Tropical Cyclone Data Impact Studies: Influence of Model Bias and Synthetic Observations

CAROLYN A. REYNOLDS, ROLF LANGLAND, AND PATRICIA M. PAULEY

Marine Meteorology Division, Naval Research Laboratory, Monterey, California

CHRISTOPHER VELDEN

CIMSS, University of Wisconsin–Madison, Madison, Wisconsin

(Manuscript received 28 September 2012, in final form 21 May 2013)

ABSTRACT

The impacts of assimilating dropwindsonde data and enhanced atmospheric motion vectors (AMVs) on tropical cyclone track forecasts are examined using the Navy global data assimilation and forecasting systems. Enhanced AMVs have the largest impact on eastern Pacific storms, while dropwindsonde data have the largest impact on Atlantic storms. Results in the western Pacific are mixed. Two western Pacific storms, Nuri and Jangmi, are examined in detail. For Nuri, dropwindsonde data and enhanced AMVs are at least as likely to degrade as to improve forecasts. For Jangmi, additional data improve track forecasts in most cases. An erroneous weakening of the forecasted subtropical high appears to contribute to the track errors for Nuri and Jangmi. Assimilation of enhanced AMVs systematically increases the analyzed heights in this region, counteracting this model bias. However, the impact of enhanced AMVs decreases rapidly as the model biases saturate at similar levels for experiments with and without the enhanced AMVs after the first few forecast days. Experiments are also conducted in which the errors assigned to synthetic tropical cyclone observations are increased. Moderate increases in the assigned errors improve track forecasts on average, but larger increases in the assigned errors produce mixed results. Both experiments allow for reductions in innovations and residuals when compared to dropwindsonde observations. These experiments suggest that a reformulation of the synthetic tropical cyclone observation scheme may lead to improved forecasts as more in situ and remote observations become available.

1. Introduction

Given the enormous potential impact, improving tropical cyclone (TC) track forecasts is of great interest around the world. The World Meteorological Organization (WMO)-sponsored program The Observing System Research and Predictability Experiment (THORPEX; see http://www.wmo.int/pages/prog/arep/wwrp/new/thorpex_new.html) Pacific Asian Regional Campaign (T-PARC; see http://www.ucar.edu/na-thorpex/tparc/SPO_PARC_revised.pdf) and the Office of Naval Research (ONR) Tropical Cyclone Structure-08 (TCS-08) field campaign were conducted from August through early October 2008 (Elsberry and Harr 2008) with the goal of improving TC intensity, structure, and track forecasts. Toward this end, five aircraft stationed in Guam, Japan, and Taiwan flew a total of 76 missions during which almost 1500 vertical soundings of atmospheric data (dropwindsondes; Hock and Franklin 1999) were taken. One of the objectives of T-PARC/TCS-08 was TC targeted observations, where additional observations were taken in sensitive regions deemed most likely to have an impact on TC forecasts [see the Langland (2005) review article for more information on targeted observations, and Reynolds et al. (2010) for information on targeted observing products produced by the Navy for T-PARC/TCS-08]. Synoptic-scale surveillance flights in sensitive regions have also been conducted for several years under ongoing programs for Atlantic storms (Aberson 2003, 2010) and western North Pacific TCs threatening Taiwan (Wu et al. 2005, 2007).

Several studies have already been published that examine the impact of the dropwindsonde data and other T-PARC observations on forecast skill in different data
assimilation/forecast systems. Weissmann et al. (2011) compared the impact of dropwindsonde data on TC track forecasts in the European Centre for Medium-Range Weather Forecasts (ECMWF), Japan Meteorological Agency (JMA), National Centers for Environmental Prediction (NCEP), and Weather Research and Forecasting Model (WRF) modeling systems. They found larger beneficial impacts on TC track forecasts in the NCEP and WRF systems than in the ECMWF and JMA systems. They attributed this difference in impact to the fact that the NCEP and WRF systems had larger errors to begin with, and less advanced data assimilation systems [e.g., three-dimensional variational data assimilation (3DVAR)] compared with the four-dimensional (4DVAR) systems of ECMWF and JMA. Similar results were found by Chou et al. (2011), with larger impacts obtained from T-PARC and Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) dropwindsonde data in the NCEP model than in the ECMWF model. Harnisch and Weissmann (2010) found that the impacts on 12–120-h track forecasts using the ECMWF system were largest from dropwindsondes that were taken in the vicinity of the storm (placed in a ring around the outer boundary of the storm), whereas remote observations (700–1200 km from the storm) by themselves had relatively small impact. They also found that observations in the core of the TC sometimes led to large differences in the analysis, but only small forecast improvements. They attributed this to the fact that current model resolutions are insufficient to effectively use the TC core data, an issue also pointed out by Aberson (2008).

Studies have also looked at the impact of lidar observations on forecasts. Harnisch et al. (2011) found that, in the ECMWF system, differential absorption lidar water vapor observations were able to substantially affect some forecasts under certain conditions, but had a small impact on average. Weissmann et al. (2012) examined the impact of airborne Doppler wind lidar profiles in the ECMWF and Navy Operational Global Atmospheric Prediction System (NOGAPS) forecasts of the T-PARC Typhoon Sinlaku. The beneficial impact of the lidar profiles was larger in the Navy system than in the ECMWF system. In the Navy system, the impact per observation of airborne lidar observations was higher than the impact of other wind observations, whereas in the ECMWF system the impact was similar to other aircraft observations. The impact on NOGAPS Sinlaku track forecasts was relatively small. The synthetic TC vortex observation data in NOGAPS substantially reduced the influence of additional observations. Removal of the synthetic observations improved the track forecast but also weakened the storm.

Because of the statistical nature of the data assimilation problem, including observation errors, errors of representativeness, and other approximations in the data assimilation system, one would not expect targeted data to improve the forecast in every case (Langland 2005). The specific impact of targeted observations will depend on how well the targeted observations cover the sensitive regions. Aberson (2008) pointed out that the assimilation of synoptic surveillance TC dropwindsonde data, while improving track forecasts on average, led to significant degradations in some cases, due either to erroneous data being assimilated or to issues with the quality control or the data assimilation system itself. Aberson also noted that because of uncertainties in the data, model, and assimilation system, some surveillance data will lead to forecast degradations. Aberson et al. (2011) found that when dropwindsonde data were obtained outside of targeted regions or did not fully sample targeted regions, the impact on the targeted storm was usually small. Their results suggest that the spread of the impact of the targeted observations into relatively data-sparse regions may result in forecast degradation, and this effect may be more pronounced if model bias is significant. Aberson (2011) examined the impact of both Pacific and Atlantic dropwindsonde data on all TCs from 27 August through 1 October 2008 in the National Oceanic and Atmospheric Administration (NOAA) Global Forecast System (GFS) and found average percent reductions in forecast errors of up to 12% in the Atlantic and 23% in the western North Pacific. Aberson also noted that storms not targeted with special observations could be impacted by dropwindsondes in other systems, as the influence of the dropwindsonde data propagates globally through the update cycle, and due to the spectral nature of the system.

Berger et al. (2011) examined the impact of enhanced satellite-derived atmospheric motion vectors (AMVs), as well as dropwindsonde data, on T-PARC TC forecasts in the Navy forecast system. [For a description of satellite products made available during T-PARC/TCS-08, please see Hawkins and Velden (2011).] Berger et al. (2011) found that the addition of Cooperative Institute for Meteorological Satellite Studies (CIMSS) hourly AMVs reduced western North Pacific TC track forecast errors on average, particularly for forecasts longer than 72 h (through improved large-scale environmental flow analyses). Adding rapid-scan AMVs further reduced forecast track errors. Adding dropwindsonde data to CIMSS hourly winds resulted in a slight degradation to the track forecasts, although with large case-to-case variability.

Forecast track errors are caused by model error as well as errors in the initial conditions. The TC itself may be
poorly represented due to limited resolution and errors in the physical parameterizations. Model errors may lead to biases in the large-scale features that steer the cyclone, such as the subtropical high. The inadequate resolution of the TC in the model will also impact the accuracy of the initial conditions, which may also be influenced by synthetic observations, as discussed below. The influence of model error is further complicated by the fact that errors in the representation of the cyclone structure itself may influence the local environment and steering flow (e.g., Wu and Emanuel 1995a,b; Fovell et al. 2009). Improved physical parameterizations have been shown to decrease forecast biases and improve TC track forecasts in the NOAA GFS (Han and Pan 2011), the Advanced Hurricane Weather Research and Forecasting Model (AHW; Torn and Davis 2012), the Hurricane Weather Research and Forecasting System (HWRF; Gopalakrishnan et al. 2012), and the ECMWF forecast system (Fiorino 2009). Features such as subtropical anticyclones have been shown to influence TC motion in composite studies (e.g., Chen et al. 2009) and case studies (e.g., Brennan and Majumdar 2011). The identification of systematic track errors (e.g., Elsberry 1995; Carr and Elsberry 2000) has led to efforts to develop techniques to assist forecasters in identification and investigation of these error mechanisms (Carr et al. 2001). If forecast errors are dominated by model error, then potential improvement from more accurate initial conditions will be limited. Model biases may also lead to biases in the analyses. Vukicevic et al. (2013) examine how the systematic spindown of vortex core circulations in short-term HWRF forecasts had a detrimental impact on subsequent analyses and forecasts.

In this paper we present the results of several data impact studies using the Navy operational global data assimilation and forecast system from the August through September 2008 T-PARC/TCS-08 time period. The current study is largely motivated by the ambiguous results concerning the impact of additional data in the western Pacific in the Berger et al. (2011) study as well as previously cited studies. We go beyond the Berger et al. (2011) study by examining the impact of enhanced AMVs and dropwindsonde data separately, examining the impact in all basins, and by exploring the data impact by basin. We evaluate whether model biases may be lessening the potential impact of additional data through evaluation of systematic analysis differences and forecast errors, and through consideration of two case studies. It was suggested in Aberson (2010), Berger et al. (2011), and Weissmann et al. (2012) that the assimilation of synthetic TC observations may be interfering with the potential impact of additional data in the vicinity of the storm. To examine this point further, two update cycle/forecast experiments are performed in which errors assigned to the TC synthetic observations are increased. In the first case, the assigned errors are increased such that the synthetics are considered less accurate than radiosondes or dropwindsondes but more accurate than AMVs. This allows for the analysis to draw closer to the dropwindsonde observations than the synthetic observations. In the second case, the assigned errors are increased further, such that the synthetic observations are considered less accurate than both dropwindsondes and AMVs. The experimental design is described in section 2. The results are presented in section 3. A brief summary and conclusions are presented in section 4.

2. Experimental design

a. System description

The data assimilation and forecasting system used is the system that has produced the operational global atmospheric forecasts for the Navy. The global data assimilation and forecasting system consists of the NOGAPS (Hogan and Rosmond 1991; Peng et al. 2004) forecast model and the Naval Research Laboratory (NRL) Atmospheric Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR; Xu et al. 2005; Rosmond and Xu 2006; Chua et al. 2009). NAVDAS-AR was transitioned to operations in September 2009, so while the operational T-PARC forecasts ran from 3DVAR analyses, these data impact studies were performed using the more advanced 4DVAR NAVDAS-AR system. The update cycle and forecast system are run at a resolution of 239 spectral triangular truncation (approximately 55 km) and 42 levels (T239L42). The update cycle length is 6 h, and long (120 h) forecasts are run at 0000 and 1200 UTC through August and September 2008.

b. Experiments

A series of six update cycle/forecast experiments are run to examine the impact of CIMSS enhanced atmospheric motion vectors (Velden et al. 2005), dropwindsonde data, and increased observation errors assigned to TC synthetic observations (Table 1). NAVDAS-AR assimilates satellite-derived AMVs produced from routine geostationary satellite image triplets at 30-min image frequency. For the T-PARC period, enhanced CIMSS AMVs were made available every hour rather than every 6 h (see Berger et al. 2011 for more details),

The Navy Global Environmental Model replaced NOGAPS in 2013.
resulting in a much larger number of AMV super-observations ingested during the 6-h assimilation window. Berger et al. (2011) also considered the impact of CIMSS rapid-scan AMVs derived from 15-min Multi-functional Transport Satellite 2 (MTSAT-2) images provided by JMA during a few high-impact periods. The impact of rapid-scan winds is not investigated here.

In the first experiment, NO_ADD, neither dropwindsonde data nor enhanced CIMSS hourly AMVs are assimilated. In CAMV, CIMSS hourly AMVs are assimilated, but dropwindsonde data are not. In DROP, dropwindsonde data (both Atlantic and Pacific) are assimilated, but enhanced AMVs are not. In CAMV_DROP, both enhanced AMVs and dropwindsonde data are assimilated. NO_ADD, CAMV, and CAMV_DROP experiments were reported on in Berger et al. (2011) for western Pacific storms only. In this study we examine the impact of dropwindsonde data and enhanced AMVs separately as well as combined, for all basins, and we take an in-depth look at case studies and model biases to aid interpretation and understanding of ambiguous results. Note that forecasts for all experiments are produced every 0000 and 1200 UTC regardless of whether or not, for example, dropwindsonde data were collected during that specific update cycle.

As has been suggested in previous work, the impact of the dropwindsonde data may be reduced by the TC synthetic observations, which are assigned small observation errors. The Navy system assimilates synthetic wind and height observations up to 400 hPa within about 600 n mi of the storm center, as described in Goerss and Jeffries (1994). The system consists of synthetic observations representing a synthetic Rankine vortex in gradient balance. The structure of the Rankine vortex is determined by the reported maximum TC wind speed. TC synthetic observations currently are assumed to be even more accurate than radiosonde and dropwindsonde observations (i.e., the assigned synthetic errors are 0.8 times the assigned radiosonde errors). Because of this, dropwindsonde data and other observations will “compete” with synthetic observations, and this may not always be beneficial. This is especially true in cases where asymmetries in the actual TC are pronounced and not captured by the symmetric Rankine vortex synthetic observations. To examine the impact of additional data when the weight given to the synthetic observations is reduced, two additional experiments (SYNTH1 and SYNTH2) are performed.

The values of the assigned zonal and meridional wind observation error standard deviations are provided in Table 2 for dropwindsonde data, synthetic observations, and AMVs, at different pressure levels. Dropwindsonde data are assigned the same error values as radiosonde observations and range from 1.8 to 2.6 m s\(^{-1}\), slightly larger than the 0.5–2.0 m s\(^{-1}\) range estimated in Hock and Franklin (1999). In SYNTH1, the error assigned to the synthetics is increased by 75% (1.4 times the radiosonde errors). Therefore, synthetics in SYNTH1 are considered less accurate than dropwindsonde data, but more accurate than AMVs. This allows the analysis to draw more closely to dropwindsonde data than synthetic observations in the vicinity of the TC. In SYNTH2, the errors assigned to the synthetics are increased by 200% (2.4 times the radiosonde errors). Therefore synthetics are considered less accurate than both dropwindsonde data and AMVs. In SYNTH1 and SYNTH2, both

| Table 1. Description of update cycle and forecast experiments. |
|-------------------|-----------------|-----------------|-----------------|
| Name              | CIMSS hourly AMVs | Dropwindsondes | Assigned synthetic observation error |
| NO_ADD            | Not assimilated  | Not assimilated | Control value   |
| CAMV              | Assimilated      | Not assimilated | Control value   |
| DROP              | Not assimilated  | Assimilated     | Control value   |
| CAMV_DROP         | Assimilated      | Assimilated     | Control value   |
| SYNTH1            | Assimilated      | Assimilated     | Increased by 75%|
| SYNTH2            | Assimilated      | Assimilated     | Increased by 200%|

| Table 2. Zonal and meridional wind observation assigned error standard deviation (m s\(^{-1}\)) for different types of observations, for selected pressure levels. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Pressure level (hPa) | Dropwindsondes/radiosondes | Synthetic observations | AMVs |
|                 |                 | Control | SYNTH1 | SYNTH2 |     |
| 1000            | 1.8             | 1.44    | 2.52   | 4.32   | 2.8 |
| 700             | 1.9             | 1.52    | 2.66   | 4.56   | 4.4 |
| 500             | 2.1             | 1.68    | 2.94   | 5.04   | 4.4 |
| 400             | 2.6             | 2.0     | 3.5    | 6.0    | 5.2 |
Dropwindsonde data and enhanced AMVs are assimilated, such that they are identical to CAMV_DROP aside from the assigned synthetic observation errors. Note that only wind observations are assimilated from the synthetics; wind, temperature, and humidity observations are assimilated from the dropwindsondes. An additional set of experiments is performed in which the synthetic observations are completely removed. These produced poor TC forecasts and are reported on briefly in section 3c.

Because current analyses are dependent on prior forecasts, there will be differences between the analyses from DROP and NO_ADD, and between CAMV and CAMV_DROP, even at times when no dropwindsonde data are assimilated. In this way the impact of the particular dataset of interest is accumulated through the two-month time period.

3. Results

a. Impact of enhanced AMVs and dropwindsonde data

Figure 1 shows the average track error in nautical miles for the first four experiments averaged for all TCs (Atlantic and Pacific basins) for 2008. The addition of hourly CIMSS AMVs (CAMV) reduces the track error for the later forecast times. The addition of dropwindsonde data (DROP) without AMVs also results in forecast error reduction at later times, although they have a smaller impact than enhanced AMVs. If both dropwindsonde data and CIMSS AMVs are assimilated (CAMV_DROP), the results are worse than the assimilation of CIMSS AMVs alone (CAMV), as was also shown in Berger et al. (2011) for the western North Pacific basin. Track errors and statistical significance (two-tailed t test) are computed using the Automated Tropical Cyclone Forecast (ATCF) system (Sampson and Schrader 2000). The improvements for CAMV over NO_ADD are statistically significant at the 95% level from 72 to 120 h.

The impact of additional observations varies by basin. The top panel of Fig. 2 shows the average TC track error for CAMV minus the average track error for NO_ADD, indicating the impact of enhanced AMVs, grouped according to basin (Atlantic, western North Pacific, and eastern North Pacific). The bottom panel of Fig. 2 shows the analogous impact of the dropwindsonde data (i.e., the average TC track error for DROP minus the average TC track error for NO_ADD). The enhanced AMVs produce the biggest reduction in track error for all lead times out to 108 h in the eastern North Pacific, perhaps not surprising given the large in situ data void in the vicinity of the TCs. The impact of enhanced AMVs in the western North Pacific is nearly neutral out to 72 h.
The impact of dropwindsonde data is shown in the bottom panel of Fig. 2. In the Atlantic, the dropwindsonde data reduce the error more than the enhanced AMVs between 24 and 108 h, with statistically significant reductions at the 95% level at 48 h and at the 90% level at 60 h. In the Pacific, the signal is mixed, with reductions in track errors of 8 n mi or greater at 60–84 h and at 108 and 120 h (statistically significant at the 90% level at 72 and 120 h), but has near-neutral impacts at other forecast times. As a percent difference, the maximum reduction due to dropwindsonde data is 11% in the Atlantic and 8% in the western North Pacific. These numbers are smaller than the Atlantic and western North Pacific reductions of 12% and 23% found for the NOAA GFS by Aberson (2011) for a somewhat shorter time period (27 August–1 October 2008). Although there are no dropwindsondes during this period in eastern North Pacific storms, there is still a small impact on these storms from the remote influence of the dropwindsonde data in other basins through the data assimilation update cycle. Aberson (2011) also observes a global impact from dropwindsonde data due to the propagation of the dropwindsonde signal through the update cycle as well as the spectral nature of the NOAA GFS. In summary, in the Navy system, the enhanced AMVs have the largest impact overall in the eastern North Pacific, with dropwindsonde data giving the most uniform improvement in the Atlantic.

Berger et al. (2011) note that there is significant case-to-case variability in data impact in the western North Pacific T-PARC storms. To examine this in more detail, we consider two storms as case studies. The first storm, Nuri, formed east of the Philippines on 17 August and traveled west northwestward crossing the northern tip of Luzon, eventually making landfall near Hong Kong on 22 August. At peak intensity Nuri had 185 km h$^{-1}$ (100 kt) sustained winds and a central pressure of 955 hPa. Figure 3 shows the observed track (black) along with the NOGAPS forecast tracks (gray) every 12 h, for the (top) NO_ADD and (bottom) CAMV_DROP forecasts. While there are small differences in the details, forecasts from both experiments show very persistent erroneous recurvature. The average track errors (shown in Fig. 5 and discussed below) confirm the lack of track forecast improvement from additional data. Tracks from the other experiments, not shown, all indicate this persistent recurvature error. NOGAPS was not the only model forecasting erroneous recurvature. Forecasts from the global models from NCEP and the Met Office also showed recurvature errors, particularly for forecasts initialized close to genesis (in contrast, ECMWF correctly indicated landfall near Hong Kong). This was reflected in the forecasts produced by the Joint Typhoon Warning Center.
Center (JTWC), which considers model forecasts from several different operational centers. JTWC predicted Nuri recurving east of Taiwan in the forecasts starting prior to 1200 UTC 18 August and had right-of-track biases through much of Nuri’s life (JTWC forecasts are available on the NRL TC web page at www.nrlmry.navy.mil/TC.html).

The second storm considered is Typhoon Jangmi (Fig. 4), which formed east of the Philippines on 24 September, traveled northwestward to make landfall in Taiwan on 28 September and then recurved into the Pacific south of Japan, eventually dissipating in early October. The JTWC revised peak intensity estimates indicate that Jangmi was a category 5 supertyphoon reaching sustained winds of 270 km h$^{-1}$ (145 kt) early on 27 September. In contrast to Nuri, most of the early Jangmi forecasts (before 27 August) for NO_ADD (Fig. 4, top) show a left of track bias and fail to capture the observed recurvature. Enhanced AMVs and drop-windsonde data mitigate the leftward bias to some degree (Fig. 4, bottom), but most forecasts initialized early in the storm’s life cycle still fail to capture the observed recurvature. The average track errors, shown in Fig. 5 and discussed below, reflect some reduction of
the average track error from the additional data. Examination of the CAMV and DROP forecasts (not shown) indicate that most of the improvement is due to the enhanced satellite winds. The persistent left of track bias for Jangmi was also noted in the low-resolution (T119) NOGAPS ensembles (Snyder et al. 2011), although an ensemble that included stochastic convection had a larger number of recurving members than an ensemble based on initial perturbations only.

As with Nuri, several operational centers had qualitatively similar error characteristics. GFS and Met Office forecasts had left-of-track biases and landfall in continental Asia for the forecasts initialized early in the storm’s life cycle (ECMWF forecasts also had

FIG. 5. The average track errors (n mi) for (top) TC Nuri and (bottom) TC Jangmi for the six experiments as denoted in key. The number of cases is shown below the x axis.
left-of-track biases for short forecast times, but captured the recurrature. JTWC predicted Jangmi to make landfall in continental Asia for forecasts started prior to 0600 UTC 27 September. It was not until the forecast from 1200 27 September that JTWC predicted recurrature east of the Asian mainland.

The TC track errors from the 0000 and 1200 UTC forecasts averaged for each storm are shown in Fig. 5. The track errors for Nuri (top panel) beyond 60 h are actually smallest for NO_ADD and DROP on average. The highest forecast errors occurred for the forecast in which the most data were included (CAMV_DROP), counter to what would be expected. (The errors for SYNTH1 and SYNTH2 will be discussed in section 3c.)

For Jangmi (Fig. 5, bottom panel), the results are more intuitive. The largest track errors occur in the NO_ADD case, with small improvements from DROP and larger improvements from CAMV, and CAMV_DROP is better still. Please note that the sample sizes are very small and the differences are not statically significant.

With such a small sample size, frequency of superior performance statistics are helpful in determining if the average differences are due to just a few cases or are more uniform throughout the sample. Figure 6 shows frequency of superior performance for several pairs of experiments (when the difference in performance is greater than 5%) for Nuri (top panel) and Jangmi (bottom panel). For Nuri, the additional data have a mixed impact on track error. Inclusion of just the enhanced AMVs results in improvements (degradations) of 5% or greater in 26% (61%) of the cases. Inclusion of just the dropwindsonde data results in improvements (degradations) in 39% (48%) of the cases. Inclusion of both the enhanced AMVs and the dropwindsonde data leads to track improvements (degradations) in 21% (74%) of the cases. Please note that these cases are not independent, as forecast errors every 24 h from every forecast are included, and the forecasts are started only 12 h apart.

For Jangmi (Fig. 6, bottom panel), the additional data are clearly more frequently helpful than detrimental. Inclusion of just the enhanced AMVs improves (degrades) the forecast by more than 5% in 53% (32%) of the cases. Inclusion of just the dropwindsonde data improves (degrades) the forecast in 53% (16%) of the cases. Inclusion of both the enhanced AMVs and the dropwindsonde data results in improvements (degradations) in 64% (24%) of the cases.

The results shown so far indicate that there is significant storm-to-storm and case-to-case variability in data impact. The results for Jangmi are intuitive, as additional observations lead to forecast track improvements more often than degradations. Nuri is the counter example, where there is considerable case to case variability, but on average the additional observations lead to degradations slightly more often than improvements.

As noted in the introduction, there are several reasons as to why assimilation of additional observations may not necessarily lead to improved analyses and/or forecasts, including poor sampling of the sensitive regions, approximations made in the data assimilation system, and observation error. Velden and Bedka (2009) note, for example, that height assignment error contributed substantially to AMV uncertainty. Of course, it is not only errors in the analysis that may lead to forecast error, but model errors as well. We examine the potential impact of model error, specifically model biases in the North Pacific subtropical high, on specific Nuri and Jangmi forecasts in the next subsection.

b. Model biases

We present analysis and forecast fields from two specific forecasts as examples of how model bias may be impacting forecast errors and limiting the potential positive impact from improvements to the initial conditions. The first forecast is for Nuri, starting from 0000 UTC 18 August. Figure 7 shows the CAMV_DROP verifying analyses and 48-h forecast valid on 0000 UTC 20 August (top panels) and the verifying analysis and 120-h forecast valid on 0000 UTC 25 August (bottom panels). The 500-hPa height fields are shaded. The 850-hPa vorticity fields (positive values only), indicating the TC position, are contoured in black for CAMV_DROP and in red for NO_ADD. In the verifying analysis for 20 August (upper left panel), the center of Nuri is over the northern tip of the Philippines, with the subtropical high to the north and east of the storm. In the 48-h CAMV_DROP forecast (upper right panel), the storm has not progressed far enough westward, with the center east of the Philippines (black contours). The weakening of the subtropical high in the forecast relative to the analysis is consistent with an erroneously weak westward steering flow, resulting in a slow track bias. By 23 August, Nuri has made landfall in China (lower left panel) in the analysis but is starting to recurve around the subtropical high, ahead of a midlatitude trough, in the CAMV_DROP forecast (lower right panel).

The erroneous weakening of the subtropical high is very consistent in all the forecasts examined. To illustrate this, the 5890- and 5905-m height contours from CAMV_DROP are denoted by the solid blue curves, and the NO_ADD analysis and forecast are denoted by the dashed blue curves, in the top panels. The 850-hPa vorticity from the NO_ADD forecast is denoted by the red contours. The NO_ADD and CAMV_DROP
subtropical highs are very similar to each other in the analyses and 48-h forecasts (slightly weaker in the NO_ADD analysis east of 135°E). These similarities indicate that the low height forecast bias is not substantially mitigated by the additional observations.

Indeed, the position of the TC in the 48-h NO_ADD forecast (red contours) is actually slightly closer to the observed position. Both the NO_ADD and the CAMV_DROP forecasts recurve to near Taiwan by 120h (lower right panel). This is also true in CAMV, DROP, and

![Graph showing percentage of cases with better or worse performance for Nuri and Jangmi](image-url)
SYNTH forecasts (not shown). The robustness of the erroneous systematic recurvature of the Nuri forecasts appears to be related to the forecast bias of a weakening subtropical anticyclone, which is not significantly modified by the additional data.

The forecasts errors for Jangmi stand in contrast to those for Nuri. With Nuri, the observed storm moves westward, making landfall in China, while the forecasts have erroneous recurvature. With Jangmi, the observed storm recurves to hit Taiwan, while the forecasts initialized close to genesis miss the recurvature and predict landfall near Hong Kong. Despite these differences, examination of synoptic fields suggests that Jangmi forecasts are also influenced by the model bias of a weak subtropical anticyclone. Figure 8 shows the same fields as in Fig. 7, but for forecasts starting from 25 September, with fields for the verifying analysis and 48-h forecasts valid on 27 September shown in the top panels, and verifying analysis and 120-h forecasts valid on 30 September in the bottom panels. As with Nuri, the subtropical anticyclone in the western North Pacific is weaker in the Jangmi forecasts than in the verifying analyses. However, because the storm is to the west of the geopotential height maximum associated with the subtropical anticyclone for Jangmi (as opposed to south of it for Nuri), this results in a weaker-than-observed northward steering flow around the western flank of the anticyclone. This in turn results in the forecasted storm continuing on a northwestward track, while the observed storm recurves around the western side of the subtropical anticyclone. Normally, a weak subtropical high typically leads to premature recurvature as the TC is influenced by an approaching midlatitude trough, but there is no obvious midlatitude trough influencing Jangmi in the early stages of its recurvature (Fig. 8, top right panel). There is some evidence, both at 500 hPa (Fig. 8, bottom panels) and in the upper troposphere (not shown), that the ridge (over China)–trough (over Japan) pattern has higher amplitude in the forecast than in the analysis, which would result in a stronger northerly flow north of
Jangmi, and may have inhibited recurvature in the forecast. The CAMV_DROP and NO_ADD analyses and 48-h forecast heights at 5890, 5905, and 5920 m are denoted by the solid and dashed blue curves, respectively. In contrast to the Nuri case, it appears that the subtropical high in CAMV_DROP is stronger than in NO_ADD in both analyses and forecasts, resulting in an improved (more northerly) position of the 48-h TC. However, even with the additional data, the subtropical high is still too weak in the forecast, and both CAMV_DROP and NO_ADD (as well as CAMV, DROP, SYNTH1, and SYNTH2) fail to capture the observed recurvature of the storm.

Both case studies suggest track errors due to an erroneous weakening of the Pacific subtropical anticyclone in the forecasts. We now examine the entire 2-month period to determine if this is a systematic error and consider what impact, if any, additional observations have on this error. Time–longitude plots illustrate the impact of additional observations on the analyses and forecasts and provide information as to whether these impacts are systematic or random in time. Figure 9 (left panel) shows a time–longitude plot of the difference between the CAMV and NO_ADD analyzed 500-hPa height, averaged from 15° to 30°N. The pattern appears random over much of the domain, but in the western Pacific (between the thick black lines) the enhanced AMVs typically increase the analyzed heights by 2–5 m over much of the 2-month period. The enhanced AMVs also tend to lower heights over the eastern Atlantic and Africa. The systematic tendency for the AMVs to increase heights in the western Pacific indicates that the additional data are correcting the low-height model bias in this region. The impact of the AMVs is slightly larger during the Jangmi period in late September than during the Nuri period in mid-August. In contrast, the drop-windsonde data result in smaller height differences on average, and do not result in systematic changes through the entire time period (Fig. 9, right panel). The smaller and less systematic impact of the drop-windsonde data...
as compared to the enhanced AMVs is consistent with the fact that the dropwindsonde data are intermittent and localized. The impact of the dropwindsonde data is not limited to the Atlantic and western Pacific regions where the dropwindsonde observations are actually taken, illustrating the nonlocal impact of observations as propagated through the update cycle (e.g., Aberson 2011).

The left panel of Fig. 10 is the CAMV 48-h forecast error (using the CAMV analysis as verification).2 The low (blue) values found between approximately 120°E and 180° confirm a low-height bias in the western Pacific, consistent with the biases found in the case studies (Figs. 7 and 8). This height bias modulates somewhat in time but is overall very persistent. Forecast biases at several different pressure levels (not shown) indicate that low height and temperature biases are evident throughout the lower to midtroposphere in the subtropical western North Pacific. In the upper troposphere (200 hPa), the height and temperature biases are modulated such that the low height biases are over the deep tropics, and high height biases exist over the subtropics. Despite the fact that the CAMV analysis starts out with systematically higher heights in this region, the model forecast bias reduces the potential forecast improvement from

![Figure 9](https://example.com/fig9.png)

**Fig. 9.** Time–longitude plot of the 500-hPa height difference (m) averaged from 15° to 30°N between the (left) CAMV and NO_ADD analyses and (right) DROP and NO_ADD analyses. Time periods for TCs Nuri and Jangmi are also indicated. Vertical black lines highlight the region between 120°E and 180°, discussed in the text.

---

2 The term “error” is used with caution, as the field would be different if a different analysis were used for verification. However, the differences in the analyses are small compared to the magnitude of the 2-day error, and using a different analysis (e.g., NO_ADD or CAMV_DROP) for verification yields qualitatively similar results. Comparison with ECMWF analyses for a few dates during this time period confirms a low forecast height bias in NOGAPS for this region.
the additional observations and improved analyses. To illustrate this, the right panel of Fig. 10 shows the difference between the CAMV and NO_ADD 48-h forecasts. There are no obvious systematic differences in the height field in the western Pacific after 48 h, and the differences are small compared to the forecast error in the left panel (note the contour change).

To investigate the impact of the enhanced AMVs on this model bias as a function of forecast time, the top panel of Fig. 11 shows 500-hPa height averaged over the entire 2-month period, from 15° to 30°N, as a function of longitude. CAMV results are given by the dashed curves, and NO_ADD results by the solid curves. The red, green, blue, and black curves correspond to 0-, 24-, 48-, and 72-h forecast times. Separation between the different color curves indicates a model drift in the time-averaged state. The largest separation between the red curves and the other colors occurs in the western Pacific, indicating that the model biases develop most strongly over the first 24 h in this region. The differences between the blue and black curves are quite small in this region, indicating that the model climate stabilizes after a few days of forecast integration. Close inspection shows that in the western Pacific, the dashed curve (CAMV) is higher than the solid curve (NO_ADD). This indicates that, despite uncertainties introduced by height assignment errors (e.g., Velden and Bedka 2009), the assimilation of the AMVs is working to correct for this bias. The curves for DROP (not shown) are basically coincident with the NO_ADD curves, indicating that dropwindsonde data do not have a significant impact on the time-mean state (as expected, given the intermittency and limited coverage of these observations).
Plotting the difference between the CAMV or NO_ADD 0-, 24-, 48-, and 72-h forecasts and the CAMV analysis shows this trend more clearly (bottom panel of Fig. 11). The dashed red curve is zero by definition (i.e., CAMV analysis minus CAMV analysis). The solid red curve (NO_ADD analysis minus CAMV analysis) obtains its largest negative value between 120°E and 180°, indicating that the analysis without the enhanced AMVs is lower than the analysis with AMVs in this region. The green, blue, and black curves also reach their lowest values in this region. At 24 h (green curves), while the CAMV forecast is still slightly higher than the NO_ADD forecast between about 130° and 140°E, the difference is very small. By 48 and 72 h (blue and black curves), the improvements from the enhanced AMVs are lost, and the CAMV and NO_ADD biases are almost identical. The lowest values in the western Pacific are approximately −8 m after 24 h, −14 m after 48 h, and −15 m after 72 h. That is, the height bias has grown approximately 8 m in the first forecast day, 6 m in the second day, and only 1 m in third day. This behavior indicates that the model bias in this region is not one that grows linearly in time, but rather one that is caused by the forecasts moving toward a particular climatology. This movement toward the model climatology is not appreciably impacted by changes to the analyses. Near 0° longitude, the enhanced AMVs reduce time-mean heights, but the forecast bias is smaller in this region.

Figure 12 is analogous to the lower panel in Fig. 11, but calculated for the (top) Nuri and (bottom) Jangmi periods only. The forecast bias during Nuri is smaller than during Jangmi (minimum of approximately −15 m at 72 h as compared to −21 m), but the impact of the enhanced AMVs on the analysis is also smaller (−2.5 m minimum versus −3.5 m) and shifted eastward. After 24 h, during Nuri, the dashed and solid green lines are very close, indicating that the enhanced AMVs impact has mostly disappeared after one day. During Jangmi, the 24-h bias is about 2 m stronger in NO_ADD (solid green) than in CAMV (dashed green). These results are consistent with some improvement of Jangmi TC tracks from enhanced AMVs, whereas no such improvement was found for Nuri. There are, of course, other potential reasons for lack of improvement from additional observations, including the influence of synthetic observations, examined in the next subsection.

The cause of this particular model bias is not known. Height biases in the Atlantic are neither as strong nor as temporally consistent as in the western Pacific (Fig. 10). Model errors and biases may be due to errors in the physical parameterizations or numerics and to limited horizontal and/or vertical resolution. Improvements in the parameterization of convection in the ECMWF system in November 2007 resulted in a substantial reduction in precipitation biases and 500-hPa height biases (Jung et al. 2010), as well as substantial improvements in TC track predictions (Fiorino 2009). Han and Pan (2011) showed that improved convection and vertical diffusion parameterizations resulted in substantial reductions in tropical wind errors and TC track errors in the NOAA GFS. Similarly, Torn and Davis (2012) illustrate how an improved representation of shallow convection in the Advanced Hurricane Weather Research and Forecasting Model led to substantial improvements in tropospheric heating profiles. Concurrent
reductions in temperature biases (and their horizontal gradients) led to reduced wind biases and improved TC track forecasts.

Previous experiments with NOGAPS have shown significant ensemble TC track improvements when resolution increased from T119 to T159, with additional smaller gains at T239 (Reynolds et al. 2011a). These improvements were substantially larger in the western Pacific than in other basins. Studies with other models have shown that model bias is sensitive to resolution (e.g., Palmer and Weisheimer 2011; Matsueda and Palmer 2011). However, it was not established that the NOGAPS TC track improvements in Reynolds et al. (2011a) were due to reductions in biases. Reynolds et al. (2011b) showed that small but statistically significant ensemble-mean TC track improvements were obtainable through inclusion of parameter variations in the NOGAPS ensemble. This suggests that uncertainty or errors in the physical parameterizations may be contributing to TC track errors.

c. Influence of synthetic observations

The synthetic observations in NOGAPS operational forecasts are currently assigned very small observation errors (Table 2) in order to have the analyses draw closely to them. We now show results in which the influence of the synthetic observations is decreased by increasing the assumed errors assigned to them. As noted above, in SYNTH1, the synthetic observations are assumed to be less accurate than dropwindsonde data but more accurate than AMVs. In SYNTH2, the synthetic observations are assumed to be less accurate than both dropwindsonde data and AMVs.

To illustrate the impact of the synthetic observations relative to dropwindsonde data and enhanced AMVs, the difference in 850-hPa vorticity analyses between CAMV_DROP and NO_ADD (upper left), CAMV (upper right), DROP (lower left), and SYNTH1 (lower right) at 0000 UTC 25 September are shown in Fig. 13. The 850-hPa vorticity is also shown for each of the experiments in red, and for CAMV_DROP in all four panels in black. The differences introduced through dropwindsonde data (upper right) are somewhat localized to the storm, extending 10° outward from the storm center to the east and west, extending beyond the immediate vicinity of the dropwindsonde locations (indicated by black dots). Because of the cyclic nature of the data assimilation system, the CAMV_DROP and CAMV differences are not only due to dropwindsonde data in the current assimilation cycle, but also represent cumulative impacts from dropwindsonde data used in previous update cycles. The differences due to enhanced AMVs (bottom left) cover a broader area consistent with the wide spatial coverage of this data [see Figs. 2 and 3 in Berger et al. (2011) for an example]. Because AMVs are based on both cloud and water vapor features, their impact on the analysis extends into the relatively cloud-free region of the subtropical high. The differences due to dropwindsonde data and enhanced AMVs (upper left) occur both locally and remotely, as expected. The differences due to moderately increasing the synthetic observation assigned error variance (lower right) are localized near the storm, and are about half the magnitude of the differences introduced through the dropwindsonde data. As these differences are a significant fraction of the size of the differences introduced through the dropwindsonde data, one might expect significant changes to individual forecast tracks. Cross sections of 850-hPa meridional wind through the
TC center at 14°N (not shown) indicate that reducing the weight of the synthetic observations allows the analyses to move closer to the dropwindsonde observations, although at this coarse resolution the model does not capture the tight horizontal wind gradients or maximum wind speeds observed by the dropwindsondes or provided by the synthetic observations. The fit to the dropwindsonde data is explored in more detail using data assimilation innovation and residual statistics below.

Figure 14 shows the TC track error for all basins for CAMV_DROP, SNYTH1 and SYNTH2. The moderate increase in assigned synthetic observation error (SYNTH1) leads to small improvements in the forecast track after 2 days. A further increase in the assigned synthetic observation error (SYNTH2) does not give consistently better results (marginally better between 60 and 108 h, and worse at 120 h). Improvements for both SYNTH1 and SYNTH2 over CAMV_DROP are statistically significant between 72 h and 96 h. Figure 15 shows the difference in the average TC track forecast error for the different basins between SYNTH1 and CAMV_DROP (top panel) and between SYNTH2 and CAMV_DROP (bottom panel). The biggest impact (statistically significant from 48 to 120 h) is seen in the eastern Pacific for almost all forecast times. This may be related to the relative sparseness of in situ observations in the area. In the Atlantic, both SYNTH1 and SYNTH2 improve TC track forecasts in the midforecast ranges (larger improvements from SNYTH2, significant at 72–96 h, than SYNTH1), but degrade forecasts at the longest forecast times, perhaps due to the small sample size of forecasts. In the western Pacific, SYNTH1 has a near-neutral impact up to 60 h and positive impact after that (significant at 84–120 h), while SYNTH2 has a mostly neutral to negative impact.

An examination of the statistics produced from the data assimilation cycle allows for a more complete...
understanding of the impact of the assigned synthetic observation error changes. Figure 16 (top panel) shows the mean absolute difference between the background and the synthetic observations (innovations; three right curves) and difference between the analysis and the synthetic observations (residuals; three left curves) for the meridional wind at 1000, 925, 850, 700, 500, and 400 hPa (the levels at which synthetic observations are produced). When comparing results for CAMV_DROP, SYNTH1, and SYNTH2, both the innovations and residuals increase as the errors assigned to those observations are increased, as expected. The residuals for SYNTH1 and SYNTH2 are statistically significantly larger (at the 95% level) than those for CAMV_DROP at all pressure levels considered.

The mean absolute differences between the background and the dropwindsonde data (innovations), and the analysis and the dropwindsonde data (residuals), for the same pressure levels are shown in the bottom panel of Fig. 16. While the differences are small, the reduced weight of the synthetics does allow for a closer fit of the analysis and the background to the dropwindsonde observations, that is, a reduction in both innovations and residuals. The innovations are statistically significantly smaller (at the 95% level) for SYNTH1 at 500 hPa and for SYNTH2 between 925 and 500 hPa. The residual reductions do not reach 95% significance, but do reach 90% significance for SYNTH1 at 700 hPa and SYNTH2 at 925 hPa.

The average TC track errors for SYNTH1 and SYNTH2 for Nuri and Jangmi are included in Fig. 5. For both storms the SYNTH1 and SYNTH2 track errors are smaller than the CAMV_DROP errors, with SYNTH1 either comparable to or slightly better than SYNTH2 at most forecast times. For Nuri, SYNTH1 improves (degrades) TC track forecasts (Fig. 6) by more than 5% in 48% (20%) of the cases. For SYNTH2, the results are comparable, with improvements (degradations) in 43% (30%) of the cases. For Jangmi, there is a very consistent improvement with SYNTH1, with 76% of the cases showing improvement and only 5% showing degradation. The results for SYNTH2 are just slightly less promising, with improvements (degradations) in 71% (11%) of the cases.
Update cycle/forecast experiments run without any synthetic observations did not perform well. Removal of the synthetics resulted in statistically significant increases in TC track error over CAMV_DROP by 24 n mi (or 68%) at 0 h and 36 n mi (16%) at 96 h. Examination of statistics produced from the data assimilation update cycle show statistically significant increases in both the innovations and residuals for the dropwindsonde data between 1000 and 500 hPa (in SYNTH1 and SYNTH2, these innovations and residuals were reduced). In addition, the number of storms identified by the TC tracker was reduced by 5% at 0 h and 11% at 120 h (for comparison, SYNTH2 showed reductions of 1% and 5%). These “no-synthetic” experiments indicate that, at this relatively coarse resolution, some type of special TC initialization is necessary. A full examination of methods to optimize the use of synthetic observations is beyond

**FIG. 15.** Difference (n mi) between the average TC track error from (top) SYNTH1 and CAMV_DROP and (bottom) SYNTH2 and CAMV_DROP as a function of basin (WPAC–western North Pacific; ATL–Atlantic; EPAC–eastern North Pacific; ALL–all basins) as denoted in key. Number of storms in each basin for both panels is denoted in italics beneath bottom panel. Forecast times for which the differences are statistically significant at the 95% (90%) level are denoted by black (gray) symbols at the top of each panel for WPAC (×), ATL (squares), and EPAC (triangles).

**FIG. 16.** Mean absolute difference (m s$^{-1}$) between the background and observations (innovations, three right curves), and analysis and observations (residuals, three left curves) for the meridional wind for (top) synthetic and (bottom) dropwindsonde observations, for CAMV_DROP (solid), SYNTH1 (dotted), and SYNTH2 (dashed). Total number of synthetic observations is 2682 at each level. Total number of dropsonde observations varies from 830 at 400 hPa to 1390 at 850 hPa. The reduction in residuals for the dropwindsondes is statistically significant at the 90% level at 700 hPa for SYNTH1 and at 925 hPa for SYNTH2. The reduction in innovations is statistically significant at the 95% level from 925 to 500 hPa for SYNTH2 and at 500 hPa for SYNTH1.
the scope of this study, but these results indicate that a reformulation of the synthetic observations may lead to track forecast improvements. Ideally, as the resolution of the forecast model improves, and as more in situ and/or remote observations become available, the need for the synthetic observations should diminish.

4. Summary and conclusions

Six experiments using the NOGAPS/NAVDAS-AR data assimilation and forecasting system have been conducted to examine the impact of Atlantic and Pacific dropwindsonde data, CIMSS enhanced AMVs, and increased assigned synthetic observation errors on TC track forecasts during the August–September 2008 T-PARC/TCS-08 period. The largest average reduction in track error (up to 18%) due to the enhanced AMVs is found in the eastern North Pacific, with reductions of up to 11% in both the Atlantic and western North Pacific basins. The dropwindsonde data show a more consistent positive impact in the Atlantic (up to an 11% reduction) than in the western North Pacific (up to 8%). The dropwindsonde data have a small impact on eastern North Pacific storms through the global propagation of the impact of observations taken in other basins. The impact of the enhanced satellite winds and dropwindsonde data are highly storm dependent. Case studies for Typhoons Jangmi and Nuri indicate that while the additional observations improved the Jangmi forecasts somewhat, they failed to improve the Nuri forecasts, on average.

Synoptic evaluation of the case studies suggests that the forecast errors for both storms are impacted by the erroneous weakening of the subtropical anticyclone in the western North Pacific during the forecast integration. Examination of the full 2-month period indicates that this low height bias is a persistent feature. The assimilation of the enhanced AMVs systematically raises the analyzed heights in the western North Pacific, counteracting this bias. However, the impact of the additional observations diminishes rapidly during the first few days of forecast integration. This bias does not increase linearly with forecast time. Rather, the forecasts move quickly to a preferred climatological state, and additional data do little to mitigate this. However, it does appear that the impact of enhanced AMVs lasts farther into the integration during Jangmi than during Nuri. This is consistent with the fact that assimilation of the additional observations improves Jangmi forecasts but has little impact on Nuri forecasts.

Experiments are also performed in which the errors assigned to the synthetic TC observations are increased. In the first experiment, the synthetic observation errors are moderately increased, such that they are considered less accurate than dropwindsonde data, but more accurate than AMVs. In the second experiment, the synthetic observation errors are increased substantially, such that they are considered less accurate than the AMVs. The moderate increase in the assigned error improves forecast tracks in the eastern North Pacific at all forecast lead times, and in the western North Pacific for lead times greater than 60 h. In the Atlantic the results are mixed. When the assigned errors are increased substantially, forecasts in the eastern North Pacific improve further, while forecasts in the western North Pacific degrade, and the results for the Atlantic are still mixed. Examination of innovation and residual statistics shows that increasing assigned synthetic observation error allows for a closer fit of the background and analysis to the dropwindsonde observations. Experiments in which the synthetic observations are removed entirely did not perform well, and indicate that some type of special consideration of the TC vortex in the update cycle is still needed in a system operating at this resolution.

These findings illustrate how large-scale model biases may limit the potential improvements available through the assimilation of additional observations and improvements to the analysis, and how this impact may vary substantially from storm to storm. The cause of this bias has not been established, but work on improving several aspects of the model formulation, including numerics, physical parameterizations, and increased horizontal resolution, is ongoing at NRL. In 2013 NOGAPS has been replaced by the Navy Global Environmental Model, with semi-Lagrangian advection, more advanced physics, and increased resolution (T359L50). In addition, the number of raw hourly AMVs available for assimilation has approximately doubled since 2008. Future anticipated improvements in the assimilation of observations include assimilating the GPS dropwindsonde observations at the observed location [accounting for drift as the dropwindsonde falls, which can be significant, as shown by Aberson (2008)], and the assimilation of high-density flight level observations. The results obtained by increasing the assigned synthetic observation errors suggest that the formulation of the synthetic observations should be revisited as model resolution is increased and as more observations are assimilated.

Acknowledgments. We thank Howard Berger for producing the ATCF track error statistics, Efren Serra for producing ATCF tracks for other operational forecast models, and Benjamin Ruston for guidance in running the data assimilation and forecast system. We thank three anonymous reviewers for suggestions that have improved the paper. This research is supported by the Chief of Naval Research through the NRL Base Program,
REFERENCES


Fiorino, M., 2009: Record-setting performance of the ECMWF IFS in medium-range tropical cyclone track prediction. ECMWF Newsletter, No. 188, ECMWF, Reading, United Kingdom, 20–27.


