SPECIAL Tropical Cyclone Predictability COLLECTION

SPECIAL THORPEX COLLECTION

The Influence of Assimilating Dropsonde Data on Typhoon Track and Midlatitude Forecasts

MARTIN WEISSMANN,* FLORIAN HARNISCH,* CHUN-CHIEH WU,+ PO-HSIUNG LIN,+
YOICHIRO OHTA,# KOJI YAMASHITA,# YEON-HEE KIM,@ EUN-HEE JEON,@
TETSUO NAKAZAWA,& AND SIM ABERSON**

* *Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany*
+ *Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan*
# *Japan Meteorological Agency, Tokyo, Japan*
@ *National Institute of Meteorological Research, Korea Meteorological Agency, Seoul, South Korea*
& *Meteorological Research Institute, Tsukuba, Japan*
** *NOAA/AOML/Hurricane Research Division, Miami, Florida*

(Manuscript received 9 February 2010, in final form 21 April 2010)

ABSTRACT

A unique dataset of targeted dropsonde observations was collected during The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC) in the autumn of 2008. The campaign was supplemented by an enhancement of the operational Dropsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) program. For the first time, up to four different aircraft were available for typhoon observations and over 1500 additional soundings were collected.

This study investigates the influence of assimilating additional observations during the two major typhoon events of T-PARC on the typhoon track forecast by the global models of the European Centre for Medium-Range Weather Forecasts (ECMWF), the Japan Meteorological Agency (JMA), the National Centers for Environmental Prediction (NCEP), and the limited-area Weather Research and Forecasting (WRF) model. Additionally, the influence of T-PARC observations on ECMWF midlatitude forecasts is investigated.

All models show an improving tendency of typhoon track forecasts, but the degree of improvement varied from about 20% to 40% in NCEP and WRF to a comparably low influence in ECMWF and JMA. This is likely related to lower track forecast errors without dropsondes in the latter two models, presumably caused by a more extensive use of satellite data and four-dimensional variational data assimilation (4D-Var) of ECMWF and JMA compared to three-dimensional variational data assimilation (3D-Var) of NCEP and WRF. The different behavior of the models emphasizes that the benefit gained strongly depends on the quality of the first-guess field and the assimilation system.

1. Introduction

Tropical cyclone (TC) track forecasts have improved significantly over the past decades. The U.S. National Hurricane Center reported a reduction of its official 24–72-h mean track forecast error of nearly 50% in the time frame 1980–2008 for the Atlantic and eastern North Pacific (NOAA/NWS/NHC 2009). Similar improvements were reported for Japan Meteorological Agency (JMA) official typhoon forecasts in the time period 1982–2007 (see the annual report for 2007 online at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/annual-report.html). A substantial part of these improvements is likely due to advanced numerical models, increased resolution, advanced data assimilation, and the steady increase of satellite observations assimilated in different models.

Targeted airborne dropsonde observations are another factor that has contributed to improvements of TC track forecasts. Several studies documented that average track forecast improvements of the order of 10%–30% can be achieved with additional airborne dropsonde observations deployed in the vicinity of tropical storms or in sensitive areas calculated by an ensemble transform Kalman

Corresponding author address: Martin Weissmann, DLR Oberpfaffenhofen, Münchner Straße 20, D-82230 Wessling, Germany.
E-mail: martin.weissmann@dlr.de

DOI: 10.1175/2010MWR3377.1

filter (ETKF), ensemble variance, adjoint sensitivity, or singular vectors (Aberson 2003, 2008; Wu et al. 2007a,b, 2009; Pu et al. 2008; Chou and Wu 2008). However, the achieved reduction of the track forecast error also depends on the errors in the analysis without additional observations, as a more accurate analysis is less likely to improve significantly by additional observations. All of the mentioned studies were carried out with global model versions using three-dimensional variational data assimilation (3D-Var), whereas many of the leading NWP centers use four-dimensional variational data assimilation (4D-Var) nowadays. The use of 4D-Var has also enabled a drastic increase of satellite observations assimilated in NWP models. These modifications are likely to reduce the beneficial influence of additional observations. However, impact studies with 4D-Var assimilation systems of global models are limited to a low number of cases (e.g., Yamaguchi et al. 2009).

Operational surveillance flights for tropical cyclones have been operated in the Atlantic and eastern Pacific for more than a decade (Aberson 2002; Burpee et al. 1996) and dropsondes by several aircraft have been regularly deployed inside hurricanes, in their environment, and in calculated sensitive regions. In contrast, operational dropsonde observations over the western Pacific Ocean have been limited to a single aircraft deployed through the Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) program (Wu et al. 2005). In autumn 2008, an international effort by several countries in North America, East Asia, and Europe was made to observe TCs in the western Pacific throughout their full life cycle from the genesis in the tropics until extratropical transition and the interaction with the midlatitude flow for the first time. These observations were conducted under the roof of the The Observing System Research and Predictability Experiment (THORPEX) program of the World Meteorological Organization (more information is available online at www.wmo.int/thorpex/). This field campaign, called the THORPEX Pacific Asian Regional Campaign (T-PARC), intended to investigate TC formation, structure changes, and extratropical transition of TCs but also contained strong efforts to perform targeted observations to improve TC track prediction.

In addition, T-PARC was connected to two other projects that coordinated their observations: the operational DOTSTAR program enhanced its flight activity and the U.S. Navy conducted the Tropical Cyclone Structure Experiment (TCS-08). Altogether, up to four aircraft were simultaneously available in a two month period: One U.S. Air Force WC-130 aircraft, which could penetrate into the eye of TCs; one U.S. Navy P-3 aircraft, which focused on rainbands and the structure of convection; the Falcon 20 aircraft of the Deutsches Zentrum für Luft- und Raumfahrt (DLR), which focused on sensitive regions for typhoon forecasts calculated by singular vector, adjoint, and ETKF methods; and the DOTSTAR Astra Jet, which usually circumnavigated the storm and also conducted observations in sensitive regions. Regular international video conferences were conducted to discuss the wide range of sensitive area calculations available and make coordinated flight plans for all four aircraft. In addition, driftsonde gondolas were launched on Hawaii, which released dropsondes while drifting toward Asia with the easterlies in the lower stratosphere. JMA conducted additional radiosonde soundings (TEMP) from research vessels and ground stations, in situ synoptic observations (SYNOP) on research vessels, and JMA's Meteorological Satellite Center produced *Multifunctional Transport Satellite (MTSAT-2)* rapid scan atmospheric motion vectors. More than 500 aircraft flight hours were spent and more than 1500 additional soundings were made. These soundings constitute a unique dataset to investigate the benefit of targeted observations for typhoon forecasting, to compare targeting strategies (Harnisch and Weissmann 2010) and to optimize the use of dropsondes in NWP models, in particular the use of observations in the TC core and eyewall region.

This study focuses on the forecast influence of special observations during the two major typhoon events of T-PARC, Typhoon Sinlaku and Typhoon Jangmi, both in September 2008 (Fig. 1). The storms formed in the Philippine Sea and near Guam, respectively, then both headed toward Taiwan where they caused severe flooding, recurved east of China, and passed south of Japan on their way eastward. The storms had peak typhoon intensities of category 4 (Sinlaku) and 5 (Jangmi) according to the Saffir–Simpson scale before hitting Taiwan. Jangmi was even the strongest tropical storm worldwide in 2008. Both storms weakened significantly while passing over Taiwan, but Sinlaku reintensified again near Japan while Jangmi also touched the Chinese coast and did not reach typhoon intensity after recurvature.

Flights for both systems were performed from early stages in the tropics through to their extratropical transition. Flights in the early stages were performed inside and in the environment of the typhoons and in areas indicated to be sensitive by targeting guidance optimized on the typhoon. In the later stages, flights investigated the internal structure of the systems, but also the interaction of the storms with the midlatitude jet and several flights aimed to sample regions indicated to be sensitive by ETKF and singular vector calculations for midlatitude verification areas downstream over the northern Pacific.

Data denial experiments with three global and one limited-area model were conducted to investigate the influence of special T-PARC observations on typhoon

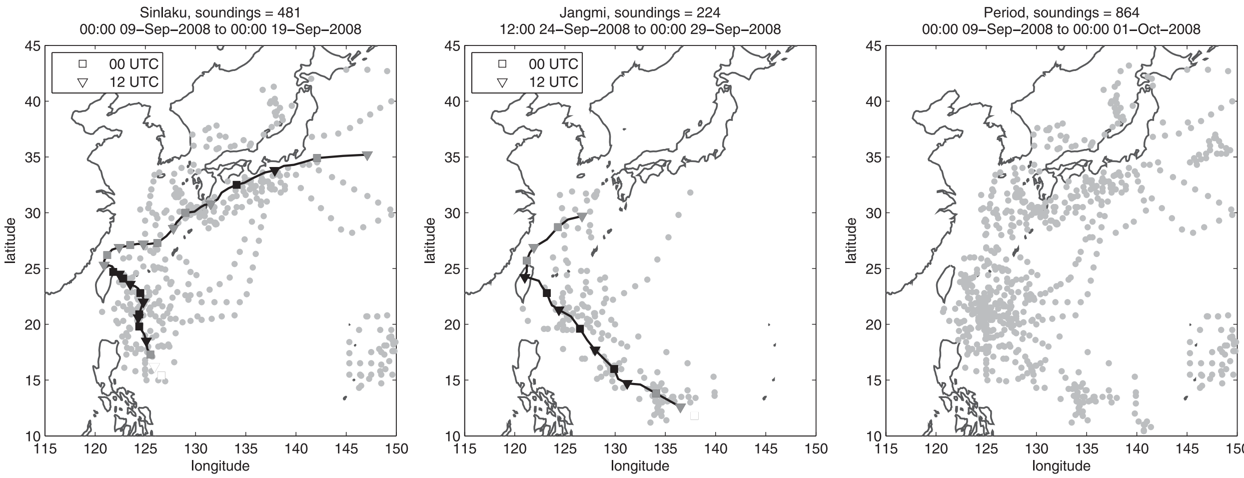

FIG. 1. JMA best tracks of (left) Sinlaku and (middle) Jangmi with dropsonde and ship TEMP locations in the respective period (gray points). Squares (triangles) along the track show the positions of Sinlaku and Jangmi at 0000 (1200) UTC starting at 0000 UTC 9 Sep 2008 (1200 UTC 24 Sep 2008). Black symbols indicate typhoon intensity and gray symbols TC intensity. (right) The location of all dropsondes and ship TEMP during the period 9 Sep–1 Oct 2008.

track forecasts and European Centre for Medium-Range Weather Forecasts (ECMWF) results were also analyzed to investigate the influence of these observations on midlatitude forecasts. The clear majority of additional observations relevant for typhoon and midlatitude forecasts consisted of airborne dropsondes. As a consequence, the influence of all additional observations will be abbreviated with ''dropsonde influence'' although extra TEMP and SYNOP observations were also included in some of the experiments.

The numerical models and experiment setup are described in section 2. The influence of the observations on typhoon track forecasts by different models is presented in section 3, followed by a discussion of the influence of T-PARC observations on ECMWF forecasts in midlatitudes over the Pacific and on the Northern Hemisphere in section 4. The discussion and conclusions are presented in section 5.

2. Model descriptions

a. JMA GSM experiment description

To evaluate the impact of the T-PARC 2008 special observations, experiments using the operational global 4D-Var system and the operational JMA global spectral model (GSM) were carried out (Tables 1 and 2). GSM data assimilation (GSM-DA) cycles were run every 6 h (2100–0300, 0300–0900, 0900–1500, and 1500–2100 UTC). Forecasts up to 84 h were executed from 4 initial times per day (0000, 0600, 1200, and 1800 UTC) and forecasts up to 216 h for selected case studies. The JMA GSM is a hydrostatic spectral model with a horizontal resolution of ~20 km (the inner-loop model for the GSM-DA is ~80 km) and 60 levels in the vertical with the top level at 0.1 hPa.

Two cycled experiments were performed for the whole period: one experiment assimilating special T-PARC observations (DROP) including Pacific dropsondes and TEMP from JMA research vessels and observatories and one experiment without special observations (NODROP).

The majority of dropsonde observations was used and erroneous observations were excluded by internal and external quality checks contained in the JMA assimilation system. JMA assimilates TC bogus data to generate realistic TC structures in the analysis fields of the operational system. In both of experiments (DROP and NODROP), TC bogus data were not used to evaluate the pure impact of the special observations on the model.

b. NCEP GFS experiment description

The operational National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) was run with a horizontal resolution of ~38 km and 64 levels in the vertical. Two model runs were made for a period covering Sinlaku and Jangmi: one without dropsonde observations (NODROP) and one that assimilates dropsondes (DROP). Special TEMP and SYNOP were not denied. All other observations from the NCEP archive were ingested into the assimilation system for both runs. The experiments were conducted in a cycled mode for the whole T-PARC period. Dropsondes in the Atlantic were also removed in the NODROP run.

The NCEP Global Data Assimilation System (GDAS) consists of a quality control algorithm, a TC vortex initialization procedure, data assimilation, and the global

TABLE 1. Comparison of the four models used in this study.

	ECMWF IFS	JMA GSM	KMA WRF	NCEP GFS
Resolution	TL799L91 (~25 km)	TL959L60 (~20 km)	30 km	T382L64 (~38 km)
DA method	12-h 4D-Var	6-h 4D-Var	6-h 3D-Var	6-h 3D-Var
Domain	Global	Global	190 × 190 grid points	Global
Bogus	No	No (Yes in operational version)	No	Vortex relocation, bogus if no vortex in first guess (rare)
Use of TC core obs	Yes	Yes	Yes	No
Denied obs	Pacific dropsondes, driftsondes, JMA ship SYNOP, JMA ship TEMP	Pacific dropsondes, JMA ship TEMP, JMA special TEMP	Pacific dropsondes	Atlantic and Pacific dropsondes, driftsondes

spectral model. The quality control involves optimal interpolation and hierarchical decision making to evaluate the observations before input to the analysis (Woollen 1991). A vortex relocation procedure (Liu et al. 2000) in which TCs in the first-guess field are relocated to the official Tropical Prediction Center position in each 6-h analysis cycle (as in Kurihara et al. 1995) ensures that the systems are located in the operationally fixed locations. The analysis scheme is the gridpoint statistical interpolation (Wu et al. 2002): the background field (the previous 6-h forecast) is combined with observations with a 3D-Var multivariate formalism (Kleist et al. 2009). Dropsonde observations within a radius of 111 km from the typhoon center (or 3 times the specified radius of maximum wind if larger than 111 km) are currently not used in the NCEP analysis (Aberson 2008).

c. KMA WRF experiment description

One experiment with dropsondes (DROP) and a control run without additional observations (NODROP) were performed for 17 selected analysis times with additional observations. In contrast to the experiments with the three global models, the WRF experiments were not cycled (i.e., the DROP and NODROP run used the same first-guess field). The adaptively observed data were assimilated by the Advanced Research Weather Research and Forecasting (ARW-WRF version 3.0.1.1) 3D-Var assimilation system. The background error covariances were computed by the National Meteorological Center (NMC) method using the operational Korea Meteorological Administration (KMA) Global Data Assimilation and Prediction System (GDAPS) forecast data (0000 and 1200 UTC) from 1 to 30 September 2008. The WRF model was used for forecasts up to a lead time of 72 h. The horizontal resolution was 30 km (190 × 190 grid points) and the GDAPS (T426L40) global model data were used for initial and boundary conditions. The physics packages included the WSM6 microphysics scheme, the Kain–Fritsch cumulus parameterization scheme, the Noah land surface model, the Yonsei University planetary boundary layer, a simple cloud-interactive radiation scheme, and Rapid Radiative Transfer Model longwave radiation schemes.

d. ECMWF IFS experiment description

Two experiments were conducted with the spring 2009 version (cycle 35r2) of the ECMWF Integrated Forecast System (IFS) for a period in September 2008 that covers the whole evolution of Sinlaku and Jangmi: a control run (NODROP) without any special observations (i.e., Pacific dropsondes, driftsondes, JMA ship SYNOP, and TEMP) and one experiment including all special observations (DROP). These experiments were cycled [i.e., special observations influence the analysis at

TABLE 2. Assigned errors for dropsonde wind components and temperature observations in different models.

	ECMWF IFS	JMA GSM	KMA WRF	NCEP GFS
200 hPa	2.50 m s^{-1}	U: 3.10 m s^{-1} V: 3.10 m s^{-1}	3.3 m s^{-1}	2.95 m s^{-1}
	0.84 K	1.30 K	1 K	1.2 K
300 hPa	2.60 m s^{-1}	U: 3.60 m s^{-1} V: 3.60 m s^{-1}	3.3 m s^{-1}	3.40 m s^{-1}
	0.70 K	1.10 K	1 K	0.9 K
500 hPa	2.10 m s^{-1}	U: 2.80 m s^{-1} V: 2.80 m s^{-1}	2.3 m s^{-1}	2.80 m s^{-1}
	0.66 K	1.00 K	1 K	0.8 K
700 hPa	1.90 m s^{-1}	U: 2.70 m s^{-1} V: 2.60 m s^{-1}	1.4 m s^{-1}	2.40 m s^{-1}
	0.77 K	1.00 K	1 K	0.8 K
1000 hPa	1.80 m s^{-1}	U: 2.10 m s^{-1} V: 2.30 m s^{-1}	1.1 m s^{-1}	2.40 m s^{-1}
	0.98 K	1.30 K	1 K	1.2 K

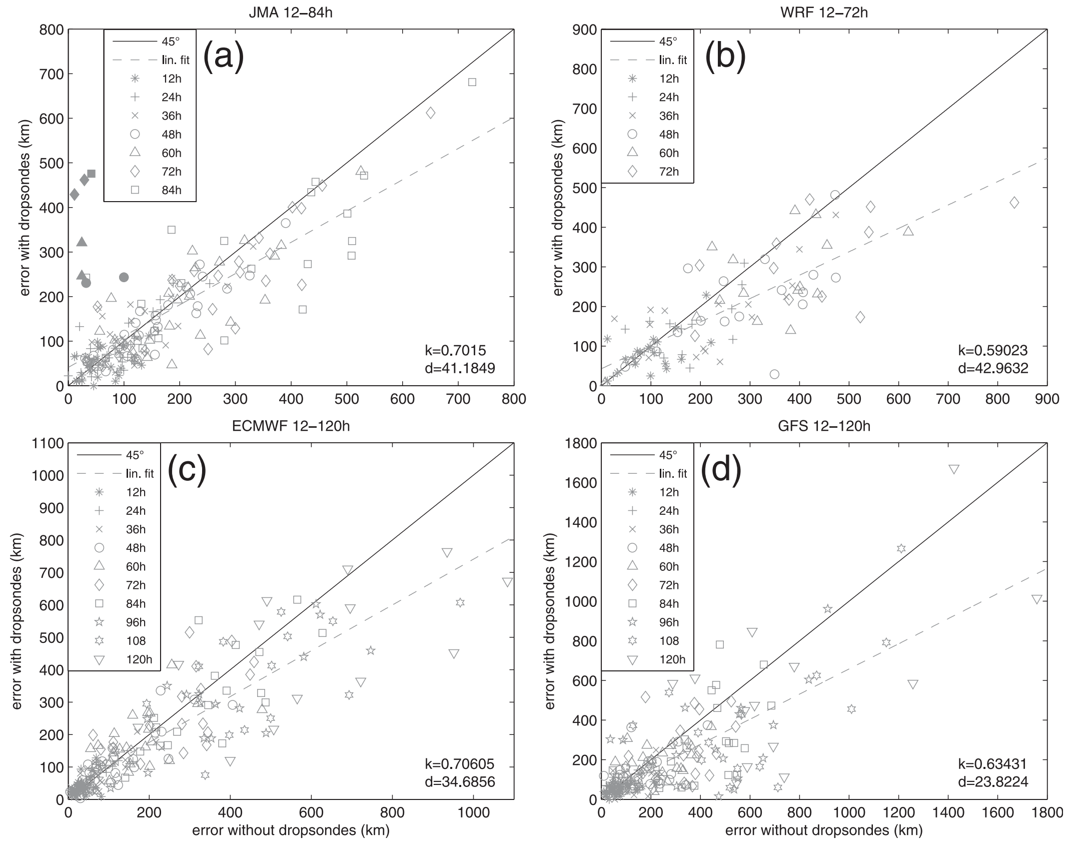

FIG. 2. Scatter diagrams of all available DROP and NODROP track forecast errors for Sinlaku and Jangmi by (a) JMA, (b) WRF, (c) ECMWF, and (d) NCEP GFS. Filled symbols in (a) indicate 48–72-h forecasts initialized at 1200 UTC 11 or 27 Sep 2008. The linear fits are shown with dashed gray lines and their slopes and *y* intercepts are given in the bottom right of the figures. Values beneath the diagonal indicate that the errors of DROP are lower than the ones of NODROP.

the targeted observation time, but also the first guess (short-range forecast) of subsequent analysis times]. In addition, an uncycled experiment (DROP_UnCy) for all analysis times with observations was carried out, which used the first guess from NODROP.

The experiments were performed with a horizontal resolution of ~25 km, 91 vertical levels, and 4D-Var data assimilation with 12-h windows (0900–2100 and 2100–0900 UTC). Forecasts up to 240 h were initialized at 0000 and 1200 UTC. For further information about the ECMWF analysis and forecasting system see Rabier et al. (2000), Mahfouf and Rabier (2000), and Richardson et al. (2009).

Following the T-PARC field campaign, it was discovered that a significant fraction of dropsondes in the operational ECMWF analysis had timing errors. Thus, the dropsonde dataset of the whole period was time corrected (Pacific and also Atlantic dropsondes) for the experiments.

The ECMWF assimilation system contains a first-guess check and a variational quality control. The first-guess check for dropsondes is strongly relaxed (nearly inactive) for latitudes of less than 30° to avoid high rejections in and near TCs. This modification was extended up to 40° latitude because Typhoon Sinlaku reintensified near 30°N. The ECMWF IFS also assimilates dropsondes in the core and eyewall of typhoons (in contrast to NCEP), but about half of these observations are usually rejected by the variational quality control.

3. The influence on typhoon track forecasts

a. JMA GSM

The official JMA best-track data was used for the verification of the typhoon track forecasts (Fig. 1). Additional observations lead to a typhoon track forecast improvement in the majority of JMA forecasts (Fig. 2a) and also the linear fit of DROP and NODROP forecasts indicates an improvement with dropsondes. However, DROP forecasts from two initial times (1200 UTC on 11 and 27 September 2008) show large degradations that level out improvements at other times (Figs. 2a, 3a, and 4). Both of these initial times have nearly perfect NODROP forecasts

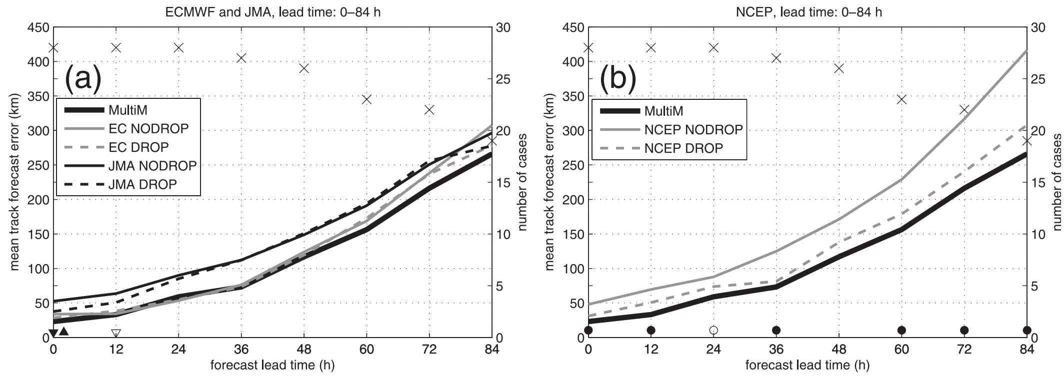

FIG. 3. Mean track forecast error of (a) ECMWF and JMA and (b) NCEP DROP and NODROP. A thick solid black line in both (a) and (b) shows the error of the multimodel DROP forecast (MultiM; average position of ECMWF, JMA, and NCEP track forecasts). The mean errors were calculated by using all initial/verification times where forecasts were available from all three models for initial times 0000 UTC 9 Sep–1200 UTC 18 Sep 2008 and 1200 UTC 24 Sep–0000 UTC 29 Sep 2008 and verification times before 1200 UTC 20 Sep and 1200 UTC 30 Sep 2008. The X symbols and the right *y* axis indicate the number of forecasts used for the calculation of the mean track error. Empty (filled) markers indicate times where the mean differences are significant at a 90% (95%) confidence level using a Student's *t* test. Circles, triangles, and inverted triangles stand for NCEP, ECMWF, and JMA significance results, respectively.

and the DROP analysis contains dropsonde observations from the U.S. WC-130 penetrating the typhoons. As a consequence, no mean track improvement by dropsondes is visible in JMA forecasts beyond a lead time of 12 h despite a significant improvement of the typhoon position at initial time (Fig. 3a, Table 3). Only a limited number of forecasts for lead times of more than 84 h are available. Thus, JMA is excluded from the comparison of track forecasts at lead times of 84–120 h. The available forecasts however, indicate no mean forecast improvement by dropsondes beyond 84 h as well (not shown).

For both DROP and NODROP, the mean error of the JMA track forecasts is about 20–40 km larger than the one of ECMWF for lead times of up to 60 h. In relative terms, this is nearly 40% of the JMA track error at 24 h reducing to 10% at 60 h. At 72–84-h lead time, the mean track errors of JMA and ECMWF are of a similar magnitude and the limited number of available longer-range

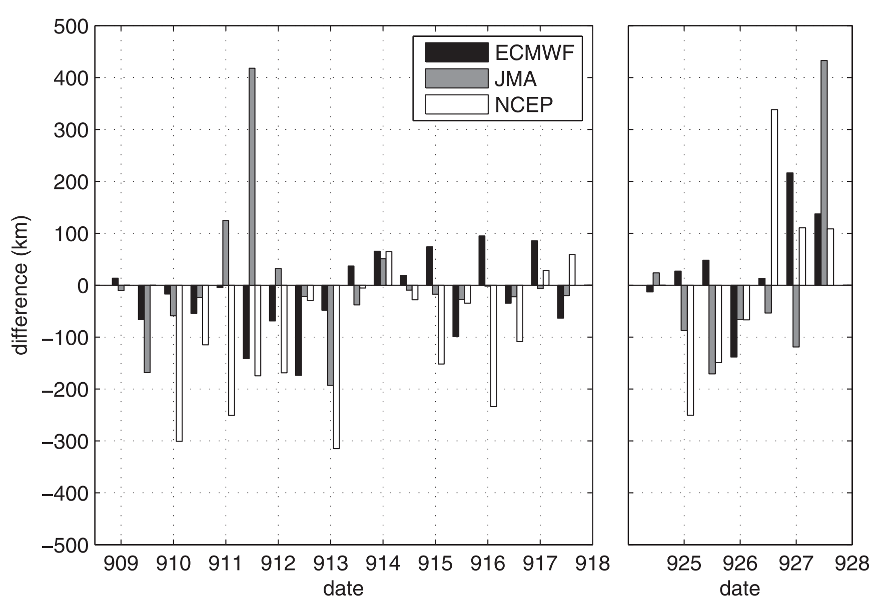

FIG. 4. Difference of 72-h DROP and NODROP track forecast error of ECMWF, JMA, and NCEP for initial times during 0000 UTC 9 Sep–1200 UTC 27 Sep 2008. Negative values indicate lower errors in the DROP experiment.

TABLE 3. Relative reduction of the mean track forecast error: Difference of DROP and NODROP mean track errors divided by the mean of both.

Lead time (h)	ECMWF	JMA	NCEP	WRF
0	−18%	−33%	−42%	1%
12	12%	−23%	−32%	−19%
24	3%	−6%	−18%	−16%
36	−4%	0%	−42%	−24%
48	−4%	1%	−22%	−32%
60	2%	2%	−25%	−25%
72	−1%	2%	−27%	−32%
84	−11%	−7%	−32%	
96	−24%		−43%	
108	−26%		−43%	
120	−28%		−30%	

JMA forecasts indicates similar mean errors of JMA and ECMWF at lead times up to 120 h as well (not shown).

The two deteriorating cases with typhoon core and eyewall observations are subject of sensitivity studies at JMA. Such strong deteriorations may be avoided through either a modification of the quality control, dropsonde thinning, modified assigned observation errors within TCs, or an exclusion of core and eyewall observations. Currently, no horizontal thinning is applied for dropsondes at JMA, which may lead to an inadequate treatment of representativeness errors or observation error correlations. Also, applying the same observation error within and outside TCs may not sufficiently represent real observation errors. However, it must be noted that TC core and eyewall observations do not lead to such forecast deterioration on several other days besides the two cases mentioned above. This may also depend on the strength of the storm or the position of the dropsondes within the TC as discussed in Harnisch and Weissmann (2010).

The largest improvements of JMA DROP forecasts are seen in the early prerecurvature stage of Jangmi, but excluding the two deteriorating initial times mentioned above, some forecast improvements are seen throughout the whole period (Fig. 4).

JMA forecasts were initialized 4 times per day, but only the ones at 0000 and 1200 UTC are used for the comparison in Figs. 2 and 3 because other initial times are not available for ECMWF. Results where forecasts initialized at 0600 and 1800 UTC are included are very similar however (not shown).

b. NCEP GFS

The NCEP model shows a strong reduction of the typhoon position error in the DROP experiment analysis and also in forecasts at all lead times compared to the NODROP experiment (Table 3; Figs. 2d, 3b, and 5).

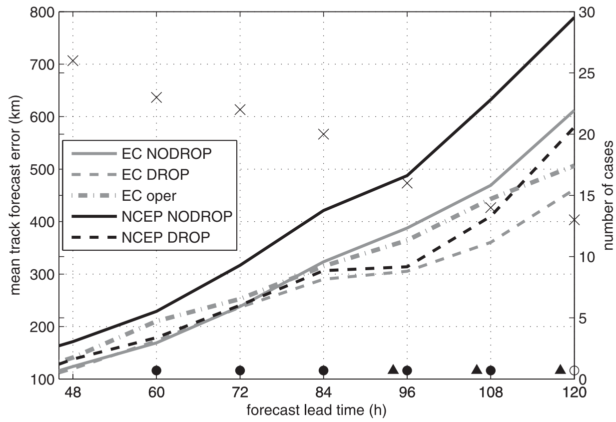

FIG. 5. Mean track forecast error of (gray) ECMWF and (black) NCEP DROP and NODROP experiments. A dashed-dotted line shows the mean error of the operational ECMWF track forecast with dropsonde timing errors. The mean errors were calculated by using all initial/verification times where forecasts were available from both models for initial times 0000 UTC 9 Sep–1200 UTC 18 Sep 2008 and 1200 UTC 24 Sep–0000 UTC 29 Sep 2008 and verification times before 1200 UTC 20 Sep and 1200 UTC 30 Sep 2008. The X symbols and the right y axis indicate the number of forecasts used for the calculation of the mean error. Empty (filled) markers indicate times where the mean differences are significant at a 90% (95%) confidence level. Circles and triangles stand for NCEP and ECMWF, respectively.

The differences of DROP and NODROP are significant at a 95% confidence level for most forecast lead times up to 120 h. Mean relative improvements are in the range of about 20%–40%, similar or even higher than improvements reported in previous studies (Aberson 2003, 2008; Wu et al. 2007b). However, NCEP is also starting from clearly higher mean NODROP track forecast errors than ECMWF and JMA. In the DROP run, NCEP track forecasts improve to mean errors lower than the ones of JMA. Up to lead times of 72 h, NCEP DROP forecasts are still worse than ECMWF NODROP, but they are in between ECMWF DROP and NODROP at longer lead times. The fact that NCEP DROP reaches such low track forecast errors despite lower resolution than the other two global models and 3D-Var illustrates the potential for improving typhoon forecasts with dropsondes.

Besides forecast improvements, the scatter diagram of NCEP DROP and NODROP forecasts (Fig. 2d) shows many improving and deteriorating cases compared to relatively few with a low difference between DROP and NODROP. This indicates that NCEP is more sensitive to the additional observations than the two other global models although the errors assigned to dropsonde wind and temperature observations are similar to the assigned errors of JMA and larger than the ones of ECMWF (Table 2). NCEP also shows a number of extreme outliers

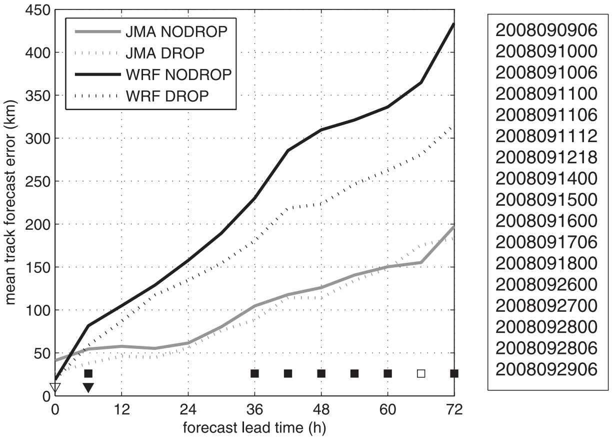

FIG. 6. Mean track forecast error of (gray) JMA and (black) WRF DROP and NODROP experiments. The track forecasts errors were averaged for 17 initial times displayed to the right of the figure and all verification times before 1200 UTC 20 Sep and 1200 UTC 30 Sep 2008, where forecasts from both models were available. Empty (filled) markers indicate times where the mean differences are significant at a 90% (95%) confidence level. Squares and triangles stand for WRF and JMA, respectively.

with larger track errors than any of the errors of ECMWF and JMA forecasts.

NCEP DROP forecasts show improvements compared to NODROP throughout nearly the whole period of the comparison (Fig. 4). The only exceptions are the last three initial times 1–2 days before landfall of Jangmi on Taiwan, which deteriorated (at 1200 UTC 26 September and afterward). Overall, the improvement in track forecasts from the added data during T-PARC to the NCEP modeling system is consistent with previous findings in the DOTSTAR program (Wu et al. 2007b).

c. KMA WRF

Forecasts by the Korean version of the WRF model are substantially improved through the assimilation of dropsondes (Fig. 6). The mean DROP and NODROP track forecast error is significantly different for forecast ranges of 36–72 h. Relative improvements of up to 32% are similar to the improvements of NCEP, but mean track errors of WRF DROP and NODROP forecasts are both above the other models. Mean WRF DROP track forecast errors for example, are 30–100 km larger than of the JMA NODROP forecast for lead times of 12–72 h. Only a very small number of WRF forecasts deteriorate whereas the majority of the forecasts improve and some forecasts even show very large improvements (Fig. 2b).

This demonstrates that the WRF 3D-Var system can gain additional information from adaptive dropsonde observations. However, the use of targeted dropsondes cannot fully compensate larger errors that are presumably due to a comparably coarse global model with 3D-Var used for the initial and boundary conditions.

d. ECMWF IFS

The scatter diagram of ECMWF DROP and NODROP typhoon track forecasts shows an improving tendency with dropsondes (Fig. 2c). These improvements are primarily due to improvements of forecasts with lead times of more than 72 h, whereas no mean improvement at shorter lead times is observed (Figs. 3a and 5). The differences of DROP and NODROP are significant at the initial time and at lead times of 96–120 h.

The improvements mainly occur in the early period of Sinlaku before recurvature (Fig. 4), whereas the influence in the later period of Sinlaku and Jangmi is rather neutral. Several of the cases with a low influence have a nearly perfect NODROP track forecasts. Errors in the other cases often appear to be connected to the land interaction with Taiwan or errors in the initial conditions upstream in midlatitudes, which cannot be reduced by additional dropsondes near the typhoons [see Harnisch and Weissmann (2010) for further discussion of individual cases]. Up to lead times of 2 days, the mean error of the ECMWF DROP and NODROP track forecasts are both as low as the mean error of the multimodel DROP forecasts (Fig. 3a). At longer lead times, the ECMWF DROP mean track error is slightly above the error of the multimodel forecast, but still the lowest mean error of all individual models. In the scatter diagram (Fig. 2c) it appears that ECMWF is the model with the lowest number of extreme improvements or degradations. This may be related to the high number of satellite observations used in the ECMWF assimilation system that presumably lead to a fairly accurate NODROP analysis, but may also pose a constraint for further improvements by a limited number of additional observations as dropsondes.

Figure 7 shows a comparison of the cycled DROP and the uncycled DROP_UnCy experiment. DROP was performed for the whole period and the dropsondes also modified the first guess of subsequent analysis cycles in addition to their direct influence on the respective analysis. DROP_UnCy was only performed for those analysis times with additional observations and used the first guess from the NODROP experiment. This comparison indicates that the largest part of the mean forecast improvement occurs through the modification of the first guess and the consequent accumulation of the influence of dropsondes in different cycles rather than through the direct influence of the dropsondes in the respective analysis.

Following the T-PARC field campaign, it was detected that dropsondes were partly assimilated with significant timing errors of several hours in the ECMWF

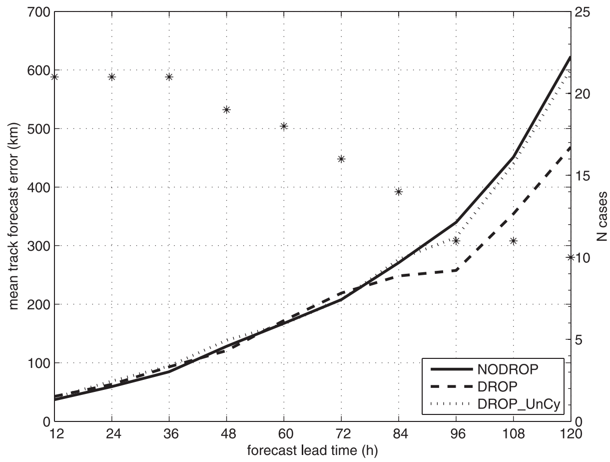

FIG. 7. Mean track forecast error of three ECMWF experiments: (solid line) NODROP, (dashed line) DROP, and (dotted line) DROP_UnCy experiment. The mean track errors were calculated by using all initial/verification times where additional observations were available in the periods 0000 UTC 9 Sep–1200 UTC 18 Sep 2008 and 1200 UTC 24 Sep–0000 UTC 29 Sep 2008 and verification times before 1200 UTC 20 Sep and 1200 UTC 30 Sep 2008. The asterisks and the right *y* axis indicate the number of forecasts used for the calculation of the mean track error.

system. These timing errors did not occur at other weather centers, but also affected dropsondes in previous years and other geographical areas. The timing errors were corrected before conducting the data denial experiments and a comparison of the operational track forecast error that includes these erroneous dropsondes has clearly higher errors than the DROP forecast with time-corrected dropsondes. Up to lead times of 3 days, the operational track forecast is even slightly worse than the ECMWF NODROP forecast and afterward the errors are in between the ECMWF DROP and NODROP experiments (Fig. 5). The operational forecasts were also performed with a slightly different model version, but a test with running the new model version with the erroneous dropsonde dataset for a limited number of cycles (not shown) confirmed that the main differences are due to the timing errors. As a consequence the operational dropsonde assimilation procedure at ECMWF was modified in June 2009 to avoid timing errors in the future.

4. Dropsonde influence on ECMWF midlatitude forecasts

The T-PARC dropsonde observations conducted for the track forecast of Sinlaku and Jangmi and during their extratropical transition near Japan also have a notable influence on midrange (days 4–5) ECMWF forecasts downstream in midlatitudes over the northern Pacific (Figs. 8a,c). Main improvements seem to result from the cycling (i.e., the modification of the first guess rather than the direct influence of the dropsondes on the analysis). Similar to the results for the ECMWF typhoon track forecasts, the DROP_UnCy run shows little difference to the NODROP experiment (Fig. 8b).

The largest improvements over the northern Pacific occur for forecasts initialized during 12–15 September (Fig. 9). Interestingly, only two flights were performed for Sinlaku during 13–15 September because the typhoon was located in Chinese airspace where flight permissions were lacking. The reduction of the forecast error downstream seems to be related to the cumulative improvements of the ECMWF track prediction of forecasts initialized during

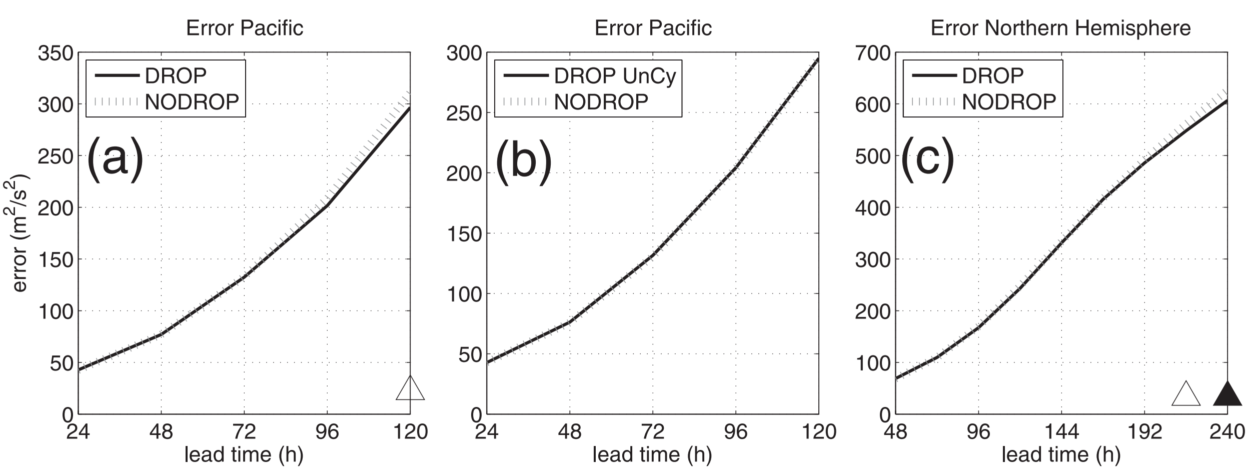

FIG. 8. (a) Mean error of 500-hPa geopotential in the ECMWF DROP and NODROP experiment in the period 9 Sep–1 Oct averaged over the Pacific 30°–65°N, 155°E–130°W. (b) As in (a), but for experiment NODROP and DROP_UnCy and only initial times where additional observations were available. (c) As in (a), but verified on the Northern Hemisphere north of 20° latitude. Empty (filled) triangles indicate times where the mean differences are significant at a 90% (95%) confidence level.

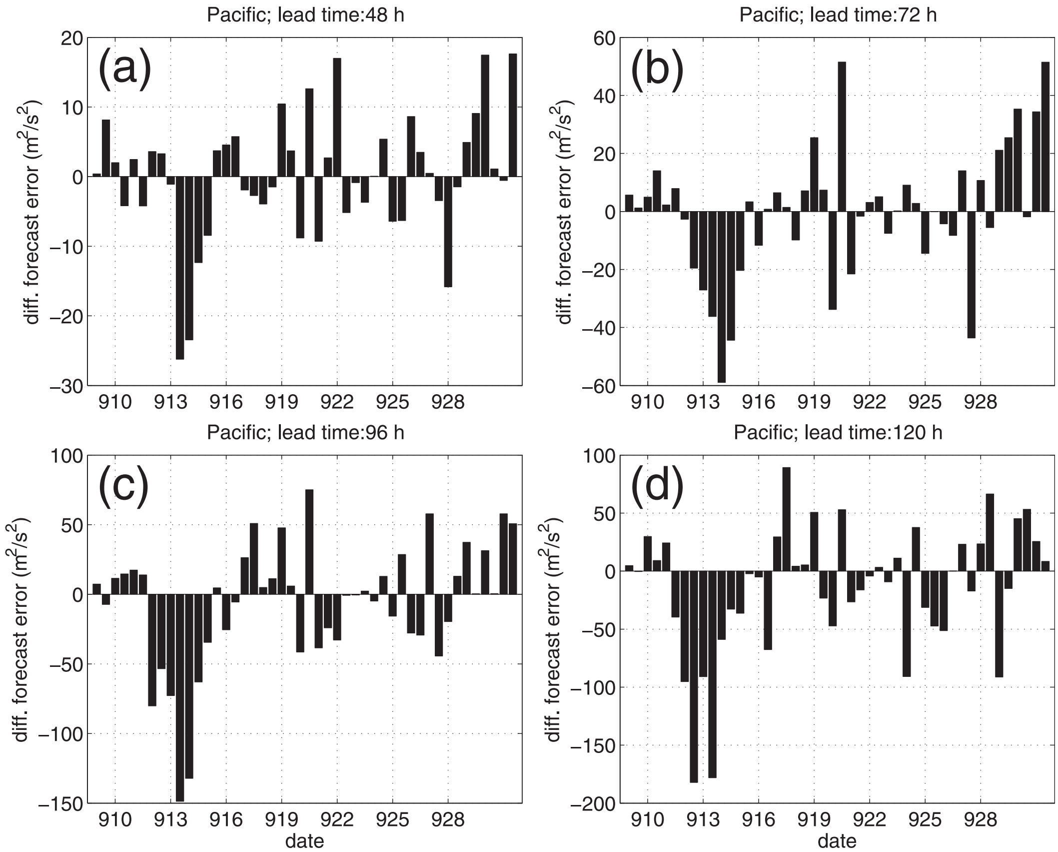

FIG. 9. Difference of 500-hPa geopotential forecast error averaged over the Pacific 30°–65°N, 155°E–130°W between the ECMWF DROP and NODROP experiment for lead times of (a) 48, (b) 72, (c) 96, and (d) 120 h in the time period 9 Sep–1 Oct. Negative values indicate lower errors in the DROP experiment.

9–13 September (Fig. 4). Thus, it seems the observations during 9–12 September improve the first-guess field for subsequent days; consequently, the errors of the typhoon track prediction are reduced and then the interaction with the midlatitude flow and the forecast over the Pacific improves. This hypothesis can explain why the largest improvements over the Pacific are seen at lead times of 4–5 days (Fig. 8a) although the verification region is centered only 60° longitude downstream of the observation region. In midlatitudes, the maximum signal could be expected to propagate into the verification region during a time frame of 2–3 days (Szunyogh et al. 2002).

Observations to investigate the extratropical transition of both systems and their interaction with the jet stream (i.e., observations during 17–21 September and after 29 September), overall have an neutral or even deteriorating influence on downstream midlatitude forecasts (Fig. 9) although several of these flights were targeted by singular vectors and ETKF sensitive area calculations optimized for verification areas over the northern Pacific. One explanation for this could be that observations during the early stages of typhoons have the largest potential to influence the analysis as they occur in data-sparse regions and a quickly evolving typhoon environment, whereas the midlatitude flow is already represented more accurately in the analysis, in particular downstream of fairly well-observed areas in China and Japan.

Additionally, the T-PARC observations lead to a reduction of the mean Northern Hemisphere forecast error at lead times of 8–10 days (Fig. 8c). Slight improvements also occur at lead times of less than 8 days. The largest improvements on the Northern Hemisphere occur for forecasts initialized in the early period of Sinlaku, but improvements are seen throughout the whole investigation period. About 60%–70% of forecasts on the Northern Hemisphere for lead times of 8–10 days improve and the difference is significant at a 95% confidence level for a lead time of 10 days.

The downstream influence is only analyzed in the ECMWF model because 1) the number of available JMA long-range forecasts is limited; 2) Atlantic dropsondes are also denied in the NCEP NODROP experiment, which may influence the global midlatitude circulation after some time; and 3) the domain of WRF is limited.

5. Discussion and conclusions

The study compares the influence of T-PARC dropsonde observations on the typhoon track forecast by three global models (ECMWF, JMA, and NCEP) and one limited-area model (KMA WRF). In addition, the influence of T-PARC observations on midlatitude forecasts downstream is investigated.

The results highlight that the influence of targeted dropsondes on typhoon track forecasts strongly depends on the modeling system. The two models using a 3D-Var system (NCEP and WRF) and lower resolution show strong improvements of 20%–40% with using dropsondes, but also comparably high errors without dropsondes. These improvements of NCEP and WRF are significant at a 95% confidence level for most lead times.

With dropsondes, the mean track errors of NCEP are overall lower than the ones of JMA and not much higher than the errors of ECMWF. This is a remarkable result given the coarser model resolution and lower amount of satellite data assimilated at NCEP. It seems that 4D-Var and the related more extensive assimilation of satellite data leads to clearly better results without dropsondes. However, the extensive use of satellite data may also limit the influence of additional observations and poses a stronger constraint for the analysis.

JMA results show an improvement of the majority of typhoon track forecasts with dropsondes, but two events with typhoon core and eyewall observations show a strong degradation, which levels out improvements of other forecasts. This indicates that there is potential for improvements by either modifying the treatment of core and eyewall observations in the analysis or discarding those observations. Recent studies showed a deteriorating influence of core and eyewall dropsonde observations on NCEP forecasts (Aberson 2008) and a neutral average influence on ECMWF forecasts (Harnisch and Weissmann 2010). As a consequence, observations within a radius of 111 km from the typhoon or hurricane center are currently not used in the NCEP analysis. How to use core and eyewall observations in global models appears to be one of the important questions for future research. Such observations certainly contain information that could partly be used to improve model analyses, but sharp gradients that are not fully resolved by global models can lead to very high differences between observations and the first-guess model, which challenge the data assimilation system. Furthermore, current data assimilation systems treat dropsondes as vertical profiles whereas dropsondes can drift significantly in high wind regimes (Aberson 2008). This may lead to additional errors.

ECMWF results show a mean typhoon track forecast improvement at lead times of 72–120 h, but the differences are only significant for lead times of 96–120 h. The ECMWF forecasts with and without dropsondes have the lowest mean track error of all individual models for lead times of up to 48 h and errors are similar to the multimodel forecast with dropsondes of ECMWF, GFS, and JMA. At longer lead times, the ECMWF forecasts with dropsondes show the best average performance of all individual models, but errors are 10–20 km above the multimodel forecast.

Despite large improvements through the assimilation of dropsondes, the errors of WRF track forecasts with dropsondes are still higher than the errors of JMA and ECMWF without dropsondes. These higher errors are presumably due to larger errors of the KMA GDAPS model that was used for initial and boundary conditions.

Typhoon intensity forecasts are excluded from the comparison because the resolution of the used models is insufficient to fully resolve tropical cyclones. In addition, both typhoons went over Taiwan and the correct intensity prediction is strongly tied to the correct prediction of the track rather than to initial condition intensity.

Mean improvements of midlatitude forecasts downstream over the Pacific also seem to be a result of improved typhoon track forecasts. In contrast, observations after typhoon recurvature and during extratropical transition of Sinlaku and Jangmi do not seem to lead to significant improvements in midlatitudes although many of these flights went into regions that were indicated to be sensitive by singular vectors or ETKF calculations optimized for the northern Pacific.

ECMWF typhoon track and midlatitude forecast improvements mainly result from the cycling (i.e., the modification of the first-guess field of subsequent cycles) rather than the direct influence of dropsondes. An uncycled experiment that assimilated dropsondes, but used the first guess from the experiment without dropsondes does not show significant typhoon track or midlatitude forecast improvements with dropsondes. The importance of cycling for typhoon track improvements in the ECMWF model is contradictory to the findings of Aberson (2010), but different models may behave differently in this respect.

Following the field campaign, it was discovered that dropsondes were partly assimilated with timing errors of several hours in the ECMWF system. The correction of these errors leads to a clear improvement of the typhoon track prediction. As a result, the operational dropsonde assimilation was corrected in June 2008. Since these timing errors also occurred in previous years and over the Atlantic, it may be expected that the ECMWF TC track forecast error in the future will be even lower than the record-setting performance reported by Fiorino (2008).

The study shows that tropical cyclone track forecast improvements can be gained with targeted dropsondes.

Given the tremendous damages caused by tropical cyclones, even small track forecast improvements can justify the expenses of airborne surveillance missions. However, the influence of assimilated dropsondes varies significantly in different modeling systems. It seems the potential for forecast improvements decreases in more complex data assimilation systems using a large amount of satellite observations. As the use of satellite observations in data assimilation is likely to increase even further in the future, alternative ways of observation targeting (e.g., the adaptive use of satellite observations or the use of airborne remote sensing observations with a larger data coverage) may have a larger potential for tropical cyclone track improvements in the longer term.

Acknowledgments. The dropsonde data were collected as part of T-PARC, TCS-08, and DOTSTAR. T-PARC and TCS-08 were sponsored by an international consortium from the United States (National Science Foundation, Office of Naval Research, Naval Research Laboratory, and U.S. Air Force), Germany (DLR, Forschungszentrum Karlsruhe), Japan (JMA), Korea (National Institute of Meteorological Research), and Canada (Environment Canada). The role of the National Center for Atmospheric Research's Earth Observing Laboratory (NCAR EOL) for campaign and data management is acknowledged. DOTSTAR is funded by the National Science Council of Taiwan, the United States Office of Naval Research, and the Central Weather Bureau Taiwan. The authors are thankful to all people who contributed to the successful realization of T-PARC, TCS-08, and DOTSTAR. The efforts of U. Schumann, D. Parsons, P. Harr, and S. Jones were crucial to bring the Falcon to T-PARC. We acknowledge the tireless efforts of the crews of the 53rd Weather Reconnaissance Squadron/Air Force Reserve Command, the NRL Scientific Development Squadron ONE (VXS-1), the DLR Falcon, and the DOTSTAR aircraft. The GFS was run with the help of the NCEP Environmental Modeling Center (EMC) who also provided computer resources. B. Kohler, P. Leighton, and B. Barry provided computer support at HRD.

The support of various people at ECMWF for the experiments and ECMWF computing resources are highly appreciated. In particular, the authors acknowledge the help of C. Cardinali, reprocessing of the dropsonde data by M. Dragosavac and Ioannis Mallas, as well as ECMWF efforts in providing the Data Targeting System software and Web site during T-PARC. M. Weissmann and F. Harnisch are part of the Deutsche Forschungsgemeinschaft (German Research Foundation) research unit Predictability and Dynamics of Weather Systems in the Atlantic-European Sector (PANDOWAE; more information is available online at http://www.pandowae.de/).

REFERENCES

Aberson, S. D., 2002: Two years of operational hurricane synoptic surveillance. *Wea. Forecasting,* **17,** 1101–1110.

——, 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.,* **131,** 1613–1628.

——, 2008: Large forecast degradations due to synoptic surveillance during the 2004 and 2005 hurricane seasons. *Mon. Wea. Rev.,* **136,** 3138–3150.

——, 2010: 10 years of hurricane synoptic surveillance (1997–2006). *Mon. Wea. Rev.,* **138,** 1536–1549.

Burpee, R. W., J. L. Franklin, S. J. Lord, R. E. Tuleya, and S. D. Aberson, 1996: The impact of Omega dropwindsondes on operational hurricane track forecast models. *Bull. Amer. Meteor. Soc.,* **77,** 925–933.

Chou, K.-H., and C.-C. Wu, 2008: Development of the typhoon initialization in a mesoscale model—Combination of the bogused vortex with the dropwindsonde data in DOTSTAR. *Mon. Wea. Rev.,* **136,** 865–879.

Fiorino, M., 2008: Record-setting performance of the ECMWF IFS in medium-range tropical cyclone track prediction. *ECMWF Newsletter,* No. 118, ECMWF, Reading, United Kingdom, 20–27.

Harnisch, F., and M. Weissmann, 2010: Sensitivity of typhoon forecasts to different subsets of targeted dropsonde observations. *Mon. Wea. Rev.,* **138,** 2664–2680.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting,* **24,** 1691–1705.

Kurihara, Y., M. A. Bender, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL Hurricane Prediction System. *Mon. Wea. Rev.,* **123,** 2791–2801.

Liu, Q., T. Marchok, H.-L. Pan, M. Bender, and S. Lord, 2000: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. NOAA Tech. Procedures Bull. 472, 7 pp. [Available online at http://www.nws.noaa.gov/om/tpb/472.htm.]

Mahfouf, J. F., and F. Rabier, 2000: The ECMWF operational implementation of four dimensional variational assimilation. Part II: Experimental results with improved physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1171–1190.

NOAA/NWS/NHC, cited 2009: 2008 National Hurricane Center forecast verification report. [Available online at http://www.nhc.noaa.gov/verification.]

Pu, Z., X. Li, C. S. Velden, S. D. Aberson, and W. T. Liu, 2008: The impact of aircraft dropsonde and satellite wind data on numerical simulations of two landfalling tropical storms during the Tropical Cloud Systems and Processes Experiment. *Wea. Forecasting,* **23,** 62–79.

Rabier, F., H. Järvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part 1: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

Richardson, D. S., and Coauthors, 2009: Verification statistics and evaluations of ECMWF forecasts in 2008-2009. ECMWF Tech. Memo 606, 45 pp.

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.,* **130,** 1144–1165.

Woollen, J. R., 1991: New NMC operational OI quality control. Preprints, *Ninth Conf. on Numerical Weather Prediction,* Denver, CO, Amer. Meteor. Soc., 24–27.

Wu, C. C., and Coauthors, 2005: Dropsonde observations for typhoon surveillance near the Taiwan region (DOTSTAR): An overview. *Bull. Amer. Meteor. Soc.,* **86,** 787–790.

——, J.-H. Chen, P.-H. Lin, and K.-S. Chou, 2007a: Targeted observations of tropical cyclones based on the adjoint-derived sensitivity steering vector. *J. Atmos. Sci.,* **64,** 2611–2626.

——, K. H. Chou, P. H. Lin, S. D. Aberson, M. S. Peng, and T. Nakazawa, 2007b: The impact of dropwindsonde data on typhoon track forecasts in DOTSTAR. *Wea. Forecasting,* **22,** 1157–1176.

——, and Coauthors, 2009: Intercomparison of targeted observation guidance for tropical cyclones in the northwestern Pacific. *Mon. Wea. Rev.,* **137,** 2471–2492.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.,* **130,** 2905–2916.

Yamaguchi, M., T. Iriguchi, T. Nakazawa, and C. Wu, 2009: An observing system experiment for Typhoon Conson (2004) using a singular vector method and DOTSTAR data. *Mon. Wea. Rev.,* **137,** 2801–2816.