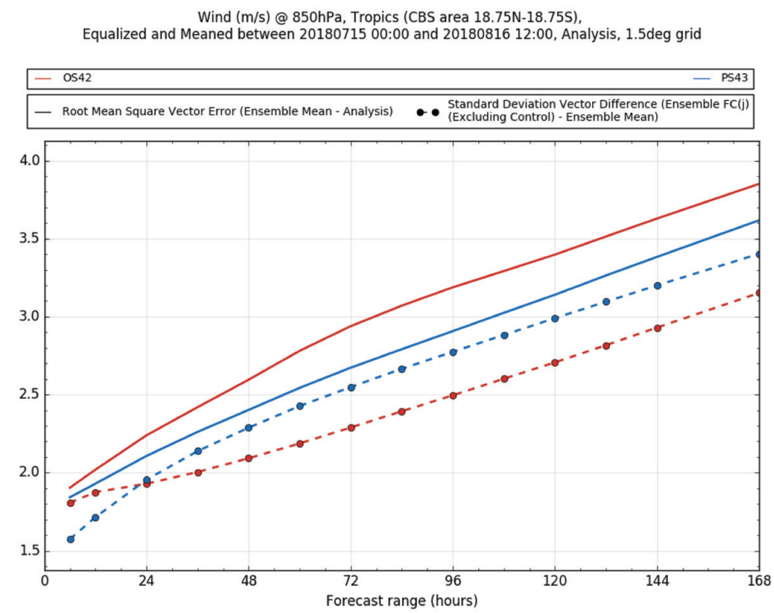


Fig. 4b.2. Root mean square error (RMSE) (solid) and spread (dashed) of wind at 850 hPa for the tropics from the current MOGREPS-G ensemble with ETKF perturbations (red), and the planned MOGREPS-G upgraded ensemble using En-4DEnVar (blue). Both are verified against ECMWF analyses.

ricane season. Multimodel ensemble products are also generated in real-time using GEFS and ECMWF ensembles, which are generally more skilful than individual ensemble means or deterministic forecasts, as shown in Figure 4c.2.

NCEP is developing the next generation GEFS based on FV3 dynamic core with more advanced stochastic perturbation techniques including Stochastic Perturbation of Physical Tendencies (SPPT), Stochastic Kinetic Energy Backscatter (SKEB) and Stochastic Humidity (SHUM) perturbations. Figure 4c.3 shows that the FV3 version of GEFS reduces track error, but increases spread relative to the operational GEFS.

In addition, GEFS will include high resolution (25 km) 31 member ensembles and will provide sub-seasonal (35-day) forecast guidance with this upgrade scheduled for operational implementation in early 2020. The GEFS development also includes production of 20-year reanalysis and 30-year reforecast datasets for calibration and evaluation.

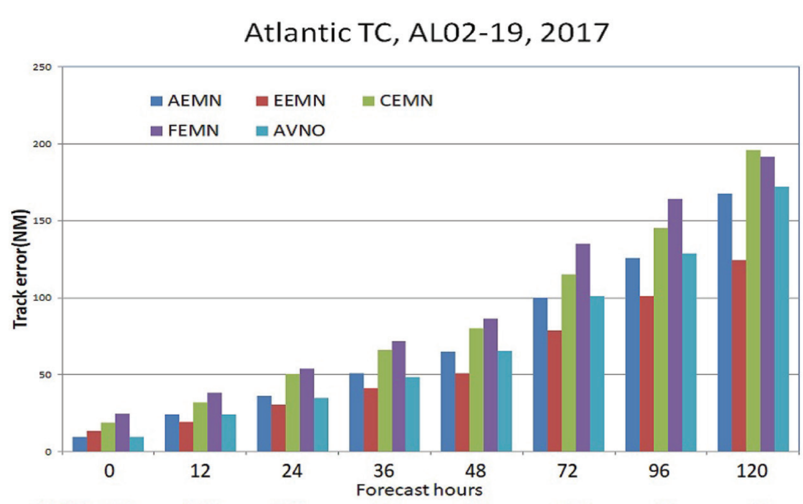


Fig. 4c.1. Track forecast errors from ensemble mean of NCEP GEFS (AEMN), ECMWF (EEMN), Canada (CEMN) and FNMOC (FEMN) compared to deterministic GFS (AVNO) for 2017 in the Atlantic.

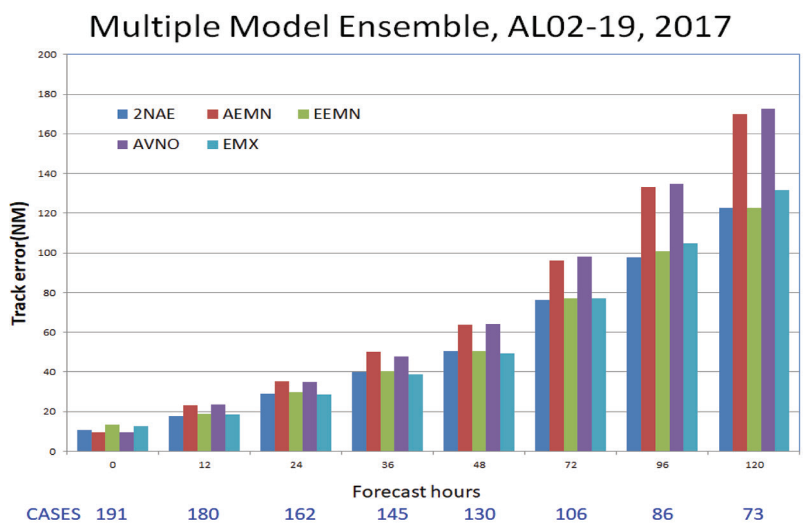


FIG. 4c.2. Track forecast errors from multi-model ensemble mean (GEFS+ECMWF ensemble, 2NAE) compared to individual ensemble means of GEFS (AEMN) and ECMWF (EEMN). Also included are deterministic forecasts from GFS (AVNO) and ECMWF (EMX).

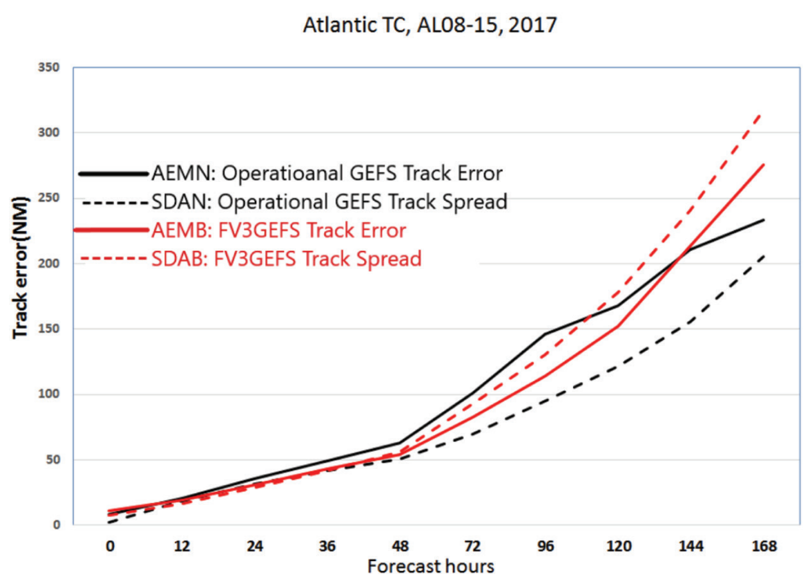


FIG. 4c.3. Track error (solid lines) and spread (dashed lines) in the operational GEFS (black) and the FV3 based GEFS (red) for selected Atlantic TCs in 2017.

High-Resolution Regional Ensembles

It is well known that the track and intensity forecasts made by deterministic dynamic hurricane model systems have its limitations due to various uncertainties existed in both observations and model, including: (1) the errors introduced by the use of imperfect initial conditions, due to observation errors, amplified by the chaotic nature of the evolution equations of the atmosphere, this is often referred to as sensitive dependence on the initial conditions; and (2) errors introduced because of imperfections in the model dynamics and model physics, such as the approximate mathematical methods to solve the equations. Ensemble Prediction System (EPS) is capable of accounting for all kinds of the uncertainties, and hence reducing track/intensity forecast errors by averaging over the ensemble members. HWRF based EPS has been running in real time parallel for the past four years with support from NOAA's Hurricane Forecast Improvement Project (HFIP).

During 2017 multi-model regional ensemble experiment, three model ensembles are used: the HWRF, Navy's

COAMPS-TC and the HMON. A 41-member, multi-regional model ensemble system consisting of HWRF (20 members), COAMPS-TC model (10 members) and HMON model (11 members) was run in real-time.

d) Japan Meteorological Agency

Forecast Performance

GEPS took over the role of JMA's previous ensemble system (Typhoon Ensemble Prediction System - TEPS) and has been providing ensemble forecasts for TC since January 2017. GEPS and TEPS annual mean position errors since 2008 are presented in Figure 4d.1. In 2017, the annual means of ensemble mean position errors for 30-, 54-, 78-, 102- and 126-hour predictions were 114 km (106 km with the GSM), 193 km (182 km), 314 km (300 km), 436 km and 542 km, respectively.

Although position errors of GEPS ensemble mean forecasts were larger than those of GSM in short-range forecasts, GEPS provides useful information on the reliability of TC track forecasts with its ensemble spread. Figure 4d.2 shows the relation between 6-hourly cumulative ensemble spreads in TC position forecasts and ensemble mean forecast position errors in 126-hour prediction. In an ideal ensemble prediction system with a large number of samples, a large position error is observed when the ensemble spread is large. The figure shows that large position errors were seen in 2017 only when GEPS predicted large spreads.

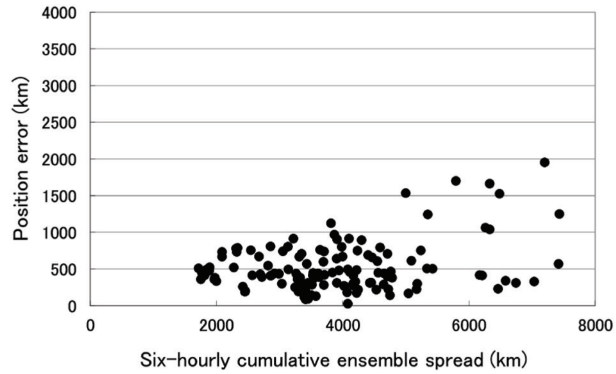


Fig. 4d.2. Relation between 6-hourly cumulative ensemble spread in TC position forecasts and ensemble mean forecast position errors in 126-hour predictions in 2017.

Multi-Model Ensemble Forecasts

JMA introduced multi-model ensemble forecasts in 2015, which consist of ECMWF, UKMO, NCEP and JMA's ensemble systems. These multi-model ensemble forecasts have been available on the RSMC Tokyo's dedicated Numerical Typhoon Prediction (NTP) website since June 2016 in order to help forecasters of the National Meteorological Services of Typhoon Committee Members.

Recently, JMA has been working on the selective consensus method and seeking for the best combination of NWP models. The research so far reveals that the position error can be reduced by up to 150 km in 72-hour forecasts with the best combination. However, what is best differs each time and it is not yet possible to find the best combination on an operational basis. Therefore, it can be said that the future tasks for track forecast are the improvement in accuracy of each model as well as consensus methods.

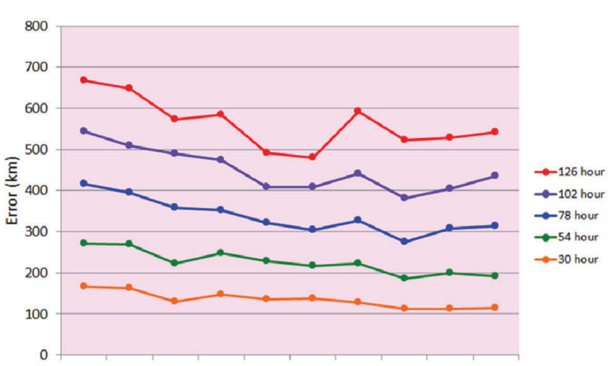


Fig. 4d.1. GEPS and TEPS annual mean position errors since 2008.

Future Plans

JMA will increase the horizontal resolution of GEPS from the current 40 km to 27 km and the ensemble members from 27 to 51 within the next five years. This will also enhance the ability to explain meteorological phenomena including tropical cyclones and achieve higher accuracy of forecasts as well as probability information.

e) Canadian Meteorological Centre

The Canadian Meteorological Centre (CMC) produces 20 ensemble members twice daily at ~39 km resolution and 45 vertical levels using the GEM, to a forecast lead time of 16 days. Once a week, these runs are extended to 32 days to provide monthly guidance. The 20 members are integrated using a multi-physics approach to ensure sufficient ensemble spread (Du et al., 2018) and are initialized using perturbations on a sub-sample of the 256 different analyses that have been generated by the Ensemble Kalman filter (EnKF) assimilation system (Houtekamer et al., 2014). The EnKF uses both perturbed observations and a set of homogeneous, isotropic perturbations to the atmospheric state to represent uncertainty within the assimilation context. The ensemble system is currently not coupled to the ocean, but will become coupled in a change planned for mid-2019.

f) Météo-France

Météo-France operates an Arpège Ensemble (PEARP), composed of 35 members and run 4 times a day. Like its deterministic counterpart, this Arpège ensemble system has a stretched grid, with a focus on Europe (TL798C2.4, 90 vertical levels from 14 m to 1 hPa). There is currently no routine evaluation for TC track prediction.

An Arome-Ensemble system, based on the one operational over France (12 members), has been developed and tested over the Indian Ocean on a few cases of TC during 2018 season (LACy and CNRM laboratories joint work). The 12 initial conditions and lateral coupling are chosen from a global ensemble (ECMWF EPS or PEARP) by clustering. Like its deterministic version, it has shown interesting added value for TC track forecasting, thanks to a more realistic simulation of TC structure. Figure 4f.1 shows the added value of the mesoscale ensemble compared to the ECMWF EPS: the fast moving motion of TC Fakir is more accurately forecasted by Arome ensemble than by ECMWF EPS.

The Arome Ensemble will be run on demand in real time during 2018-2019 Southwest Indian Ocean TC season at the LACy laboratory.

5. Operational Forecasting Centres

a) RSMC Miami (National Hurricane Center, USA)

The National Hurricane Center (NHC) has made tremendous improvements over the past couple of decades in lowering the error of the official track forecasts for TCs. Figure 5a.1 shows a time series of NHC's 24 hour through 120 hour track errors since 1990. The 24–72 hour track forecast errors have been reduced by 70 to 75% since 1990 and error reductions of about 60% have occurred over the past 15 years or so for the 96- and 120-hour forecast periods. In 2017, records for accuracy were set at all time periods and the errors were about 15% lower than the previous records at several forecast times. The primary reason for this success are the advancements in technology, specifically the improvements in the observing platforms and the various modelling systems that NHC uses to make forecasts. The horizontal and vertical resolution, and physics in the models today are far superior to what forecasters had available in the 1990s or prior decades. In addition, NHC has found ways to outperform the individual dynamical models by using a balance of model consensus approaches and experience.

Consensus models are not true forecast models per se, but are merely combinations of results from other models. One way to form a consensus is to simply average the results from a collection or ensemble of models, but other more complex techniques can also be used. The Florida State University Super-ensemble (FSSE) for example, combines its individual components on the basis of past performance and attempts to correct for biases in those components (Williford et al., 2003). A consensus model that considers past error characteristics can be described as a weighted or corrected consensus. On average, these consensus models have been the most accurate track forecast aids over the past several years and NHC forecasters value these models most when making a track prediction. An evaluation over the three years 2015-17 (Figure 5a.2) indicates that the HFIP Corrected Consensus Approach (HCCA), FSSE, and NHC's Track Variable Consensus model (TVCN) were the

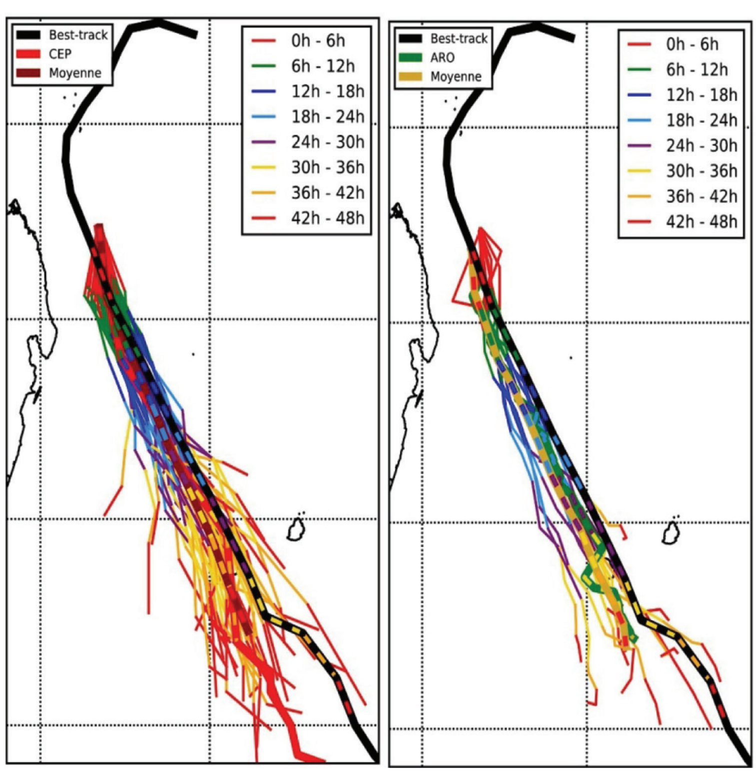


FIG. 4f.1. Best track (black) and spaghettis (colour) of ECMWF EPS (left) and Arome Ensemble (right).Tracks for Cyclone Fakir for 0000 UTC 23 of April 2018. Each colour represents a 6 hour interval of forecast that can be compared to the best-track inner colour. Credit : Sélim Kébir

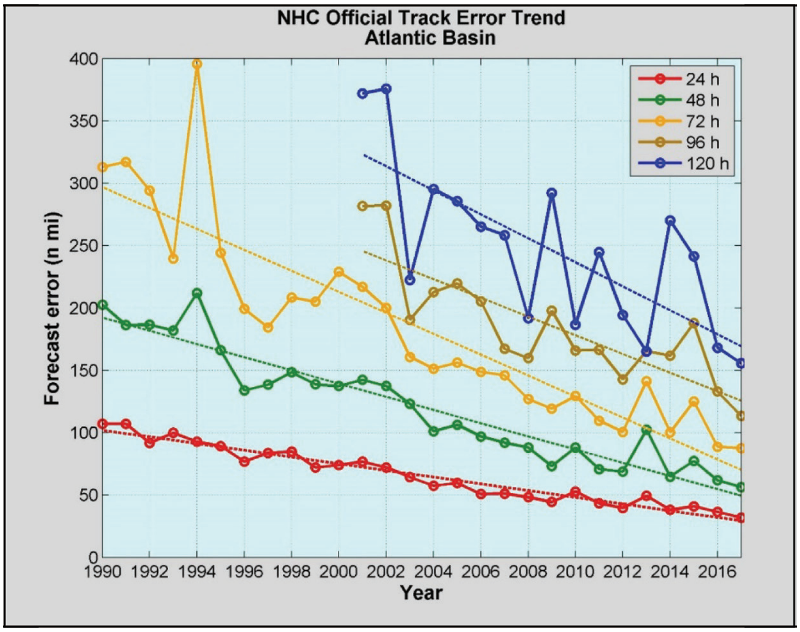


Fig. 5a.1. Recent trends in NHC official track forecast error for the Atlantic basin.

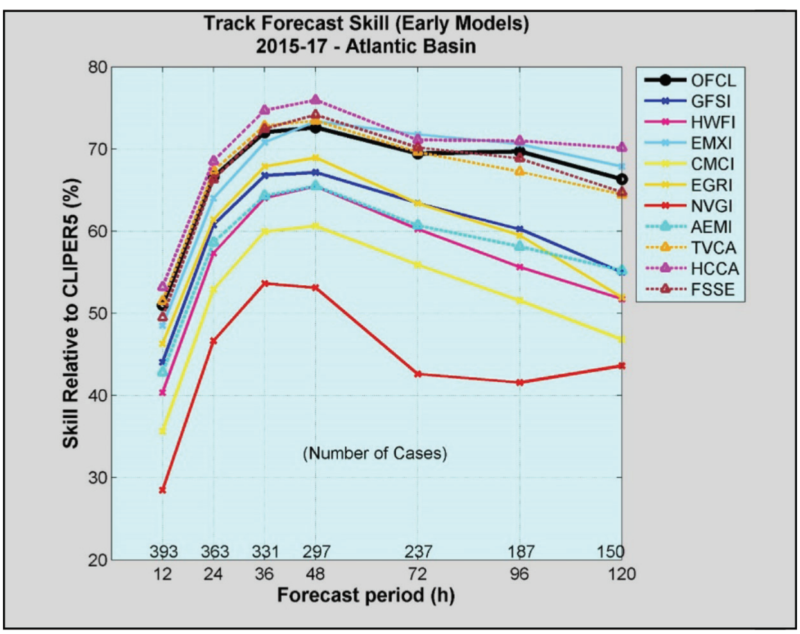


Fig. 5a.2. Homogenous comparison for selected Atlantic basin early track models for 2015-2017.

best-performing models. Table 5a.1 lists all of the models that NHC uses.

Looking ahead to the future, there are a few anticipated challenges in terms of track forecasting. NHC has been experimenting in extending the forecasts out further in time to day 7 (currently 5-day forecasts are made). However, not all of the models are run out to 7 days, which reduces the guidance to NHC and limits the utility of the current consensus model composition at days 6 and 7. The global model ensembles have proven to be quite useful for longer range track prediction. However, it has been shown that the GFS ensemble suite is not dispersive enough to fully capture the uncertainty and possible scenarios. Another issue is how to communicate 7-day track forecasts. The NHC currently uses a combination of deterministic forecasts (i.e. cone graphic) and probabilistic graphics (i.e. wind speed probabilities, storm surge probabilities) to display their prediction and TC hazards. However, it is not known if the

TABLE 5a.1. National Hurricane Center forecasting aids.

Tracker Name	Forecast Aid Description	Type
OFCL	Official NHC forecast	
HWRF	HWRF Model	Regional model
HMON	HMON	Regional model
GFSO	NWS/Global Forecast System	Global model
AEMN	GFS ensemble mean	Global model
UKM	UK Met Office model, full resolution tracker	Global model
EGRR	UK Met Office model, reduced resolution tracker	Global model
UEMN	UKMET ensemble mean	Global model
NVGM	NAVGEM	Global model
CMC	Environment Canada global model	Global model
NAM	NWS/NAM	Regional model
CTCX	COAMPS-TC using GFS initial and boundary conditions	Regional model
EMX	ECMWF global model	Global model
EEMN	ECMWF ensemble mean	Consensus
TABS	Beta and advection model (shallow layer)	Single-layer trajectory
TABM	Beta and advection model (medium layer)	Single-layer trajectory
TABD	Beta and advection model (deep layer)	Single-layer trajectory
CLP5	CLIPER5 (Climatology and Persistence model)	Statistical (baseline)
TCLP	Trajectory-CLIPER model	Statistical (baseline)
OFCI	Previous cycle OFCL, adjusted	Interpolated
HWFI	Previous cycle HWRF, adjusted	Interpolated-dynamical
HMNI	Previous cycle HMON, adjusted	Interpolated-dynamical
CTCI	Previous cycle CTCX, adjusted	Interpolated-dynamical
GFSI	Previous cycle GFS, adjusted	Interpolated-dynamical
UKMI	Previous cycle UKM, adjusted	Interpolated-dynamical
EGRI	Previous cycle EGRR, adjusted	Interpolated-dynamical
NVGI	Previous cycle NVGM, adjusted	Interpolated-dynamical
EMXI	Previous cycle EMX, adjusted	Interpolated-dynamical
CMCI	Previous cycle CMC, adjusted	Interpolated-dynamical
AEMI	Previous cycle AEMN, adjusted	Consensus
UEMI	Previous cycle UEMN, adjusted	Consensus
FSSE	FSU Super-ensemble	Corrected consensus
GFEX	Average of GFSI and EMXI	Consensus
TCON	Average of EGRI, GFSI, and HWFI	Consensus
TCCN	Version of TCON corrected for model biases	Corrected consensus
TVCN	Average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus
TVCA	Average of at least two of GFSI EGRI HWFI EMXI CTCI	Consensus
TVCE	Average of at least two of GFSI EGRI HWFI EMXI CTCI HMNI EMNI	Consensus
TVCX	Average of at least two of EMXI (double weight) GFSI EGRI HWFI CTCI	Consensus
TVDG	Average of at least two of GFSI (double weight) EMXI (double weight) EGRI (double weight) CTCI HWFI	Consensus
TVCC	Version of TVCN corrected for model biases	Corrected consensus
HCCA	Weighted average of AEMI, GFSI, CTCI, DSHP, EGRI, EMNI, EMXI,HWFI, LGEM	Corrected consensus

current methodology would be appropriate to extend in forecast time.

A recent study (Landsea and Cangialosi, 2018) discussed the potential limits of TC predictability and how these limits could be reached in the near future. Although it remains unknown when this might occur, it is agreed upon that perfect forecasts are not possible.

b) Joint Typhoon Warning Center (USA) ***Consensus Forecast Aids***

JTWC was one of the first TC forecasting sites in the world to implement consensus forecasting in the late 1990s

(Sampson and Schrader, 2000). At that time, limitations in model availability meant that the consensus was fairly limited. However, over the years, the availability of high-quality NWP models capable of generating a high confidence track forecast has dramatically increased. Beginning in the early 2000s, JTWC began to have access to increasing numbers of these models and started to use the consensus methodology to improve track forecasts. Today, the JTWC utilizes an internally generated, non-weighted, consensus forecast track aid, (CONW), consisting of 10 individual members, including a mix of global, regional and ensemble models (Table 5b.1).

The JTWC consensus requires a minimum of two of the 10 members be present in order to generate the CONW tracker. NRL and JTWC annually review the performance and reliability of various models to assess the sensitivity of CONW accuracy to each member and to optimize overall accuracy of the consensus. For instance, in 2017, the USAF Global Air-Land Weather Exploitation (GALWEM) model (AFUM) replaced the JMA TC ensemble mean track forecasts (JENS) in the track forecast consensus. An example of the current CONW members as displayed in the Automated Tropical Cyclone Forecast (ATCF) system is shown in Figure 5b.1.

Forecast Performance

The JTWC provides TC track, intensity, and wind field forecasts for US Government customers in the North Pacific, South Pacific, and Indian Ocean basins. Figure 5b.2 shows the mean track errors for the Northwest Pacific (since 1970) and Southern Hemisphere regions (since 1985).

TABLE 5b.1. Models used in JTWC's CONW consensus

Tracker Name	Model Name	Type	Date in CONW
AVNO	GFS	Global	2002
NVGM	NAVGEM	Global	2002
AFUM	USAF GALWEM	Global	2017
EGRR	UK Met Office	Global	2002
ECMF	ECMWF	Global	2007
JGSM	Japan GSM	Global	2002
HWRF	HWRF	Regional	2014
CTCX	COAMPS-TC	Regional	2017
AEMN	GEFS ensemble mean	Global	2014
EEMN	ECMWF ensemble mean	Global	2016

JTWC began producing 120-hour forecasts for the Northwest Pacific in 2000, and in the Southern Hemisphere in 2010. In 2000, mean track error at 72 and 120 hours were near 400 km and 600 km. In 2017, mean track errors at 72 and 120 hours were near 250 km and 415 km. The same general trends are evident in the southern hemisphere as well. Taking into account intra-seasonal variability, the implementation of consensus track forecasting at JTWC along with great improvements in numerical modeling capabilities has significantly and steadily reduced forecast track errors, particularly in the later forecast periods.

c) RSMC Tokyo (Japan Meteorological Agency)

Forecast Performance

Annual mean errors in TC track forecasts covering 24-,

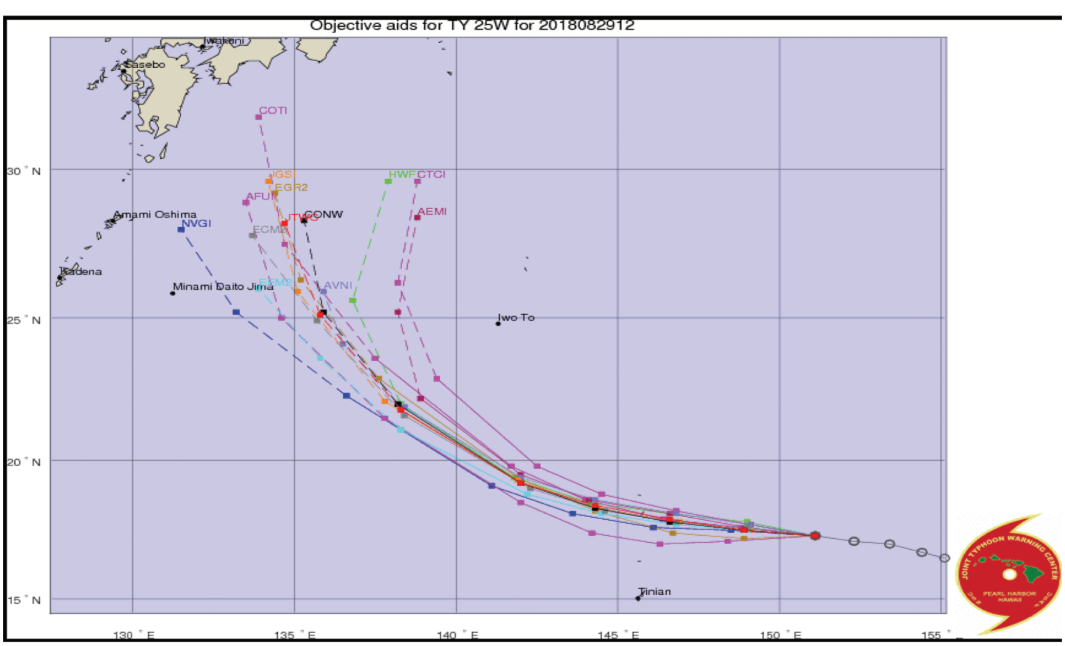


FIG. 5b.1. Example of current CONW trackers in ATCF display

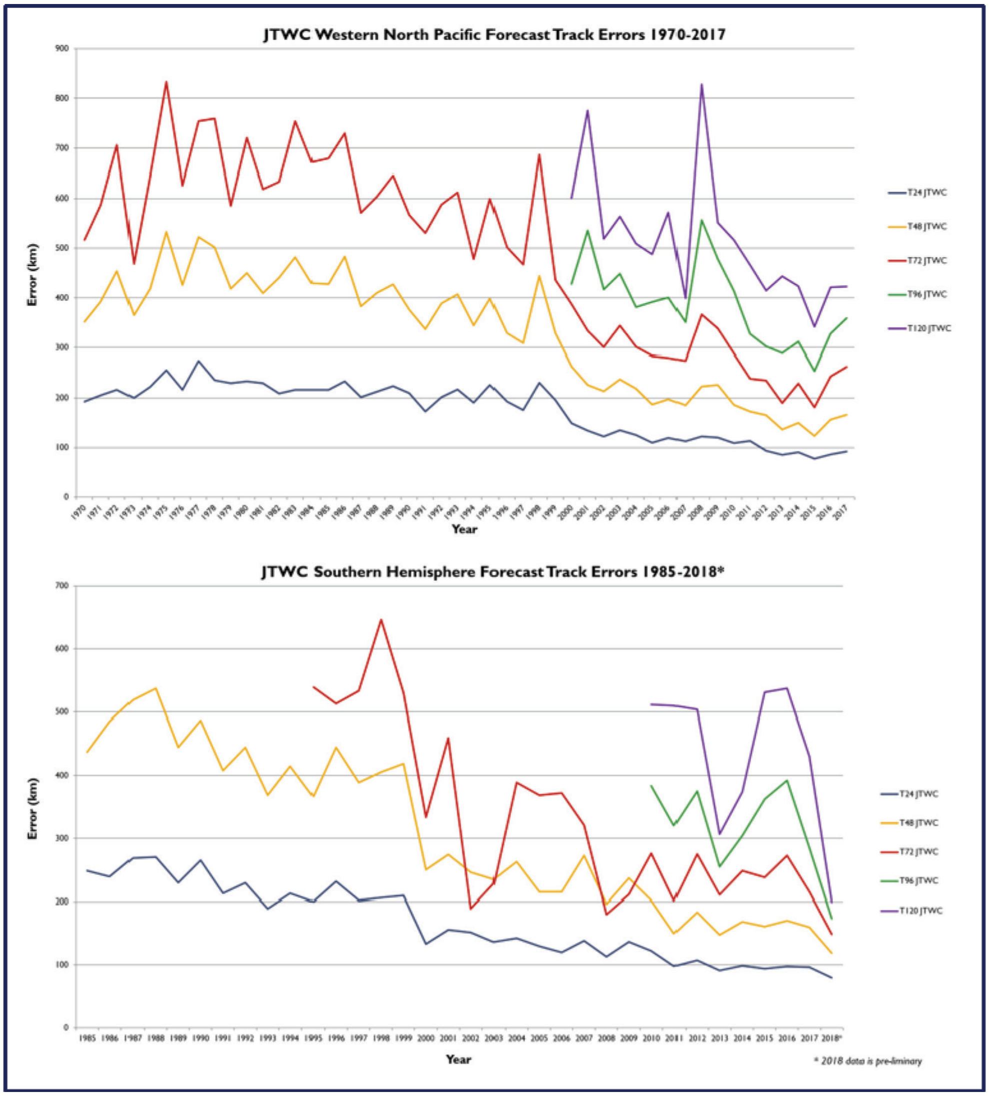


FIG. 5b.2. JTWC TC forecast errors (km) for the Northwest Pacific (top) and Southern Hemisphere (bottom).

48-, 72-, 96- and 120-hour operational track forecasts have steadily reduced in the long term (Figure 5c.1). Results for 2016-17 were affected by the characteristics of TCs in those seasons since the forecast skill against persistence (not shown) mostly continued increasing compared to 2015.

Reduction of Forecast Circle Radii

The reduction of forecast circle radii is a remarkable improvement in performance related to track forecasts since IWTC-8. Based on TC track forecast improvements made in recent years via NWP model enhancement and other forecast techniques, JMA reduced the radius of forecast circles in its official forecasts by 20 - 40% (depending on TC direction and speed) in June 2016. This change addresses the issue of over-dispersiveness of warning areas. The size of forecast circles is determined so that forecast track falls within the circles in a probability of about 70%. For each forecast time, circle size is defined based on the speed and direction of movement. Furthermore, for forecast times 96 and 120 hours, circles are dependent on the forecast reliability estimated by the results of GEPS for each TC. Changes in forecast circle size in typhoons with two directions are shown in Figure 5c.2.

Consensus Forecasting

Research on a selective consensus technique for TC track forecasts using multi-model ensembles was conducted in 2014 by the Meteorological Research Institute of JMA (Nishimura and Yamaguchi, 2015). Based on this, JMA verified the accuracy for track forecasts from 2012 to 2014 using four NWP models (JMA, ECMWF, UKMO, NCEP) and proved the effectiveness of the simple consensus meth-

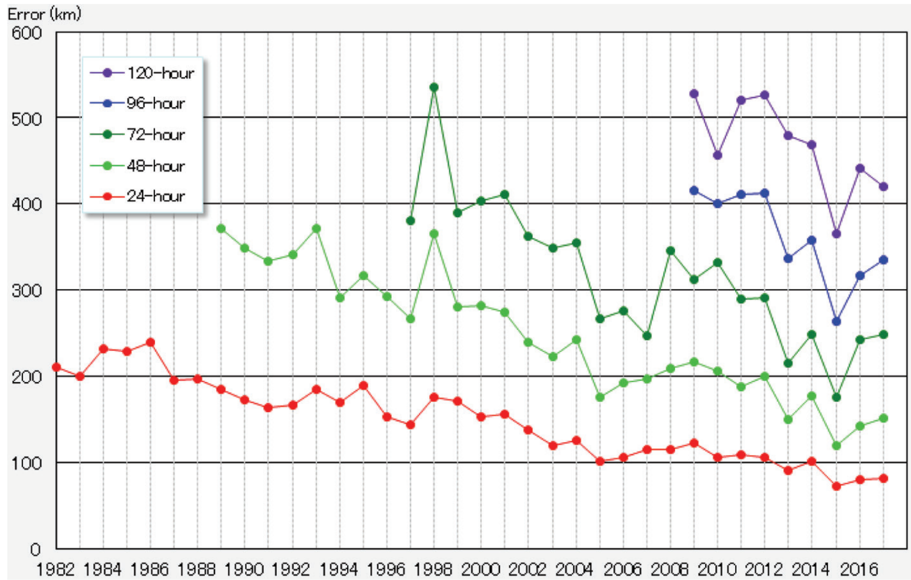


Fig. 5c.1. Annual mean position errors in JMA 24-, 48-, 72-, 96- and 120-hour operational track forecasts

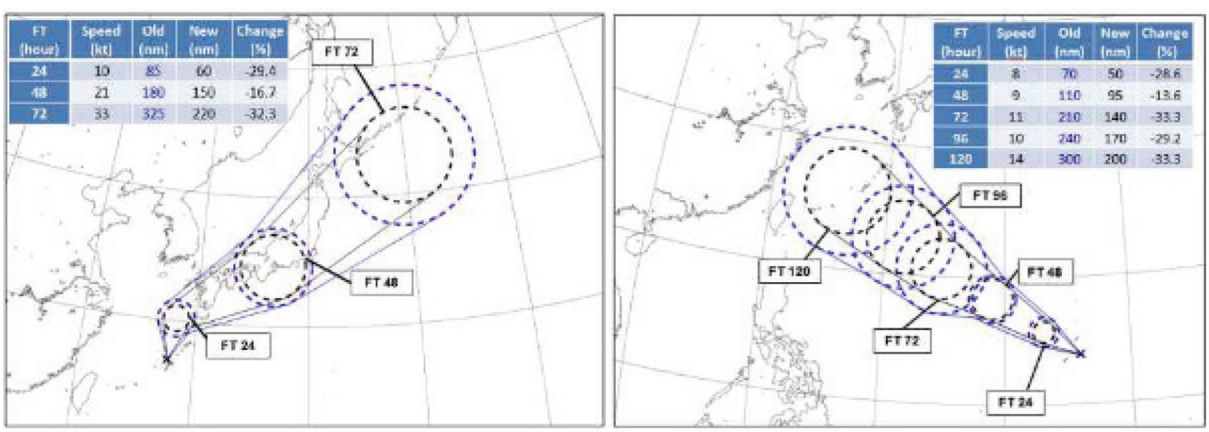


Fig. 5c.2. Changes in forecast circle size; blue: old circles; black: new circles. Left: Northeastwards moving Typhoon Chan-hom (2015). Right: Northwestwards moving Typhoon Vongfong (2014)

od, with the combined four NWP model consensus achieving the highest accuracy at all forecast times (24, 48 and 60 hours). Therefore, this method has been adopted for JMA's operational TC forecasts as the first guess since 2015. In addition, JMA has started improving a model and method to use in the following season by conducting verification at the end of every year.

JMA's operational forecasts were mainly based on GSM until 2014, so accuracy was almost the same as GSM. In 2015, the four NWP model consensus was adopted for JMA's operational tropical cyclone forecasts as the first guess. Figure 5c.3 shows that operational forecasts (black bars) had lower errors than GSM (pink bars) and the four model consensus (GEUA; red bars) gave the most accurate results.

d) Bureau of Meteorology (Australia) Forecasting Method and Performance

Performance of Bureau of Meteorology (BoM) in operational track forecasting is summarized in Figure 5d.1 and Table 5d.1, with values averaged over the past five years. The BoM issues a forecast track out to 120 hours with 'uncertainty areas' which represents the range of possible tracks between analysis time and a certain time.

The BoM's standard track forecast process involves a consensus of model guidance, shifting to the analysis position and using average motion to generate a track. Once a system has developed, the standard consensus is the latest

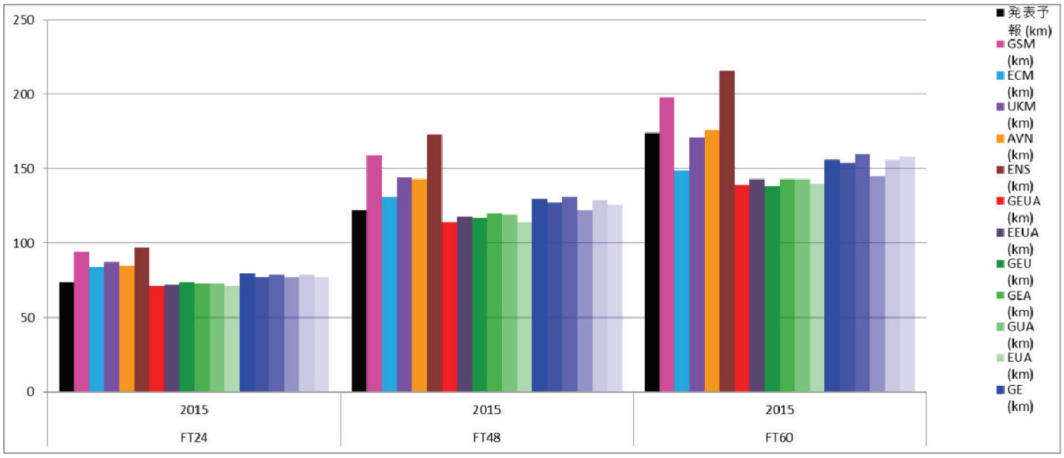


FIG. 5c.3. Track errors (km) for combinations of the four NWP models for 2015, with operational errors (black bars). Forecast times 24, 48 and 60 hours.

(deterministic run) from ECMWF, GFS, UKMO, HWRF, ACCESS-TC (3-day high resolution) or ACCESS-TCX (5-day high resolution), COAMPS (CTCX), JMA, and optionally ACCESS-G or ACCESS-R. When or if available, ECMWF and UKMO ensemble means are also used. The shifting to analysis position and then average track motion avoids jumps in the consensus track when/if model guidance stops being available. The uncertainty areas (for 24-, 48-, 72- and 120-hour forecasts) are initially based on radii calculated from the consensus tracks and climatological track error. The forecasters then have the option to adjust the shape of the uncertainty areas based on guidance (deterministic and ensemble).

The BoM has a requirement, in some cases to start issuing tropical system forecasts prior to a tropical disturbance forming. Due to limited guidance at that time range, the usual consensus is two or more runs of the ECMWF, GFS, UKMO, JMA and ACCESS-G models. Depending on the situation, the inclusion of models and past runs is based on comparison to how well the resulting spread matches ensemble guidance. Due to uncertainty in analysis position (or no analysis position at all) the track is generally based on average guidance position or average initial position and average motion (not shifted to analysis). The included models and ensemble guidance is used to inform the uncertainty areas. These early tracks generally exhibit greater forecast error due to errors in where the system develops compounding with errors in track forecasting.

Forecasting Challenges

The BoM's main challenging scenarios are:

- Small systems with different steering flows depending on depth of the system, particularly since the small systems can be under forecast by the lower resolution global models.

- Small, well structured systems which could move offshore and then rapidly intensify.

- Two or more systems interacting which increases forecast error.

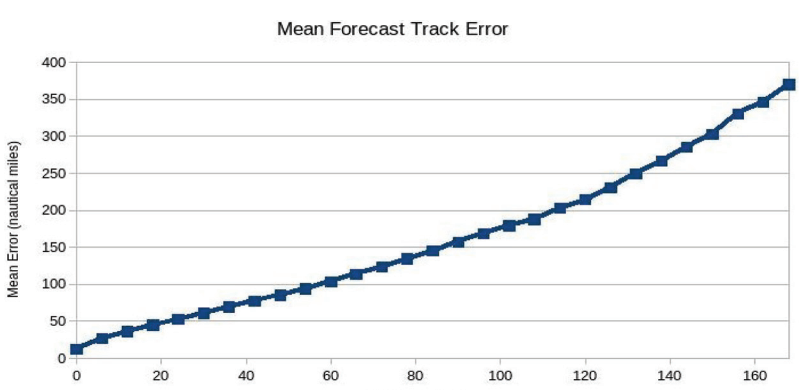


FIG. 5d.1. Mean forecast track error (great circle distance) based on best track or operational best track for seasons 2013-14 to 2017-18 inclusive.

Table 5d.1. Percentage of times the analysis position at the forecast hour was within the uncertainty area for seasons 2013-14 to 2017-18 inclusive (preliminary results). BoM has only been doing 120-hour uncertainty areas since November 2015.

Forecast Hour	Number in area	Total number	Percentage in area
24	614	750	81.9
48	516	641	80.5
72	395	510	77.4
120	90	106	84.9

- Bifurcation and representing that in a forecast track. We generally try and forecast the most likely scenario and then modify the uncertainty area (which is initially uniform in all directions) to show the remainder of the risk, with the risk being based on guidance.

Future Developments

In the future, we'd like to move to a forecast method which uses a 'super ensemble' (multiple deterministic models, ensemble models and multiple runs) to generate the track and uncertainty areas. The uncertainty area for a specific time would be based on a contour of a heat map, with the contour value having been calibrated on past events and representing the area the system should be in some percentage of the time.

With improving NWP skill and availability of ensemble guidance, BoM aims to increase automation of the track generation process but there is still concern at how to interpret outliers and to identify scenarios requiring forecaster scrutiny. This includes cases of both high and low model spread.

In line with an increased focus on hazard rather than the TC, forecasters are focussing on the occurrence of wind thresholds for example point based onset of gales. While track is a key element, this approach requires a more sophisticated method of combining track, structure and intensity details rather than viewing these parameters in isolation.

e) RSMC La Réunion (Météo-France)

TC track forecasting at RSMC La Réunion essentially relies on the track forecasts provided by the main NWP models available at the Centre. Indeed, while a TC forecaster can beat any numerical model for an individual forecast at a given time, it has become virtually impossible to statistically beat them for a large sample of forecasts, like for a whole cyclone season.

Given this reality, the challenge for the TC track forecasting is then to optimize the forecasts provided by the models. The best that can be done is to follow the currently deemed 'best performers' or the consensus of the models that have been assessed to outdo the individual best models. A few years ago, during a certain period the ECMWF HRES model outperformed the other models by such a margin that any consensus built by adding one or several other models just degraded the performance of the track forecast. Currently, this situation does not prevail anymore, as the GFS model has caught up with the ECMWF model in the Southwest Indian Ocean basin. Therefore the outputs from these two models form the base of the track forecasts for a more or less weighted consensus, the main adjustment variable being the weight given to each in the consensus.

However, in certain circumstances a consensus would not be appropriate. This is the case when there is a too large discrepancy between the models with track forecasting options that really differ. In such a situation it is the role of the TC forecaster to 1) try to understand the origin of the differences by examining the different fields and 2) to make a choice (with no guarantee of making the right one despite his expertise).

While deterministic track forecasts remain the main output of RSMC La Réunion, they say nothing about the inherent uncertainty in forecasts. Since December 2011 the RSMC La Réunion website displays dynamical probabilistic cones of uncertainty around the official track forecasts of the RSMC. Instead of including a cone of uncertainty based on the average error climatology, a more sophisticated probabilistic method is used which more realistically takes into account the real degree of uncertainty of each individual TC track forecast situation. The spread information included in the ensemble forecasts (EPS from the ECMWF) is used to better assess the uncertainty and construct an EPS-based probabilistic adaptive cone to convey this uncertainty. Dupont et al. (2011) shows that this methodology has skill over just using climatology.

The TC track forecast performance of the RSMC La Réunion is shown in Figure 5e.1. While short range forecasts have demonstrated very little or no improvement, forecasts at 48 hours lead time and beyond continue to show improvement. In the past few seasons the forecast errors for 60 hours and beyond have shown the most spectacular reduction, which means that the natural trend of increasing error with time becomes drastically flatter. So much so that the gap between 36-hour forecast errors and 72-hour forecast errors has been divided by more than three since IWTC-8. Also, 72-hour forecasts are now better than the 48-hour forecasts were just one or two years ago.

f) RSMC Nadi (Fiji Meteorological Service)

RSMC Nadi area of responsibility (AOR) is from the Equator to 25°S and 160°E to 120°W covering over 20 million square miles of ocean. Apart from a few ship reports there are no drifting buoys for open waters observations. Land based observations are also very few in most of the island countries which fall in the AOR. In brief this is a data sparse region.

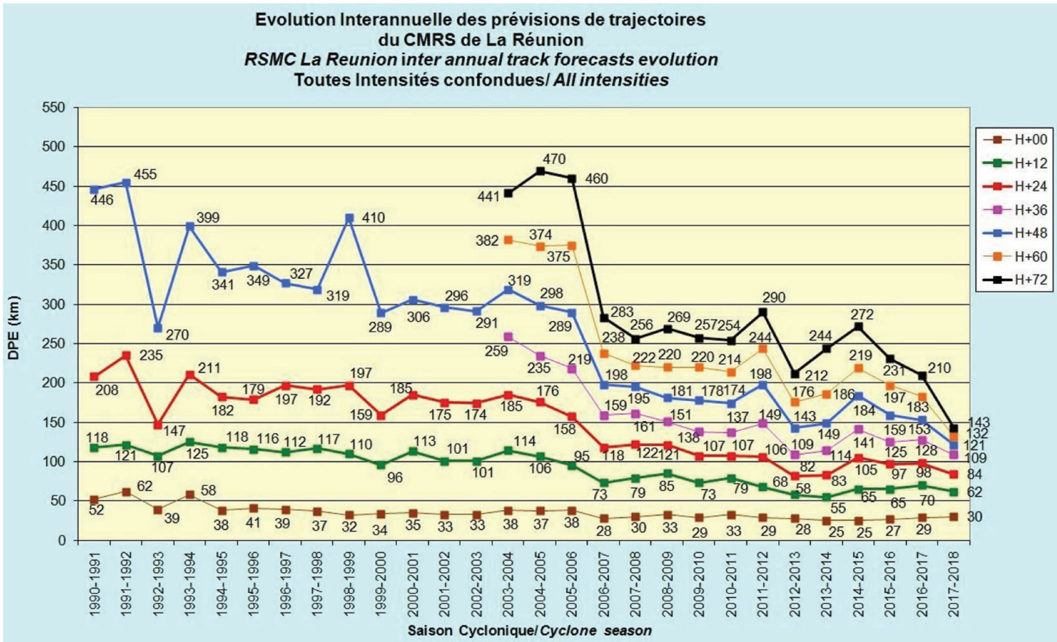


FIG. 5e.1. RSMC La Reunion Direct Positional Errors of track forecasts (in km).

RSMC, Nadi issues a 3-day Tropical Cyclone Outlook every day at 0400 UTC from 1 November to 30 April.

TC intensity forecasting especially beyond 24 hours remains a challenge at RSMC Nadi. Each TC is different and behaves somewhat differently to similar environmental condition. For most of the systems, normal development of Dvorak T = 1 per day is applied. Midget systems which intensify rapidly are the most difficult ones to forecast where the Dvorak constraint is usually broken.

For analyzing a system, we receive Himawari satellite images every 30 minutes from HimawariCloud which is of high resolution and HimawariCast every 10 minutes at lower resolution. The Dvorak technique is applied for analyzing the intensity. For low-level circulation centre location, satellite loops, ASCAT passes, land based observations (if close to land) are used.

The intensity analysis can only be verified if there is an ASCAT pass (useful when the TC intensity is below 50 knots) or land-based observation stations near to the TC centre.

The Australian TC Module is used for TC official forecasts, tracks maps and issuance of most products.

For forecast intensity and track, RSMC Nadi is dependent on global model guidance which is imported from the JTWC website to TC Module. RSMC Nadi does not run any locally developed numerical models. Guidance from GFS, UKMO, JTWC, GFDL and JMA are available from the JTWC collaboration site. ECMWF is entered manually from Tropical Tidbits. A consensus forecast track is prepared from all the above models using TC Module. Sometimes the forecast track is shifted when the TC is approaching a land area. The track is moved a little closer to land area mainly for warning purpose. This is after the experience with TC Evan.

For the intensity forecasts, model guidance is used together with the Dvorak rules for intensification and weakening.

Examples of TC track forecast errors for TCs in the RSMC Nadi AOR in the 2016-17 and 2017-18 seasons are shown in Figures 5f.1 and 5f.2.

g) RSMC New Delhi (India Meteorological Department)

Forecasting Tools

For short range forecasting (up to 24 hours) IMD uses synoptic, statistical, satellite and radar guidance. NWP guidance is mainly used for 24-120 hour forecasts. Consensus forecasts that gather all or part of the numerical forecast tracks and use synoptic and statistical guidance are utilised to issue official forecast (IMD, 2013).

The NWP models used by IMD include individual deterministic models, a multi-model ensemble (MME) and single model ensemble prediction system (EPS). The deterministic models include GFS, the regional WRFDA-WRF-ARW model with 9 km and 3 km horizontal resolutions, HWRF, Unified Model (12 km resolution) and Unified regional model (4.5 km resolution) adapted from UKMO. IMD also makes use of NWP products prepared by some other operational NWP centres such as ECMWF, NCMRWF, JMA, UKMO and Météo-France.

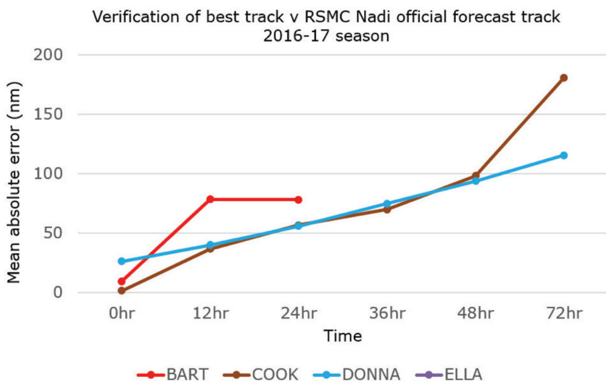


Fig. 5f.1. RSMC Nadi track forecast errors for cases in the 2016-17 season

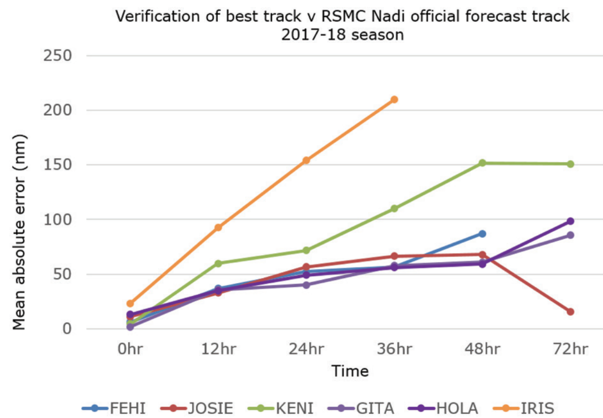


Fig. 5f.2. RSMC Nadi track forecast errors for cases in the 2017-18 season

The MME technique (Kotal and Roy Bhowmik, 2011) includes five member models; WRF (ARW), GFS (IMD), GFS (NCEP), ECMWF and JMA. The MME product is available about 9 hours late, so, for example, the 36-hour MME forecast is used for 24-hour official forecasts.

The Ensemble forecast products from ECMWF, NCEP, UKMO, CMC and JMA are available near real-time. A super-ensemble is also developed based on above ensembles. In India, NCMRWF and IMD run the Unified Model Ensemble Prediction System (UM-EPS) and Global Ensemble Forecasting System (GEFS) respectively to provide 7-day forecasts based on 0000 UTC initial condition with a resolution of 12 km each (RSMC, New Delhi, 2018).

Forecast Performance

There has been a significant improvement in TC track forecasting over the north Indian Ocean by IMD in recent years. The average track forecast errors of IMD during 2014-18 were 81, 128, 180, 260 and 285 km respectively for 24-, 48-, 72-, 96- and 120-hour forecasts (RSMC, New Delhi, 2018). The forecast performance of individual NWP models, the MME and IMD's official forecast is shown in Figure 5g.1. It is found that the MME outperforms the individual models. The official forecast accuracy is similar to MME forecast accuracy. In this figure, the operational forecast error has been compared with model errors with a 12 hour lag (i.e. 24-hour model error is compared with 12-hour official error), as the model products used for official forecasts are available with almost 12 hours delay. Considering individual NWP models, it is observed that the HWRF model has the lowest error up to 36 hours. This is followed by the ECMWF model for these lead times. The UKMO model shows the lowest errors for the longer lead times.

The track forecast errors based on 2014-2018 as compared to previous five years are shown in Figure 5g.2.

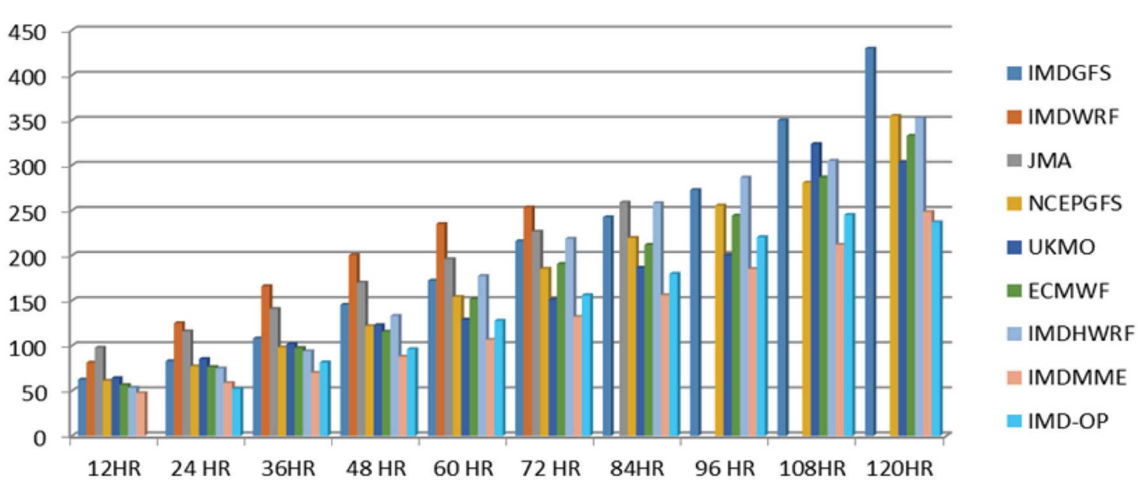


Fig. 5g.1. The average track forecast errors of various NWP models used by IMD and IMD's official forecasts during 2014-18

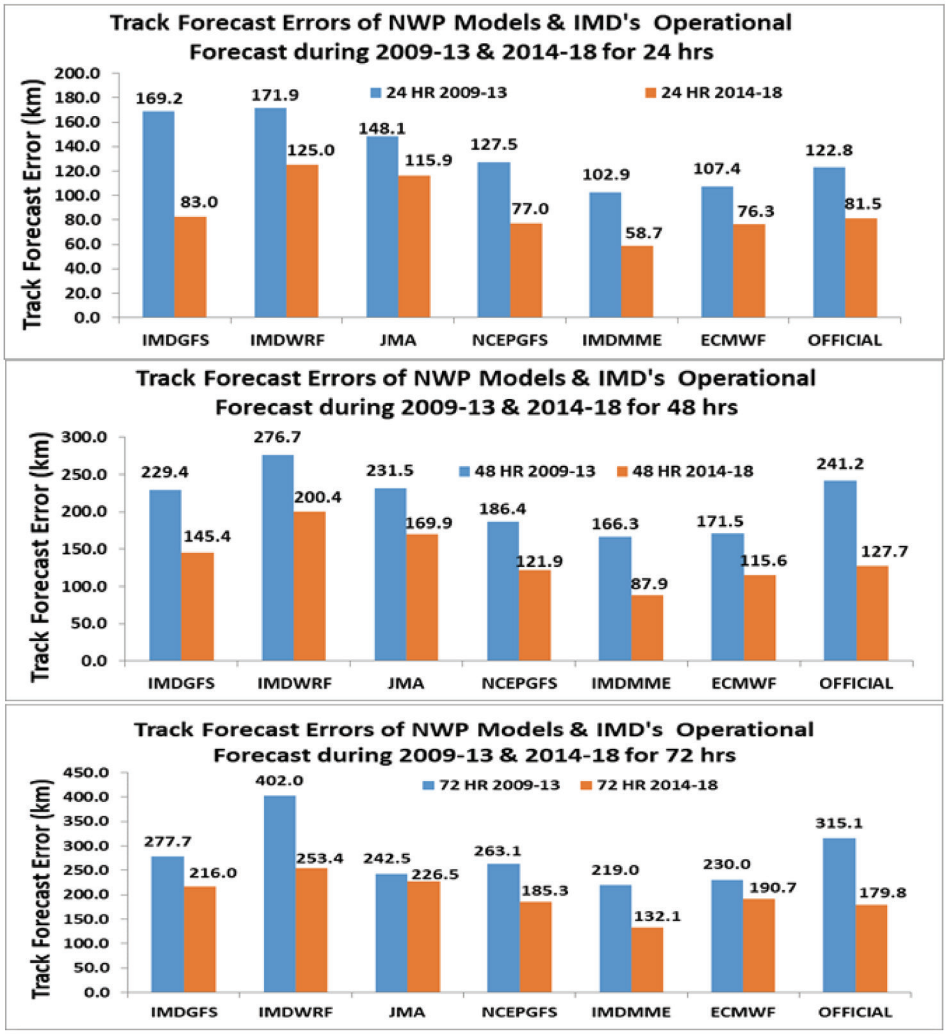


FIG. 5g.2. Comparative track forecast errors of IMD and NWP models for TCs over the north Indian Ocean during 2009-13 and 2014-18

There is an improvement of 25-35% for 12- to 24-hour lead times and of 35-45% for 36- to 72-hour lead times. There has been continuous improvement in track forecasts by IMD (Mohapatra et al., 2013a) due to modernization of IMD (Mohapatra et al., 2013b).

Cone of Uncertainty

Considering the improvement in track forecasting, the radii of the cone of uncertainty based on past five years average track forecast error have been reduced by about 20-25% since 2014 (Mohapatra et al., 2012, 2017; Figure 5g.3). It is being revised every five years and revision is due in 2019. It has its own limitation due to its static nature, especially in the case of recurving tracks. A dynamical cone of uncertainty, which has not been introduced to date, needs to be implemented.

Sudden change in track

Situations that are difficult to forecast TC track include recurving TCs, rapid movement of TCs during landfall, Slow movement or stationarity of TCs near the coast (Mohapatra and Bandyopadhyay, 2012) and sudden change in direction a few hours before landfall. It is found that the error is higher by about 5-20% for 12- to 72-hour lead times in case of TCs with rapid track changes as compared to the mean track forecast errors based on the data of 2014-18.

Pre-genesis Forecasting

RSMC New Delhi issues TC track forecasts valid up to 120 hours from the stage of deep depression (28-33 knots), in anticipation of intensification into a TC (34 knots or more). This practice has been operative since 2009. It has been further revised in 2018 with track forecasts from the

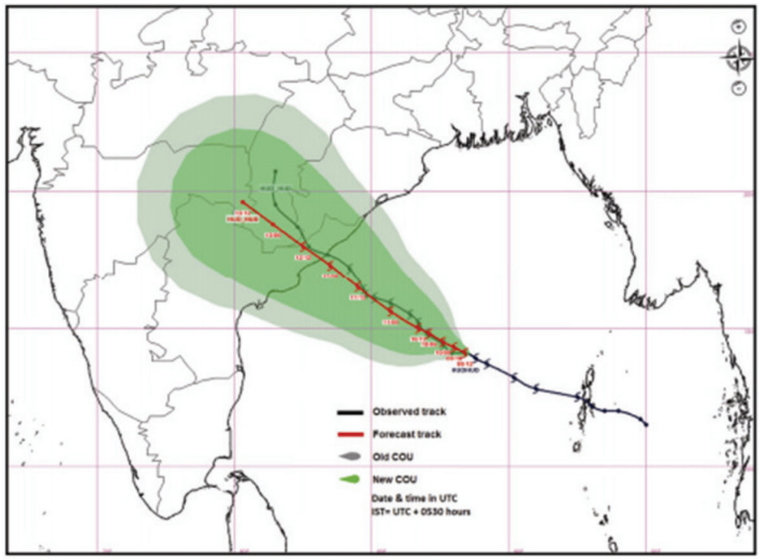


FIG. 5g.3. Observed and forecast tracks along with old and new cones of uncertainty in case of TC Hudhud. Initial time 1200 UTC 9 October 2014

stage of depression (17-27 knots) with a validity period of 72 hours. These forecasts are issued five times a day. However, the track forecast error in the pre-genesis stage is relatively high. This may be due to relatively high initial errors in the estimation of the centre of the depression unlike that of TCs and most of the models do not use vortex initialization or relocation at this stage due to lack of TC vital information from the forecasters.

End-of-life Track Forecasting

RSMC New Delhi issues track forecasts objectively indicating forecast latitude and longitude, whilst the former TC is expected to remain as a depression, even after landfall. No location in terms of latitude and longitude is given in the forecast of a low pressure area (< 17 knots). When the depression weakens into a low pressure area, forecast responsibility is handed over to the local state meteorological centre. Forecasts of the movement of low pressure areas are not provided by IMD.

h) Canadian Hurricane Centre

Regardless of the activity in the Atlantic Basin throughout any season, CHC tend to respond to 4 or 5 TCs or transitioning systems in our Response Zone on any given year. We almost always inherit tracks from the NHC as TCs approach from the south, so by default we have a good initial guidance from their tracks. Historically, we have become accustomed to utilizing other models other than the Canadian Global Deterministic Prediction System as its performance was substandard. However, significant progress has been made with the global model since the introduction of the coupled atmospheric-ocean physics.

Climatologically, tropical systems at our latitudes tend to have a 'well-behaved' track as the upper flow is usually well defined. The along-track error component usually dominates the cross-track component for this reason. The error components we use to define our error ellipses (upon which our cone of uncertainty is based) are in Table 5h.1.

These numbers have been in use for a few years and are likely in need of updating in the next few years. The biggest challenge in track forecasting for our latitudes remains the speed of systems as they begin to interact with the mid-latitude upper flow.

CHC find it difficult to find forecast utility of the ensemble information. We do expose some of our clients to 'track spaghetti plots' (i.e. ensemble member tracks) at times to give an indication of NWP possibilities. We are also considering having a look at these member tracks superimposed over the climatological error cone to see how dispersed any specific ensemble run is compared to the error cone values to see if information is revealed. It might be

TABLE 5h.1. Error components used to derive cone of uncertainty

Hour	Cross-track (nm)	Along-track (nm)
0	0	0
12	21	30
24	40	56
36	55	83
48	73	110
72	112	145
96	156	189
120	200	235

a simple assessment: the longer the member tracks remain within the error cone the better behaved the system might be. It's more complicated to consider the cross and along track components as well. We may discover that some TCs are suited to longer track forecasts than others.

6. Summary

This report has summarised the latest configurations of many NWP models used for operational TC track forecasting and included performance statistics and future developments. It has also summarised forecasting techniques and recent performance statistics of many operational TC warning centres. The results presented show the continued reduction in TC track forecast errors by both NWP models and TC warning centres. There has been significant development of ensemble prediction systems and their usage by operational warning centres, although challenges remain as to how to communicate the inherent uncertainty in TC forecasts to the wider public.

7. Acronyms of Meteorological Centres and Numerical Models

Arpège: Météo-France Global Model
CHC: Canadian Hurricane Centre
CMC: Canadian Meteorological Centre
COAMPS: Coupled Ocean/Atmosphere Mesoscale Prediction System (USA)
CPHC: Central Pacific Hurricane Center
ECMWF: European Centre for Medium-Range Weather Forecasts
ENS: ECMWF Ensemble System
GEOS: Goddard Earth Observing System Model
GEPS: Global Ensemble Prediction System (JMA)
GFDL: Geophysical Fluid Dynamics Laboratory
GFS: Global Forecasting System (USA)
GSM: Global Spectral Model (JMA)
HRES: ECMWF High Resolution Model
IMD: India Meteorological Department
JMA: Japan Meteorological Agency
JTWC: Joint Typhoon Warning Center (USA)
KMA: Korea Meteorological Administration
LFM: Local Forecast Model (JMA)
MetUM: Met Office Unified Model (UK)
MOGREPS: Met Office Global and Regional Ensemble Forecasting System (UK)
MSM: Meso-Scale Model (JMA)
NASA: National Aeronautics and Space Administration
NCEP: National Centers for Environmental Prediction (USA)
NCMRWF: National Centre for Medium Range Weather Forecasts (India)
NHC: National Hurricane Center (USA)
NRL: Naval Research Laboratory, Monterey (USA)
PAGASA: The Philippine Atmospheric, Geophysical and Astronomical Services Administration
TEPS: Typhoon Ensemble Prediction System (JMA)
UKMO: United Kingdom Met Office

References

Arakawa, A. and V.R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.,* **17**, 173-265.

BNOC Operations Bulletin No. 112, 2016: "APS2 Upgrade to the ACCESS-TC Numerical Weather Prediction System. http://www.bom.gov.au/australia/charts/bulletins/BNOC_Operations_Bulletin_112.pdf

Bannister, R.N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Q.J.R. Meteorol. Soc*., **134**, 1971-1996.

Biswas, M. K., Bernardet, L. and J. Dudhia, 2014: Sensitivity of hurricane forecasts to cumulus parameterizations in the HWRF model. *Geophys. Res. Lett.*, **41**, 9113-9119.

Bowler, N.E., A.M. Clayton, M. Jardak, P.M. Jermey, A.C. Lorenc, M.A. Wlasak, D.M. Barker, G.W. Inverarity, and R. Swinbank, 2017: The effect of improved ensemble covariances on hybrid variational data assimilation. *Q.J.R. Meteorol. Soc*., **143**, 785–797.

Brousseau, P., Y. Seity, D. Ricard, and J. Léger, 2016: Improvement of the forecast of convective activity from the AROME□ France system. *Q.J.R. Meteorol. Soc.,* **142**, 2231-2243. doi:10.1002/qj.2822

Buehner, M., R. McTaggart-Cowan, A. Beaulne, C. Charette, L. Garand, S. Heilliette, E. Lapalme, S. Laroche, S.R. Macpherson, J. Morneau, and A. Zadra, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part I: the global system. *Mon. Wea. Rev.,* **143**, 2532–2559.

Bush, M. and co-authors, 2019: The Met Office Unified Model/ JULES Regional Atmosphere and Land configurations (RAL) – 1st release. *Submitted to Geosci. Model Dev*.

Cangialosi, J.P., 2018: National Hurricane Center Forecast Verification Report – 2017 Hurricane Season, National Hurricane Center, Miami, FL, 73 pp. http://www.nhc.noaa.gov/verification/pdfs/Verification_2017.pdf.

Cangialosi, J.P., 2018: The State of Hurricane Forecasting. National Hurricane Center Blog – Inside the Eye. https://noaanhc.wordpress.com/2018/03/09/the-state-of-hurricane-forecasting

Chan, J.C.L. and J.D. Kepert, 2010: *Global Perspectives on Tropical Cyclones: From Science to Mitigation, WMO.* https://doi.org/10.1142/7597

Charney, J.G. and N.A. Phillips, 1953: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. *J. Meteor.,* **10**, 71-99.

Chen, S., T. J. Campbell, H. Jin, S. Gaberšek, R. M. Hodur, and P. Martin, 2010: Effect of two-way air–sea coupling in high and low wind speed regimes. *Mon. Wea. Rev*., **138**, 3579–3602.

Clayton, A.M., A.C. Lorenc, and D.M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Q. J. R. Meteorol. Soc*. **139**, 1445–1461, doi: 10.1002/qj.2054.

Cummings, J. A., 2005. Operational multivariate ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131**, 583-604.

Davidson, N. and H. Weber, 2000: The BMRC High-Resolution Tropical Cyclone Prediction System: TC-LAPS, *Mon. Wea. Rev*. **128**, 1245-1265.

Davidson, N.E., Y. Xiao, Y. Ma, H.C. Weber, X. Sun, L.J. Rikus, J.D. Kepert, P.X. Steinle, G.S. Dietachmayer, C.F. Lok, J.

Fraser, J. Fernon, and H. Shaik, 2014: ACCESS-TC: Vortex Specification, 4D-VAR Initialization, Verification and Structure Diagnostics. *Mon. Wea. Rev.,* **142**, 265-1289.

Doyle, J.D., Y. Jin, R. Hodur, S. Chen. Y. Jin. J. Moskaitis, S. Wang, E.A. Hendricks, H. Jin, T.A. Smith, 2014: Tropical cyclone prediction using COAMPS-TC. *Oceanography,* **27**, 92-103.

Doyle, J.D., Y. Jin, R. Hodur, S. Chen. H. Jin, J. Moskaitis, A. Reinecke, P. Black, J. Cummings, E. Hendricks, T. Holt, C. Liou, M. Peng, C. Reynolds, K. Sashegyi, J. Schmidt, S. Wang, 2012: Real time tropical cyclone prediction using COAMPS-TC. *Advances in Geosciences,* **28**, World Scientific Publishing Company, Singapore, 15-28.

Du, J., J. Berner, R. Buizza, M. Charron, P. Houtekamer, D. Hou, I. Jankov, M. Mu, X. Wang, M. Wei and H. Yuan, 2018: Ensemble Methods for Meteorological Predictions. *Handbook of Hydrometeorological Ensemble Forecasting*. Springer, Berlin, Heidelberg.

Dupont, T. and co-authors, 2011: Verification of Ensemble-Based Uncertainty Circles around Tropical Cyclone Track Forecasts. *Weather and Forecasting*, **26**, 664-676.

Fovell, R. G., Corbosiero, K. L. and H. C. Kuo, 2009: Cloud microphysics impact on hurricane track as revealed in idealized experiments. *J. Atmos. Sci.*, **66**, 1764-1778.

Girard, C., A. Plante, M. Desgagné, R. McTaggart-Cowan, J. Côté, M. Charron, S. Gravel, V. Lee, A. Patoine, A. Qaddouri, M. Roch, L. Spacek, M. Tanguay, P.A. Vaillancourt, and A. Zadra, 2014: Staggered vertical discretization of the Canadian Environmental Multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type. *Mon. Wea. Rev.,* **142**, 1183–1196.

Gopalakrishnan, S.G. and co-authors, 2018: 2017 HFIP R&D Activities Summary: Recent Results and Operational Implementation, HFIP Technical Report HFIP2018-1, Available online at http://www.hfip.org/documents/HFIP_AnnualReport_FY2017.pdf

Harris, L. M., S.-J. Lin and C. Tu, 2016: High-resolution climate simulations using GFDL HiRAM with a stretched global grid. *J. Clim*., **29**, 4293–4314, doi:10.1175/JCLI-D-15-0389.1

Heming, J.T., 2016: Met Office Unified Model Tropical Cyclone Performance Following Major Changes to the Initialization Scheme and a Model Upgrade. *Wea. Forecasting*, **31**, 1433–1449.

Heming, J.T., 2017: Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. *Meteorological Applications*, **24**, no.1, 1-8.

Heming, J.T. and J.S. Goerss, 2010: Track and Structure Forecasts of Tropical Cyclones. *Global Perspectives on Tropical Cyclones: From Science to Mitigation, Chapter 10, WMO.* https://doi.org/10.1142/7597

Hersbach, H. and co-authors, 2018: Operational global reanalysis: progress, future directions and synergies with NWP. *ERA Report Series no.27*. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom. https://www.ecmwf.int/en/elibrary/18765-operational-global-reanalysis-progress-future-directions-and-synergies-nwp

Hogan, T. F., M. Liu, J. A. Ridout, M. S. Peng, T. R. Whitcomb, B. C. Ruston, C. A. Reynolds, S. D. Eckermann, J. R. Moskaitis, N. L. Baker, J. P. McCormack, K. C. Viner, J. G. McLay, M. K. Flatau, L. Xu, C. Chen., and S. W. Chang, 2014: The Navy Global Environmental Model. *Oceanography*, **27**, 116-125.

Houtekamer, P.L., X. Deng, H.L. Mitchell, S. Baek, and N. Gagnon, 2014: Higher Resolution in an Operational Ensemble Kalman Filter. *Mon. Wea. Rev.,* **142**, 1143–1162.

IMD, 2013: Cyclone Warning Services: Standard Operation Procedure. *Cyclone Warning Division, IMD, New Delhi*.

Kotal S.D. and S.K. Roy Bhowmik, 2011: A multimodel ensemble technique for cyclone track prediction over north Indian Sea. *Geofizika*, **28**, 275-291.

Landsea, C.W and J.P. Cangialosi, 2018: Have We Reached the Limits of Predictability for Tropical Cyclone Track Forecasting? *Bulletin of the American Meteorological Society*, in press.

Lin, S.-J., 1997: A finite-volume integration method for computing pressure gradient force in general vertical coordinates, *Q. J. Roy. Meteorol. Soc*., **123**, 1749-1762.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models, *Mon.Wea. Rev.,* **132**, 2293-2307.

Lin, S.-J., and R.B. Rood, 1997: An explicit flux-form semi-lagrangian shallow-water model on the sphere. *Q. J. Roy. Meteorol. Soc*., **123**, 2477-2498.

Li, X. and Z. Pu, 2009: Sensitivity of numerical simulations of the early rapid intensification of Hurricane Emily to cumulus parameterization schemes in different model horizontal resolutions. *J. Meteor. Soc. Jpn.*, **87**, 403–421.

Marchok, T. P., 2002: How the NCEP Tropical Cyclone Tracker works. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P1.13.

Martin, P. J., 2000: Description of the NAVY coastal ocean model version 1.0. NRL Rep. NRL/FR/7322-00-9962, 42 pp.

Martin, P. J., J. W. Book and J. D. Doyle, 2006: Simulation of the northern Adriatic circulation during winter 2003. *J. Geophys. Res.*, **111**, C03S12. doi:10.1029/2006JC003511.

Mogensen, K.S., L. Magnusson and J-R. Bidlot, 2017: Tropical cyclone sensitivity to ocean coupling in the ECMWF coupled model. *J. Geophys. Res. Oceans*, **122**, 4392-4412.

Mohapatra, M., D.P. Nayak and B.K. Bandyopadhyay, 2012: Evaluation of Cone of Uncertainty in Tropical Cyclone Track Forecast over north Indian Ocean issued by India Meteorological Department. *Tropical Cyclone Research and Review*, **2**, 331-339.

Mohapatra, M., D.P. Nayak, R.P. Sharma and B.K. Bandyopadhyay, 2013a: Evaluation of official tropical cyclone track forecast over north Indian Ocean issued by India Meteorological Department. *J. Earth Syst. Sci.,* **122**, 589–601.

Mohapatra M, D.R. Sikka, B.K. Bandyopadhyay and A. Tyagi, 2013b: Outcomes and Challenges of Forecast Demonstration Project (FDP) on Landfalling Cyclones over the Bay of Bengal. *Mausam,* **64,** 1–12.

Mohapatra, M., B. Geetha and M. Sharma, 2017: Reduction in uncertainty in tropical cyclone track forecasts over the North Indian Ocean. *Current Science*, **112**, 1826-1830.

Mohapatra M. and B.K. Bandyopadhyay, 2012: Characteristics of Sudden Changes in Tropical Cyclone Tracks over North Indian, Ocean. http://www.wmo.int/pages/prog/arep/wwrp/pdf%20files/6%20Nov/MMohapatra.pdf.

Moorthi, S., H.-L. Pan, P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. *NWS Technical Procedures Bulletin*, **484**, pp.14. Available at http://www.nws.noaa.gov/om/tpb/484.htm.

Nasrollahi, N., AghaKouchak, A., Li, J., Gao, X., Hsu, K. and S. Sorooshian, 2012: Assessing the impacts of different WRF precipitation physics in hurricane simulations. *Wea. Forecasting*, **27**, 1003-1016.

Nishimura, M. and M. Yamaguchi, 2015: Selective ensemble mean technique for tropical cyclone track forecasts using multi-model ensembles. *Tropical Cyclone Research and Review*, **4**, 2, 71-78.

Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *J. Comput. Phys*., **227**, 55-78, doi:10.1016/j.jcp.2007.07.022.

Qaddouri, A. and V. Lee, 2011: The Canadian Global Environmental Multiscale model on the Yin-Yang grid system. *Quart. J. Roy. Meteor. Soc.*, **137**, 1913-1926.

RSMC, New Delhi, 2018: Report on cyclonic disturbances over the north Indian Ocean during 2017. *RSMC, IMD, New Delhi.*

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (Version 3.2). *Bull. Amer. Meteor. Soc,* **81**, 1131–1240.

Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. 2011: The AROME-France Convective-Scale Operational Model. *Mon. Wea. Rev.,* **139**, 976–991. https://doi.org/10.1175/2010MWR3425.1

Shepherd, T. J. and K. J. Walsh, 2017: Sensitivity of hurricane track to cumulus parameterization schemes in the WRF model for three intense tropical cyclones: impact of convective asymmetry. *Met. Atmos. Phys.*, **129**, 345-374.

Short, C. J. and J. Petch, 2018: How well can the Met Office Unified Model forecast tropical cyclones in the Western North Pacific? *Wea. Forecasting*, **33**, 185-201.

Smith, G.C., J. Bélanger, F. Roy, P. Pellerin, H. Ritchie, K. Onu, M. Roch, A. Zadra, D.S. Colan, B. Winter, J. Fontecilla, and D. Deacu, 2018: Impact of Coupling with an Ice–Ocean Model on Global Medium-Range NWP Forecast Skill. *Mon. Wea. Rev.,* **146**, 1157–1180.

Suselj, K., T. F. Hogan, and J. Teixeira, 2014: Implementation of a stochastic eddy-diffusivity mass flux parameterization into the Navy Global Environmental Model. *Wea. Forecasting*, **29**, 1374–1390. doi:https://doi.org/10.1175/WAF-D-14-00043.1.

Termonia, P. and co-authors, 2018: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, *Geosci. Model Dev*., **11**, 257-281. https://doi.org/10.5194/gmd-11-257-2018.

Titley, H.A. and R. Stretton, 2018: A Probabilistic Evaluation of Global Tropical Cyclone Forecasts from the Upgraded Met Office MOGREPS-G Ensemble, and the Value of Multi-Model Ensembles. *33rd Conference on Hurricanes and Tropical Meteorology, Florida. Extended abstract*.

Tsuyuki, T., R. Sakai, and H. Mino, 2002: The WGNE intercomparison of typhoon track forecasts from operational global models for 1991-2000. *WMO Bulletin*, **51**, No.3, 253-257.

Walters, D. and co-authors, 2017: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geosci. Model. Dev.,* **10**, 1487-1520.

Williford, C.E., T. N. Krishnamurti, R. C. Torres, S. Cocke, Z. Christidis, and T. S. V. Kumar, 2003: Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999. *Mon. Wea. Rev*., **131**, 1878-1894.

Wood, N., A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat, and J. Thuburn, 2014: 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.

Yamaguchi, M., J. Ishida, H. Sato and M. Nakagawa, 2017: WGNE Intercomparison of Tropical Cyclone Forecasts by Operational NWP Models: A Quarter Century and Beyond. *Bull. Amer. Meteor. Soc.*, **98**, 2337–2349.

Zadra, A. and Coauthors, 2014a: Improvements to the Global Deterministic Prediction system (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). Canadian Meteorological Centre Tech. Note, **88**.
http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.

Zadra, A., R. McTaggart-Cowan, P.A. Vaillancourt, M. Roch, S. Bélair, and A. Leduc, 2014b: Evaluation of Tropical Cyclones in the Canadian Global Modeling System: Sensitivity to Moist Process Parameterization. *Mon. Wea. Rev.,* **142**, 1197–1220.