The Impact of Dropwindsonde Data on GFDL Hurricane Model Forecasts Using Global Analyses

ROBERT E. TULEYA

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, New Jersey

STEPHEN J. LORD

National Centers for Environmental Prediction/NOAA, Washington, D.C.

(Manuscript received 9 April 1996, in final form 24 October 1996)

ABSTRACT

The National Centers for Environmental Prediction (NCEP) and the Hurricane Research Division (HRD) of NOAA have collaborated to postprocess Omega dropwindsonde (ODW) data into the NCEP operational global analysis system for a series of 14 cases of Atlantic hurricanes (or tropical storms) from 1982 to 1989. Objective analyses were constructed with and without ingested ODW data by the NCEP operational global system. These analyses were then used as initial conditions by the Geophysical Fluid Dynamics Laboratory (GFDL) high-resolution regional forecast model.

This series of 14 experiments with and without ODWs indicated the positive impacts of ODWs on track forecasts using the GFDL model. The mean forecast track improvement at various forecast periods ranged from 12% to 30% relative to control cases without ODWs: approximately the same magnitude as those of the NCEP global model and higher than those of the VICBAR barotropic model for the same 14 cases. Mean track errors were reduced by 12 km at 12 h, by ~50 km for 24–60 h, and by 127 km at 72 h (nine cases). Track improvements were realized with ODWs at ~75% of the verifying times for the entire 14-case ensemble.

With the improved analysis using ODWs, the GFDL model was able to forecast the interaction of Hurricane Floyd (1987) with an approaching midlatitude trough and the storm's associated movement from the western Caribbean north, then northeastward from the Gulf of Mexico into the Atlantic east of Florida. In addition, the GFDL model with ODWs accurately forecasted the rapid approach and landfall of Hurricane Hugo (1989) onto the U.S. mainland. An assessment of the differences between analyses indicates that the impact of ODWs can be attributable in part to differences of ~1 m s^{-1} in steering flow of the initial state.

In addition to track error, the skill of intensity prediction using the ODW dataset was also investigated. Results indicate a positive impact on intensity forecasts with ODW analyses. However, the overall skill relative to the National Hurricane Center statistical model SHIFOR is shown only after 2 or 3 days. It is speculated that with increased data coverage such as ODWs both track and intensity error can be further reduced provided that data sampling can be optimized and objective analysis techniques utilizing asynoptic data can be developed and improved.

1. Introduction

It has been speculated that the relatively sparse data surrounding tropical cyclones is a major factor limiting forecast skill of tropical storms and hurricanes. Burpee et al. (1984) hypothesized that enhancing wind and thermodynamic observations in the hurricane environment and core would improve the initial representation and subsequent model track forecasts of hurricanes. Since 1982, the Hurricane Research Division of the National Oceanic and Atmospheric Administration (NOAA) has conducted a series of 18 experiments with research aircraft to enhance observations through the deployment of Omega dropwindsondes (ODWs). A typical ODW flight pattern involved two aircraft with ~20 ODWs deployed by each aircraft every 150–200 km over a period of ~8 h. The ODW flight pattern for the case of Hurricane Floyd at 0000 UTC 11 October 1987 is shown in Fig. 1.

Early sensitivity studies, for example, Burpee et al. (1984), were generally inconclusive for a number of reasons. First, analysis schemes at that time were not able to ingest hurricane-scale observations in global models due to both resolution and analysis methodologies. In addition dynamic forecast models used specifically for hurricane track prediction were still rather crude and unable to utilize additional information effectively. Last, the statistical variability of hurricane track forecasts could mask any ODW impact for only

Corresponding author address: Dr. Robert E. Tuleya, GFDL/NOAA, Princeton University, Forrestal Campus, U.S. Route 1, P.O. Box 308, Princeton, NJ 08542.
E-mail: rt@gfdl.gov

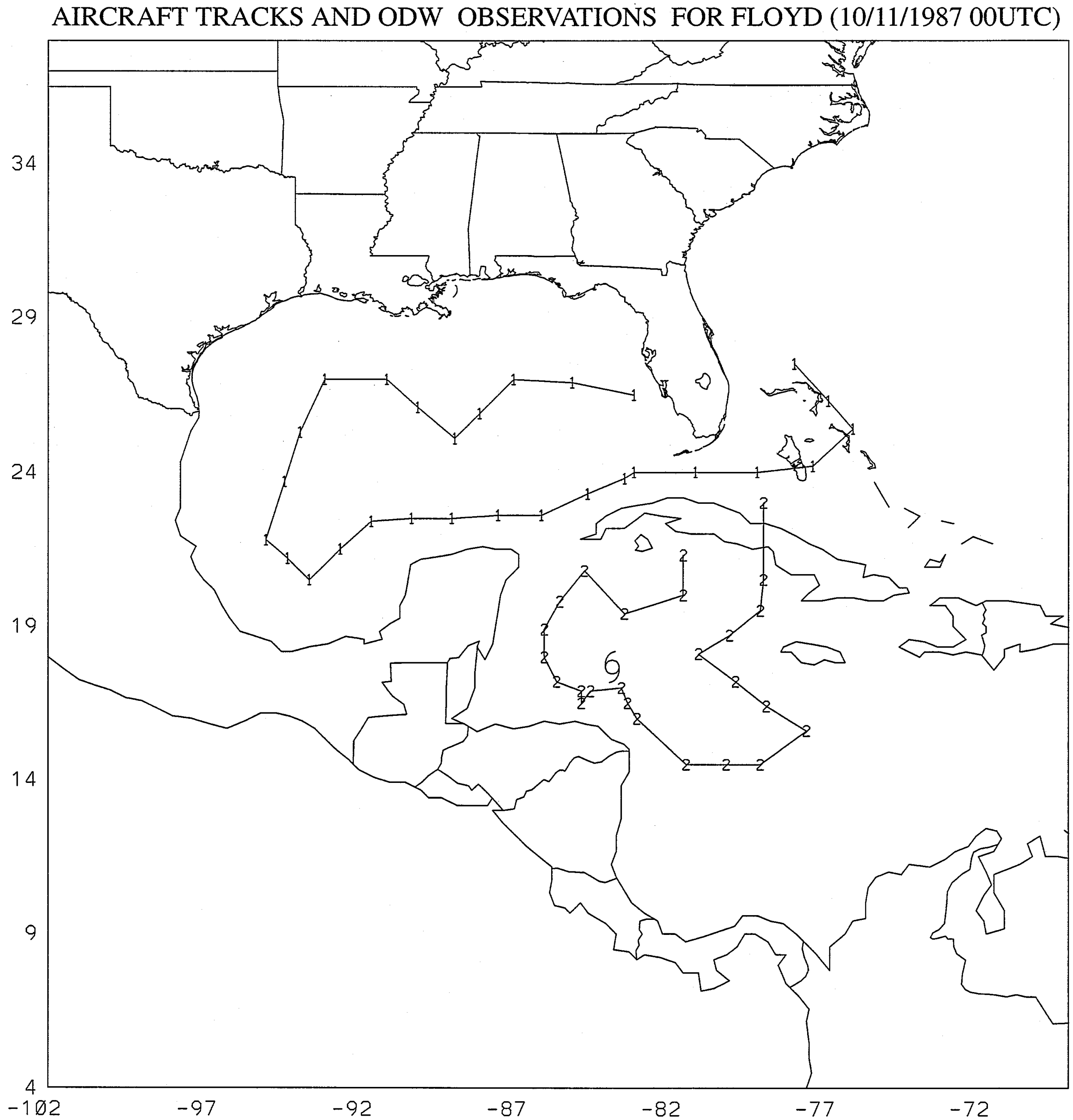

FIG. 1. The ODW flight patterns involving two aircraft are shown for the case of Hurricane Floyd (11 October 1987). The sites of ODW deployments are indicated by the numbers 1 and 2, with the observed 0000 UTC hurricane location depicted by the tropical storm symbol.

a small number of cases; in general, about a dozen cases were needed to determine any significant signal. In the last five years both improved analysis methods and model development, including increased horizontal resolution, combined with additional cases have motivated additional research on this topic.

Recently, dynamic models have demonstrated improvements when additional observations from hurricanes are used. Franklin and DeMaria (1992) showed statistically improved storm tracks from 14 cases using the VICBAR barotropic model (DeMaria et al. 1992). The VICBAR system uses its own analysis system with the National Centers for Environmental Prediction (NCEP) global model analysis (T126) as a first guess and as lateral boundary conditions for the forecast. Significant improvements of ~25% using ODWs in the NCEP global analysis–forecast system were demonstrated by Lord (1993). Also demonstrated in this study was the effectiveness of a synthetically specified vortex in representing a tropical cyclone vortex in a global

TABLE 1. Storms investigated.

Storm (year)	No.	Days	Initial time (UTC)
Debby (1982)	2	15, 16 Sep	0000
Josephine (1984)	3	10–12 Oct	0000
Gloria (1985)	1	25 Sep	0000
Emily (1987)	2	24, 25 Sep	0000
Floyd (1987)	2	11, 12 Oct	0000
Florence (1988)	1	9 Sep	0000
Hugo (1989)	2	20, 21 Sep	0000
Jerry (1989)	1	14 Oct	1200

analysis. Both ODW and the synthetic specification reduced track errors and their combination was even more effective. Results indicated that both ODWs and synthetic data increased the skill of track forecasts using the NCEP model with the synthetic data having the largest impact on 12–36-h forecasts and the ODWs impacting all forecasts out to 72 h.

Coincident with these data sensitivity studies, the Geophysical Fluid Dynamics Laboratory's (GFDL) MMM (multiply nested movable mesh model) system was being developed and applied to real data forecasting and operational use. Originally used as a research model (Kurihara and Tuleya 1974), the nested version was developed (Kurihara and Bender 1980) and first used in a real data case by Kurihara et al. (1990) in forecasts of Hurricane Gloria (1985). The GFDL MMM system was improved with vortex specification and other refinements (Kurihara et al. 1995) and became operational at NCEP in June 1995.

In a recent paper, Burpee et al. (1996) summarized track forecast improvements due to ODWs for a consensus of operational models over a 17-case ensemble. The purpose of this paper is to demonstrate, in more detail, the impact of ODWs on the forecast skill of the GFDL MMM system for a variety of initial conditions. The new GFDL MMM system may now be compared with other models for identical cases. In contrast to other recent work on impact of ODWs, intensity skill as well as track skill will be investigated. For direct comparison to the earlier studies of Lord (1993) and Franklin and DeMaria (1992), the 14-case ensemble will be utilized here. The addition of three recent cases [one of Hurricane Andrew (1992) and two of Hurricane Emily (1993)] does not appreciably change the results.

2. Use of the GFDL prediction system

a. GFDL model description

Kurihara et al. (1995) described the GFDL model, which is one of the operational hurricane forecast models at NCEP. It is a primitive equation model formulated in latitude, longitude, and sigma coordinates with 18 vertical levels (Table 1 of Kurihara et al. 1990). The grid system consists of two movable inner meshes of $1/6°$ and $1/3°$ resolution within a fixed $75° \times 75°$ outer domain of 1° resolution (Bender et al. 1993). The GFDL forecast system consists of three steps: 1) an interpolation of the NCEP global analysis onto the MMM regional domain, 2) a vortex specification to produce a realistic storm-scale initial condition near the hurricane, and 3) a model forecast from the above-obtained initial conditions. The GFDL model typically produces a 72-h forecast from the global NCEP analysis with the vortex specification derived from a storm message supplied by the National Hurricane Center. Based on semioperational and operational forecasts for the 1993–95 Atlantic and eastern Pacific seasons, the GFDL track forecasts appear superior to other NOAA operational products especially beyond one day (e.g., Aberson and DeMaria 1994).

b. NCEP data assimilation for the dropwindsonde cases

The NCEP Global Data Assimilation System consists of a global model, an analysis procedure, a quality control algorithm, and a synthetic data (wind bogusing) procedure. More details of the NCEP model can be found in Kanamitsu (1989) and Kanamitsu et al. (1991). The $\sim$1° global analysis utilizes spectral statistical interpolation (SSI, Parrish and Derber 1992) as configured for operational use in June 1991. Information on the synthetic data procedure and initial analyses used in the present paper can be found in Lord (1991, 1993). In the synthetic data procedure, wind pseudo-observations based on operationally estimated intensity and storm position are ingested at mandatory pressure levels from 1000 to 300 mb over an area 300 km on a side near the storm center. The synthetic data scheme was designed to give a more realistic initial state of storm structure and steering current to the global NCEP analysis and, therefore, to improve the subsequent NCEP forecast. The ODW data (wind, temperature, and moisture) were assimilated in a similar fashion to the current operational method (Burpee et al. 1996). The cases and analyses used in this study are from Lord (1993) and correspond to his CNTL (control, no ODW or synthetic data), ODW (ODWs only), SYN (synthetic data, no ODWs), and ODW/SYN (synthetic data ODWs) for each of the 14 cases listed in Table 1. These cases are shown in Fig. 2 and include storms in the western North Atlantic or Gulf of Mexico with four making landfall in the continental United States.

c. GFDL model results without dropwindsondes

To examine the impact of ODWs using the GFDL model, one may use the synthetic or control runs of the global analyses. In addition the GFDL modeling system can be used with and without its specified vortex procedure. It is interesting to examine first the sensitivity of the GFDL model to these variations in initial con-

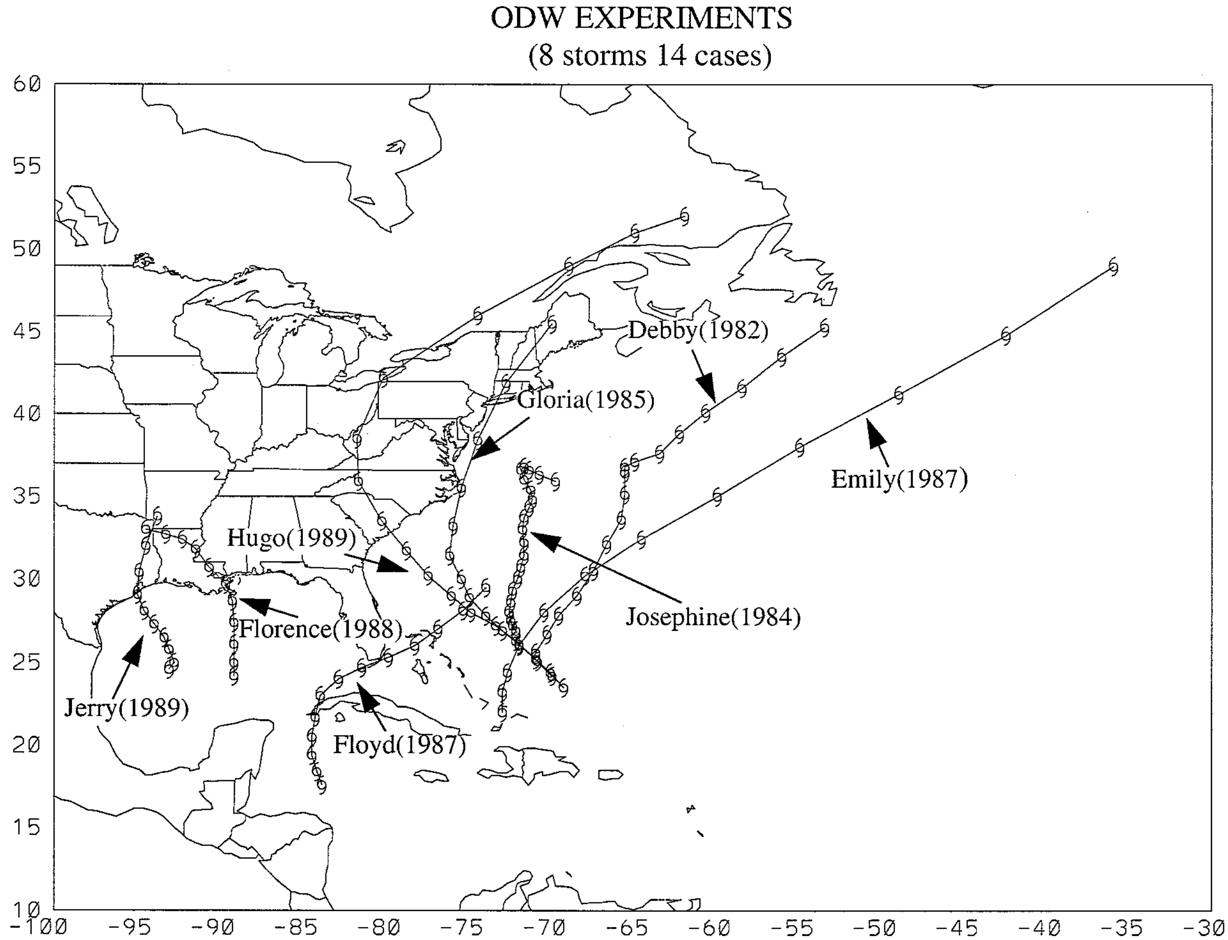

FIG. 2. The tracks and positions of the eight storms studied. Five storms were investigated at several initial times for a case total of 14. The positions are indicated every 6 h by tropical storm symbols.

ditions. The GFDL track errors are shown in Fig. 3 for the four possible cases: 1) control case, experiment GCNTL (no ODW or synthetic vortex); 2) NCEP synthetic vortex in the global analysis, experiment GSYN; 3) control case with the GFDL-specified vortex, experiment GSPE; and 4) NCEP synthetic vortex and GFDL-specified vortex, experiment GFDL. The results are shown relative to CLIPER, the standard climatology and persistence model. The liability of running the GFDL MMM with an ill-defined vortex in the global analysis is clearly shown in the control experiment suite with no bogus vortex specification. Some cases exhibit skill [e.g., Hurricane Gloria (1985), Kurihara et al. (1990)] when the global analysis is able to resolve a well-defined tropical system in the appropriate location without adding any idealized storm data. Overall for the total of these 14 control cases, however, the control experiments GCNTL show no relative skill until after 2 days. On the other hand, the implementation of either idealized vortex (NCEP synthetic, GFDL specified, or both) demonstrated considerable skill in track forecasting. This can be seen in Fig. 3 in which skill relative to CLIPER is shown from 12 h onward. If the synthetic NCEP vortex is used together with the GFDL-specified vortex (i.e., the case for the present operational GFDL system in which the scheme first filters out the global NCEP synthetic system and replaces it with the GFDL specified vortex), a high degree of skill is achieved (generally 40% better than CLIPER after 1 day), similar in performance to when either the NCEP synthetic or the GFDL-specified vortex is used alone. One major reason to install the NCEP synthetic system operationally was to have a reasonable, trackable system from the start. It appears that the GFDL model is capable of credible track forecasts using the NCEP synthetic vortex without utilizing its own vortex specification method. Notice that in these cases, forecasts with the NCEP synthetic vortex alone yield slightly better results than when the GFDL-specified vortex was used, but there is no statistical difference in track error among GSYN, GSPE, and GFDL. However, in order to track weak systems in operational mode, the GFDL-specified vortex was necessary in the day-to-day implementation of the GFDL system (Bender et al. 1993).

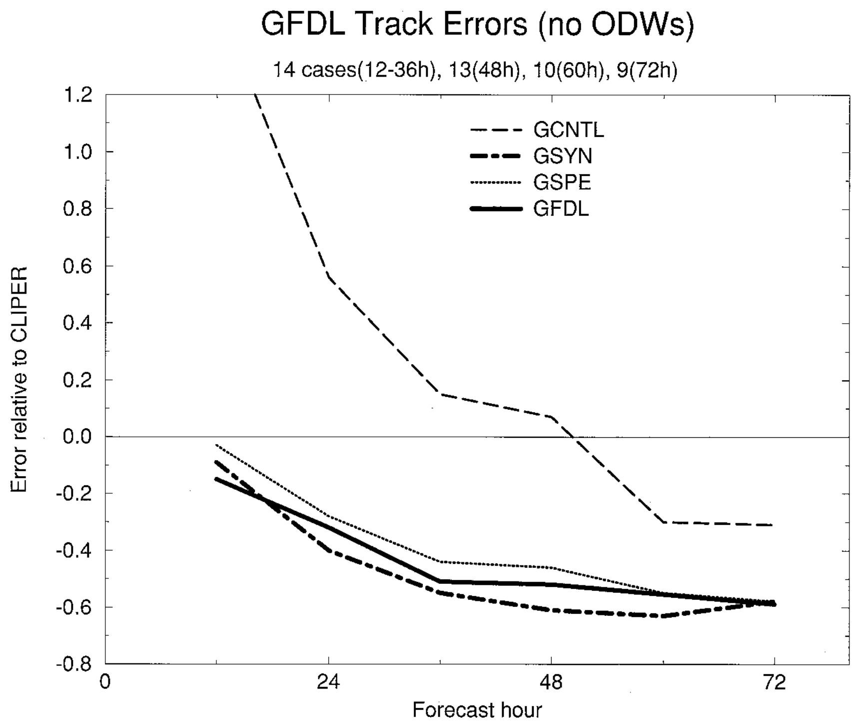

FIG. 3. Mean track errors for four modes of the GFDL model relative to CLIPER for the 14 cases studied. Increased skill is indicated by increasingly negative values. GCNTL is the suite of experiments initiated from analyses with no NCEP synthetic vortex. GSPE are those initiated from analyses with no NCEP synthetic vortex but the GFDL vortex specification scheme implemented. GSYN is the suite of experiments initiated from analyses with the NCEP synthetic vortex alone. GFDL is the suite of experiments initiated using both the NCEP synthetic vortex and the GFDL vortex specification scheme. The initial conditions used are those without ODWs.

The GFDL MMM is the first operational dynamical model to forecast intensity. Its feasibility was demonstrated in Bender et al. (1993), but consistent predictive skill relative to SHIFOR for a large case size has yet to be demonstrated. Despite the lack of skill, Fig. 4 shows the advantage in the GFDL-specified vortex system in reducing the intensity error in the first 2 days of the forecast. The GFDL-specified system forecasts were about the same for the case with or without the NCEP synthetic vortex (i.e., GFDL and GSPE). On the other hand the suite of experiments using the NCEP synthetic vortex alone (GSYN) shows improvement over the control (GCNTL), which lacks skill for the entire 72-h forecast period. Both GCNTL and GSYN are limited by the coarse resolution of the global assimilation system; GCNTL, without the synthetic vortex, also suffers from data sparsity near the storm center. Skill relative to SHIFOR, the standard climatology and persistence intensity technique, is not reached until 36 h in GSYN but is reached at ~30 h in GFDL. In general, significant skill relative to SHIFOR is achieved in GSYN and GFDL at and after these times for this limited case study. In the first day or so, the level of forecast intensity skill is not large in either GFDL system as one sees evidence of an initial adjustment bias in Fig. 4. Values of 400% worse and 200% worse than SHIFOR are shown at 12 h, decreasing to 200% and 50% worse at 24 h for the NCEP synthetic and the GFDL-specified series, respectively. The GFDL system overintensifies weak systems and underpredicts intense systems regardless of the forecast hour (Fig. 5). Intensity errors range from as high as 25 m s^{-1} overprediction to as low as 23 m s^{-1} underprediction. Note that there are only four cases of overprediction and four cases of underprediction for observed wind speeds greater than and less than 25 m s^{-1}, respectively. This bias remains if one examines the intensity errors relative to the initial intensity as well. Research and model development is continuing on removing or alleviating this initial bias, which is also evident from the operational GFDL forecast system (Kurihara et al. 1995).

3. Results using the GFDL prediction system

a. Impact of dropwindsondes on the GFDL forecast tracks

In this section, the impact of ODWs will be shown using the GFDL model for four suites of initial conditions, which are divided into two pairs. The first pair uses the NCEP model initial conditions with synthetic data (without and with ODWs), and the second pair is

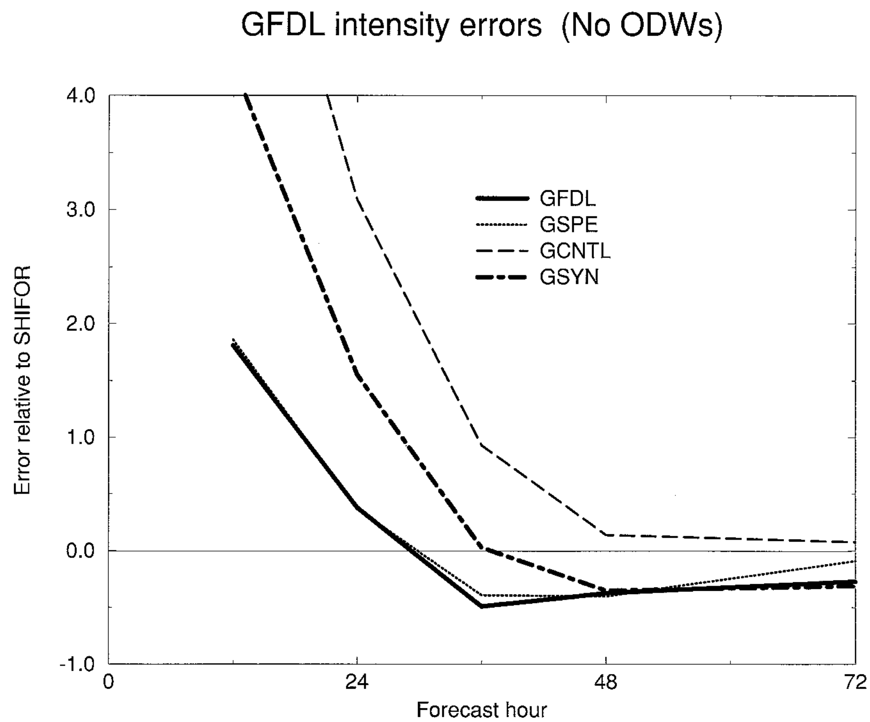

FIG. 4. Mean intensity errors for four modes of the GFDL model relative to SHIFOR for the 14 cases studied. Similar to the standard CLIPER diagram (Fig. 3), positive skill has negative values. The initial conditions used are those without ODWs.

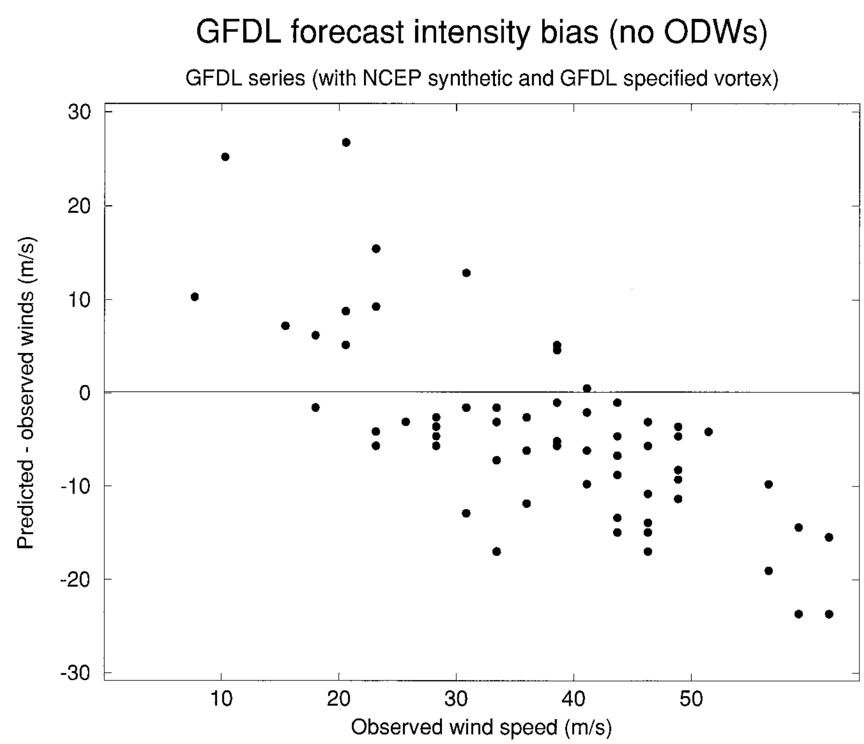

FIG. 5. A plot of the forecast intensity bias (predicted − observed) versus the observed wind speed for the GFDL suite of experiments for the 14 cases studied. Data are plotted for 12-, 24-, 36-, 48-, and 72-h forecast times.

TABLE 2. List of experiments.

Experiment	Initial analysis/ forecasts	Assimilation of ODW data
GSYN-ND	NCEP synthetic	No ODWs
GSYN-WD	NCEP synthetic	ODWs
GFDL-ND	GDFL specified	No ODWs
GFDL-WD	GFDL specified	ODWs

the same as the first except that the GFDL specified vortex is used as well (Table 2). The two suites without ODWs correspond to the GSYN and GFDL experiments discussed earlier. The mean track errors for each 12-h forecast period are smaller with ODWs than without ODWs (Fig. 6). The mean improvement ranges from 12 km at 12 h, ~50 km for 24–60 h, to 127 km for the nine cases at 72 h. Although the relative errors of GSYN and GFDL change with forecast period, both suites of experiments indicate similar improvement with ODWs despite the different method of vortex specification. Since the track forecasts are better, environmental wind and basic steering current improvements apparently are retained with both bogusing techniques.

One can also examine the case-to-case variability in addition to the mean results. This is important for testing the robustness of the ODW impact on track forecast errors whose case-by-case variation is known to be large. A majority of cases indicate improvement using ODWs for all forecast periods. Improvements ranged from 9 of 14 cases at 12 h to all cases at 72 h, with improvement overall at ~77% of the verifying times. The mean track error for all forecast periods showed a considerable difference at the 95% significance level except at 12 and 48 h, which indicated improvement at the 90% significance level. Figure 7 shows a scatter diagram of forecast errors for experimental suites GFDL and GSYN for the 14 cases at all forecast periods. Most of the cases lie above the diagonal line, indicating improvement using ODWs.

While the mean errors show an overall improvement with ODWs, there is considerable case-to-case variation (Fig. 7). A comparison of improvement using GFDL and GSYN shows positive correlation with ~2/3 of the verifying times of the 14-case ensemble indicating the same sign of improvement. For both vortex initialization methods the most improvement was shown for the case of the 72-h forecast hour of Hurricane Floyd on 11 October with improvements of 504 and 457 km for GSYN-WD and GFDL-WD, respectively. On the other hand the worst responses were for GSYN-WD for Hurricane Florence at 48 h (158-km degradation) and GFDL-WD at 48 h for Hurricane Floyd on 12 October (261-km degradation). Note that all forecasts with ODWs that were worse had errors of less than 300 km

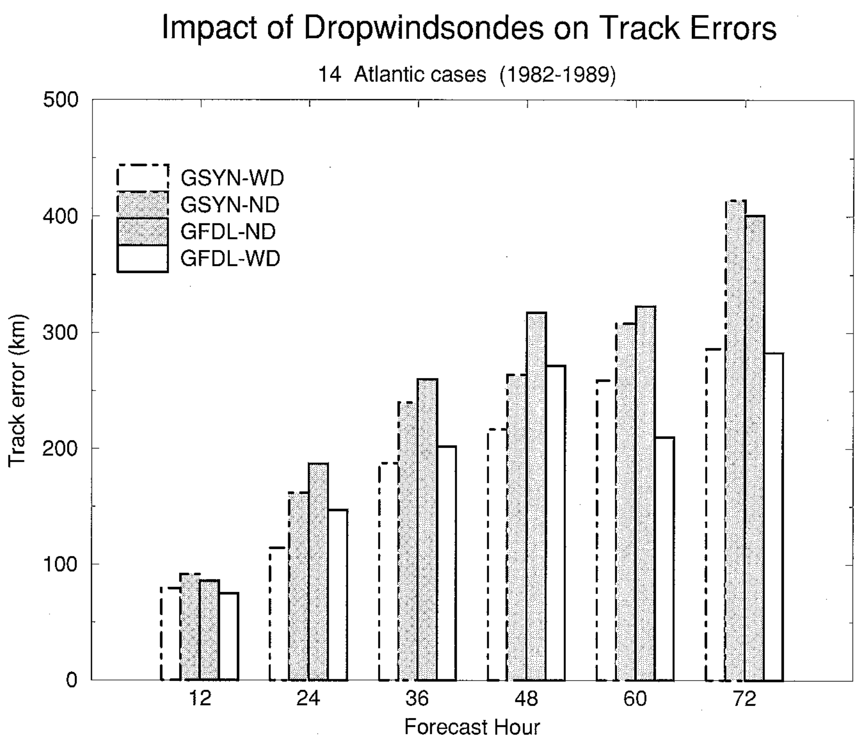

FIG. 6. Mean track errors for the GFDL model with and without ODWs for the 14 cases studied. Shaded values at each time period indicate mean values without ODWs, while unshaded values indicate mean values with ODWs. The two left values enclosed by dashed lines are those using the NCEP synthetic vortex alone as initial conditions, while the two right values enclosed by solid lines are those using additionally the GFDL vortex specification method.

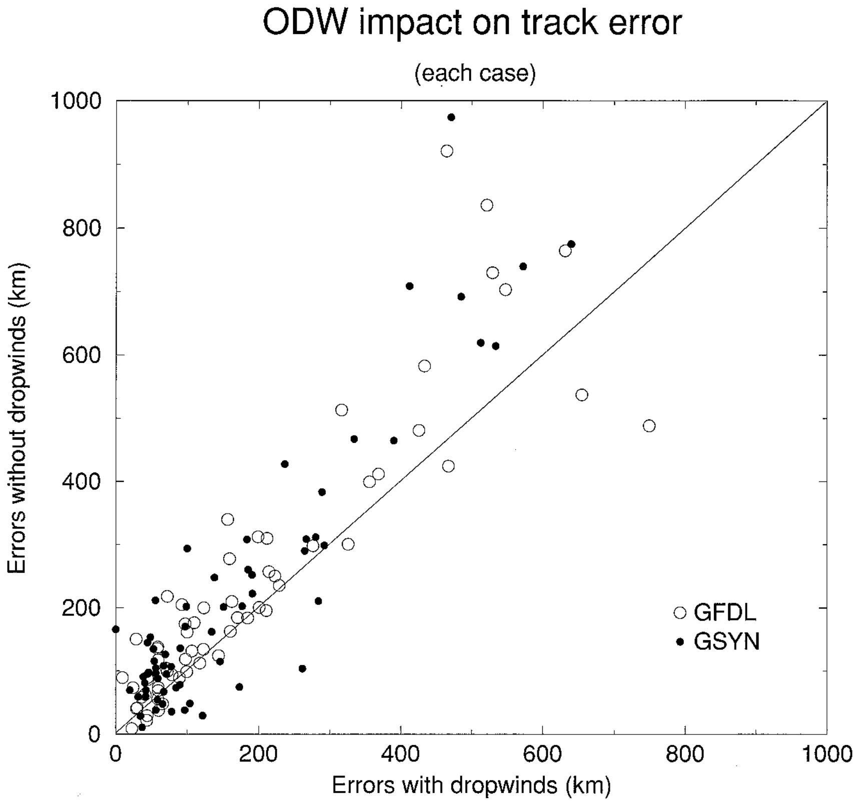

FIG. 7. Plot of individual GFDL model track forecast errors using analyses without vs with ODWs at 12, 24, 36, 48, and 72 h. Open circles and closed dots refer to experiments using NCEP synthetic vortex with and without the GFDL vortex specification system, respectively.

for GSYN and 750 km for GFDL. Overall only 15 and 13 of the 63 total verifying times were degraded for GSYN-WD and GFDL-WD systems, respectively.

b. Comparison of improvement with previous studies

It is useful to compare previous results of ODW impact with the current study. Table 3 lists the forecast track error for a homogeneous sample of the 14 synoptic cases without ODWs for the NCEP model with the synthetic vortex, VICBAR, and the GFDL model experiments, GFDL and GSYN, with the NCEP synthetic vortex implemented (with and without the GFDL-specified vortex). Consistent with error statistics from larger sample size, the GFDL model suites display substantially superior forecasts beyond 12 h. It is therefore more instructive to examine the relative improvement that ODWs have on each model by comparing the fractional improvement in the mean forecast error (Fig. 8). The mean improvement using the two GFDL model suites is >10% at 12 h and 15%–30% thereafter. The NCEP model improves about as much as the GSYN model suite for the first 36 h, increasing from 10% to >20% beyond 24 h. The model suite GFDL improves somewhat less initially with values approaching 20%. Apparently the GFDL vortex specification system has some influence on the impact of the ODW data. This technique may dilute the ODW data signal near the storm center. It is left for further study to evaluate the exact impact of different filtering and bogusing strategies on limiting the usefulness of additional high-density data such as ODWs. Although both indicating improvement beyond 36 h, the NCEP and GFDL model suites behave quite differently. One main difference between the NCEP and both GFDL experiments is that at the 72-h forecast period where the GFDL forecasts continue to improve at the same or even greater degree while the NCEP global model improvement drops off below 10%. This is due primarily to the difficulty the T126 global model has in sustaining a viable vortex past 60 h.

TABLE 3. Track errors (km) for a homogeneous suite of cases without ODWs for depression, tropical storm, and hurricane stages. The number of cases for each time period is in parentheses.

Track model	GFDL	GSYN	NCEP	VICBAR
12 h (14)	86	92	105	88
24 h (14)	187	162	230	196
36 h (14)	260	240	336	338
48 h (13)	318	264	403	537
72 h (8)	429	437	494	976

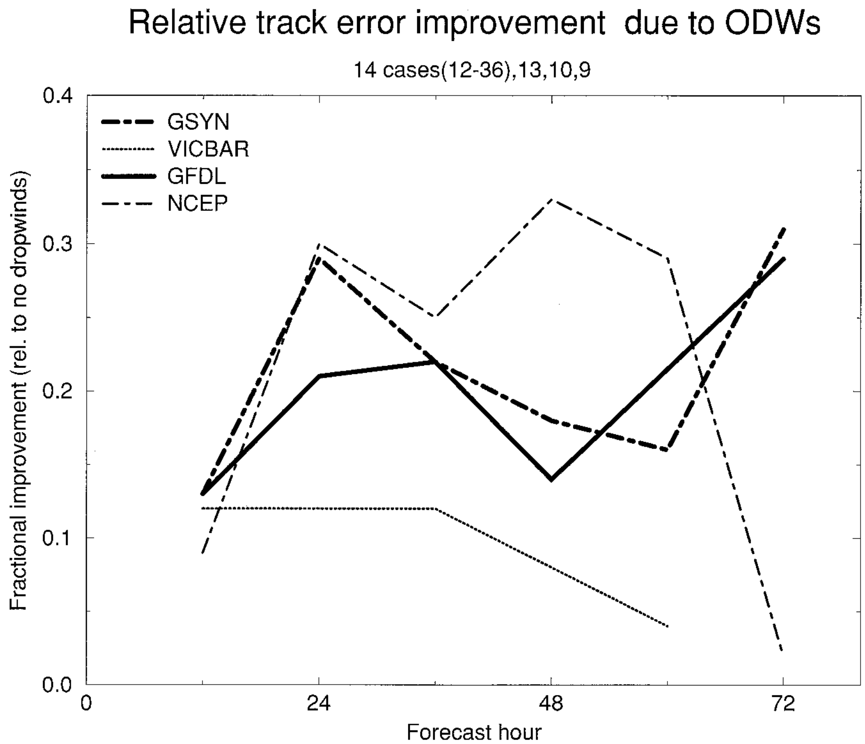

FIG. 8. The relative track error improvement due to ODWs for the GFDL model, the NCEP spectral model (T126 with synthetic vortex), and the VICBAR model. Fraction improvement is calculated as the difference in the track errors between model suites with and without ODWs normalized by the track error without ODWs. GFDL and GSYN refer to experiments using the NCEP synthetic vortex with and without the GFDL vortex specification system, respectively.

One can also examine the individual case improvement between the NCEP and the two GFDL suites. The GSYN and the NCEP global runs are highly correlated with ~44 of 60 verifying times of the 14-case ensemble indicating the same sign of improvement. This is even more evident for the 12- and 24-h periods when 23 of 28 verifying times show trends of the same sign. As expected, the correlation between the GFDL and the NCEP global series is not as strong due to the GFDL-specified vortex changing the initial conditions. During later periods the GSYN and GFDL series of experiments show a more similar response to ODWs than that of the NCEP response. This suggests that the results become more model sensitive and less dependent on details of the initial conditions for later times in the forecast period.

In addition, a positive relationship was found between the magnitude of the 850–200-mb vertical wind shear and the magnitude of the ODW track improvements in the NCEP, GSYN, and GFDL runs with corresponding correlation coefficients of 0.44, 0.40, and 0.10, respectively. The correlations were calculated at 12, 24, 36, 48, and 72 h using the shear calculated for a $10° \times 10°$ area following the storm in the GSYN forecasts. Large ODW improvements were found in highly sheared, rapidly moving storms such as Hurricane Emily for the NCEP, GSYN, and GFDL runs. In contrast, the VICBAR series shows a relatively small fractional improvement of ~10% to 36 h, decreasing after that, with little correlation (correlation coefficient of 0.04) found between ODW improvement and vertical shear. This indicates that both the barotropic and baroclinic structures are improved by ODWs and that this is subsequently manifested in more improved track forecasts in the baroclinic models. The VICBAR model can improve only through barotropic changes induced by ODWs and its forecast skill is limited by two-dimensional assumptions beyond ~36 h. This can be further seen in individual case comparisons (J. L. Franklin 1995, personal communication), which indicate 98% of the ODW improvements are confined to within 100 km in VICBAR, whereas in the NCEP and GFDL models greater than 20% of occurrences had improvement greater than 100 km. Additionally in a manner similar to the GFDL suite, the VICBAR system may not fully utilize the ODW data near the storm center. Nevertheless ~45 of 60 occurrences indicate coincident improvement of the VICBAR and the GSYN experiments. Overall it appears that the NCEP, VICBAR, and GFDL suites were able to benefit from the additional ODW data and that for most cases the relative improvement is at least qualitatively similar. At the later forecast periods, however, the GFDL model shows absolute skill levels above those of either VICBAR and the NCEP model experiments, and thus should be more able to retain the initial information from the ODWs. This supposition is confirmed by the fact that the GFDL model suites show ODWs having a significantly large positive influence into day three.

c. Comparison of individual cases with and without dropwindsondes

Section 3a indicated that tracks of the 14 cases were improved for ~70% of the forecast verification times when using ODWs and the GFDL model. The initial analyses with and without ODWs and their corresponding differences are now examined for particular cases. For this purpose the initial analyses and forecast tracks will be analyzed using the GSYN series. In this study the ODW impact on storm track was traced directly to the initial analysis, although clearly as section 3b indicates, model forecast skill has a strong bearing on the evolution of this impact.

The case for Hurricane Floyd (0000 UTC 11 October 1987) had large forecast errors without ODWs but showed considerable improvement due to the well-sampled ODW distribution (Fig 1). Figure 9 shows the observed track of Hurricane Floyd and the forecast with and without ODWs. The ODW case better forecasts the initial north-northwest movement and subsequent recurvature northeastward. Figure 10 (top) displays both the initial sea level pressure and wind analysis near 500 mb. At this level differences of ~5 m s^{-1} were typical between wind analyses with and without ODWs for this and all other cases. The movement of Hurricane Floyd was affected by a southward extension of the midlatitude upper-level trough above 500 mb and related weakness in the subtropical ridge. In order to analyze this case further, the analyses with and without ODWs were then spatially filtered to remove the disturbance field (see Kurihara et al. 1993). The vertical mean difference between these two filtered analyses (Fig. 10, bottom) indicates near-storm differences of ~0.4 m s^{-1} with larger differences to the north in the direction of storm movement. The analysis with ODWs shows an enhanced weakness in the subtropical ridge with the increased westerlies being consistent with the more initial eastward and subsequently more rapid northeastward movement of the storm.

The case for Hurricane Hugo (0000 UTC 20 September 1989) was analyzed with and without ODWs in a similar manner. In contrast to the upper-level trough interaction in Hurricane Floyd, the track of Hurricane Hugo was dictated by a deep northwestward flow on the southwest side of both surface and midlevel highs. Figure 11 shows the observed and two forecast tracks with and without ODWs, and Fig. 12 (top) displays the synoptic situation. Notice the more accurate initial northwestward movement and subsequent more rapid movement of Hurricane Hugo toward the coast in the experiment with ODWs. Although ~6 h too slow in making landfall, the ODW experiment predicted landfall ~4 h earlier than the no-ODW experiment. The vertical mean difference field (Fig. 12, bottom) clearly showed an enhanced flow around the Atlantic high with the ODW data steering the storm more rapidly northward toward the coast. Near-storm differences of ~0.5 m s^{-1} are observed with larger values in the path of the storm again being consistent with the differences in forecast storm track between ODW and no-ODW experiments.

Similar results were also noted for other cases including Hurricane Debby (0000 UTC 16 September 1982), which took a jog in response to an upper-level trough. This small-scale movement (Fig. 4 of Burpee et al. 1996) was forecast particularly well by the GFDL model using the analysis with ODWs. On the other hand, the worse case of GSYN ODW degradation was examined: that of Hurricane Florence, which moved northward from the southern gulf making landfall in the Louisiana delta region. As in the case of Hurricane Floyd, Hurricane Florence was influenced by an upper-level trough in the westerlies. With ODWs, this trough was not as sharp leading to a more relatively westerly mean flow across the path of the storm. This led to an erroneous eastward movement with ODWs. A key unanswered question is whether such degradations can be traced to errors in the data, biases in either analysis scheme or model, or are due to the random fluctuations of predictability.

Steering currents may be defined in a gross sense (e.g., Chan and Gray 1982; Dong and Neuman 1986; Carr and Elsberry 1990; Franklin et al. 1996) as a vertically averaged flow along the perimeter of the storm. This was done for an azimuthal ring, vertically integrated throughout the atmosphere, 600 km from the center of each of the 14 storms at the initial time for the analyses with and without ODWs. In the case of Hurricane Floyd on 11 October the ring analysis had a steering direction 70° north of west at 2.3 m s^{-1} for the analysis without ODWs. The ODW analysis had a steering flow more toward the north, 83° north of west at 2.3 m s^{-1}. This again was consistent with the model track evolution in which the storm in experiment GSYN-WD moved more eastward than that in experiment GSYN-ND. In the case of Hurricane Hugo on 20 September the ring analysis had a steering magnitude of 3.4 m s^{-1} at ~73° north of west for the analysis without ODWs compared to 4.5 m s^{-1} in the analysis with ODWs. This again was consistent with the model track evolution in which the storm in experiment GSYN-WD moved northwestward more rapidly than that in experiment GSYN-ND. A similar relationship existed for most of the 14 cases with the steering current differences reflecting the ultimate difference in the model storm tracks. The mean magnitude of the difference in the steering current between the two anal-

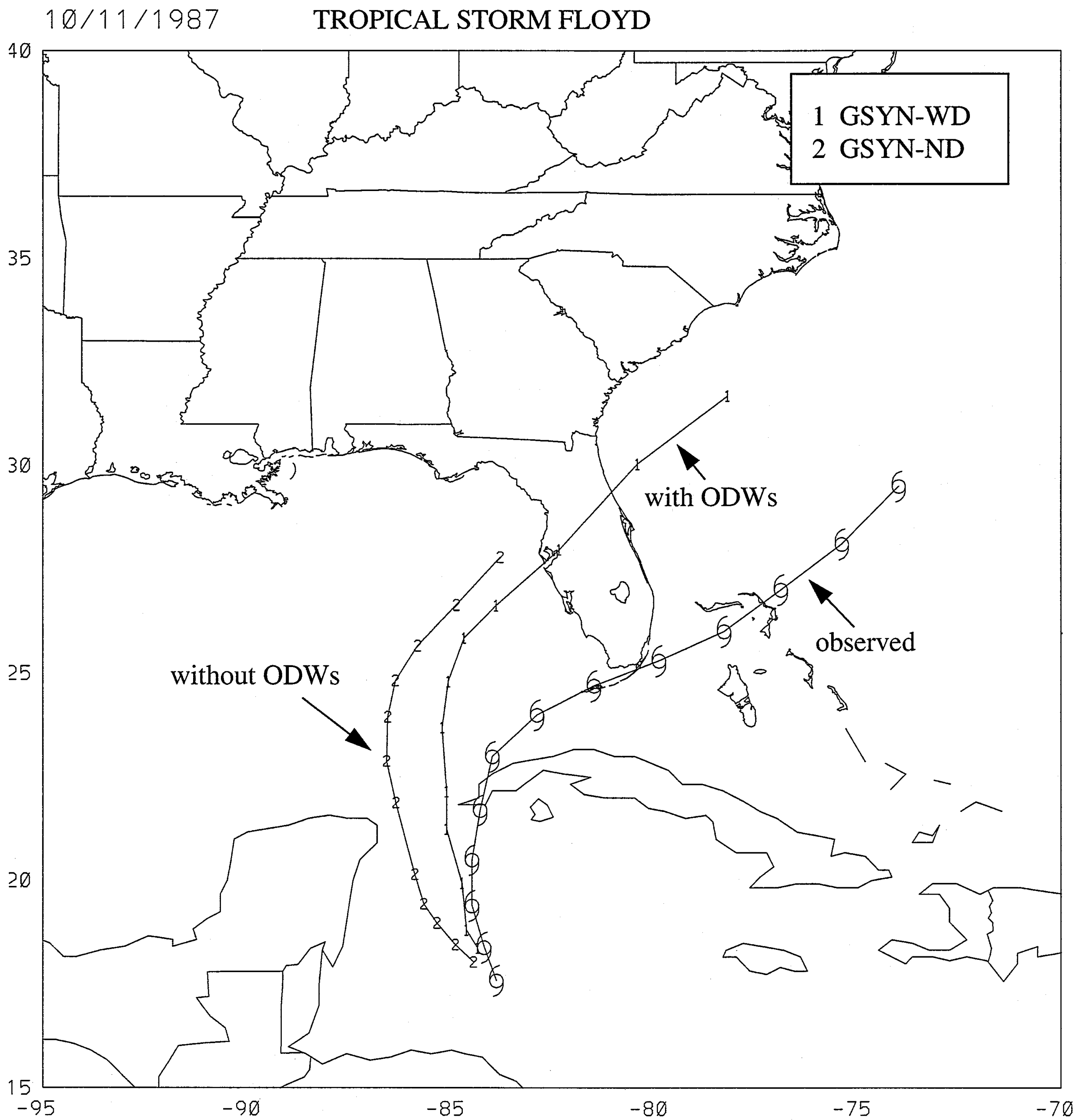

FIG. 9. Tracks of GFDL 3-day forecasts starting from 0000 UTC 11 October 1987 for Hurricane Floyd with and without ODWs ingested in the analyses. Results are for experimental series GSYN without the GFDL vortex specification system. The best track positions, including that at the initial time, are plotted with a tropical storm symbol, whereas the model forecasts are numbered and start at 6 h. Best track and model forecast positions are plotted every 6 h.

yses was ~0.6 m s^{-1} for those cases ranging from a difference of 0.1 m s^{-1} in Hurricane Jerry to 1.4 m s^{-1} for Hurricane Emily (24 September 1987). In summary, forecast track differences between analyses with and without ODWs can be explained qualitatively either by differences in steering currents defined in a traditional fashion (ring average) or by differences in the filtered fields of the vertical mean current over the storm. Vertically averaged wind changes due to ODWs are very subtle, typically ~0.5 m s^{-1}, but a sustained difference of this magnitude can amount to more than 40 km in forecast track over 1 day. Given that errors in forecast wind fields grow rapidly with time (Lorenz 1965), mean improvements considerably larger than 40 km day^{-1} can be expected for forecasts of several days or more.

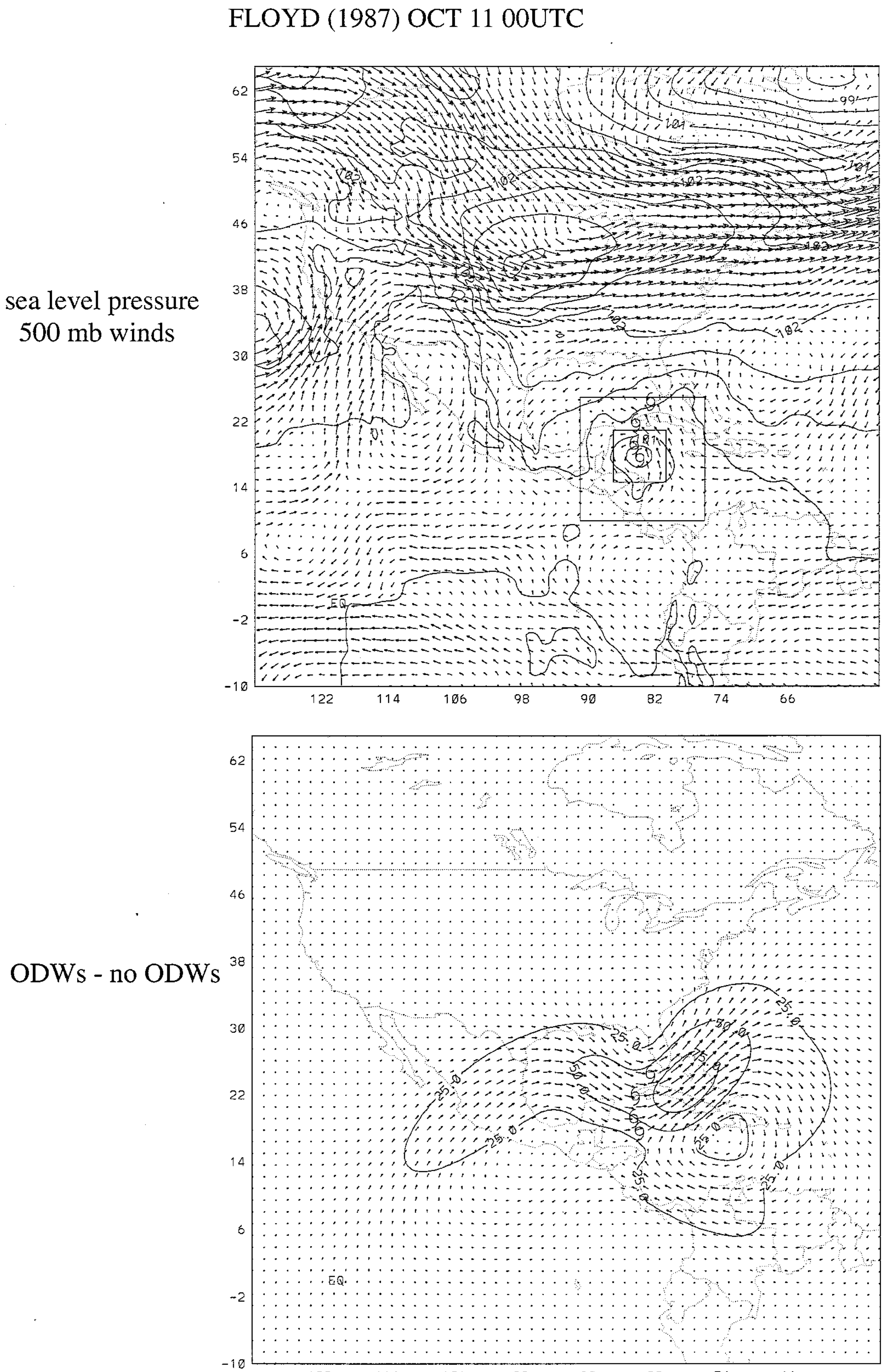

FIG. 10. Sea level pressure (contour interval of 4 mb) and vector winds at model level 10 (~500 mb) at 0000 UTC 11 October 1987 for initial conditions with ODWs (top). The difference (10^{-2} m s^{-1}, 0.25 m s^{-1} contour interval) in wind analyses, ODWs minus no ODWs, when the data are both spacially filtered (see Kurihara et al. 1993) and vertically integrated (bottom). The results are also shown in vector fashion with the observed position of Hurricane Floyd plotted for the first 36 h.

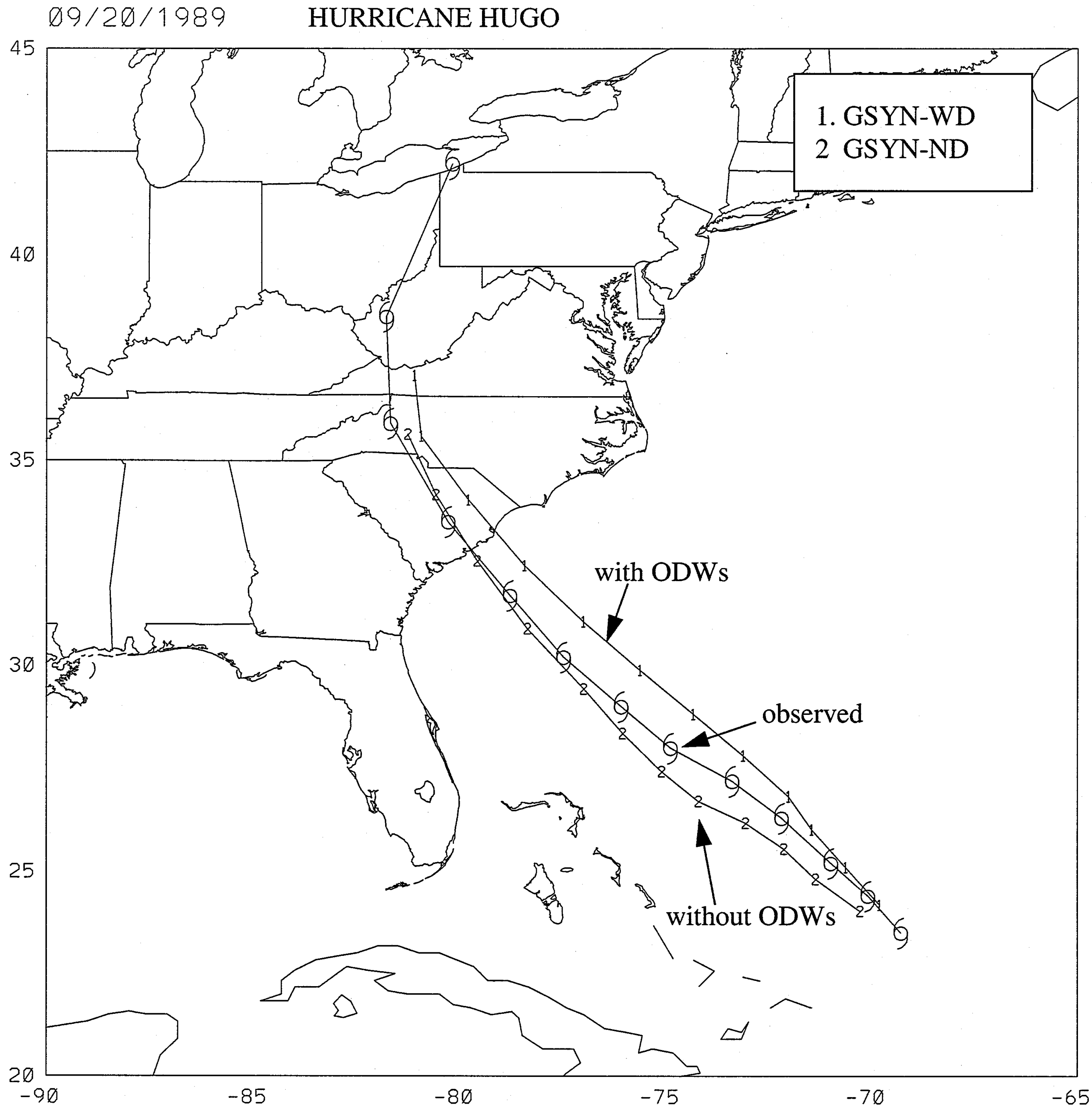

FIG. 11. Tracks of GFDL 3-day forecasts starting from 0000 UTC 20 September 1989 for Hurricane Hugo with and without ODWs ingested in the analyses. See Fig. 9 for other details.

d. Impact of dropwindsondes on intensity

Although the physical factors controlling intensity are not well understood, and achievable skill relative to SHIFOR is difficult, it is nevertheless informative to investigate whether there is any sign of increased skill when ODWs are used. Note that for these 14 cases skill relative to SHIFOR was achieved in forecasting without ODWs after ~30 h using the GFDL model (section 2, Fig. 4). ODWs produce improvement in intensity forecasts, especially for the GSYN experiments (Fig. 13). The mean error reductions of 2, 3.5, 1.5, 2, 1, and 1.5 m s^{-1} at 12, 24, 36, 48, 60, and 72 h, respectively, using ODWs indicate 95% statistical significance for all forecast periods except at 48 h. The GSYN experiments show positive impact of ODWs for the first 2 days when skill is low relative to SHIFOR (Fig. 4). The addition of ODW data in the hurricane environment combined with the NCEP synthetic vortex improves intensity forecasts when no mesoscale vortex specification method such as Kurihara et al. (1993) is used. The GFDL vortex specification system, however, initializes and forecasts a more realistic storm intensity even without ODWs in

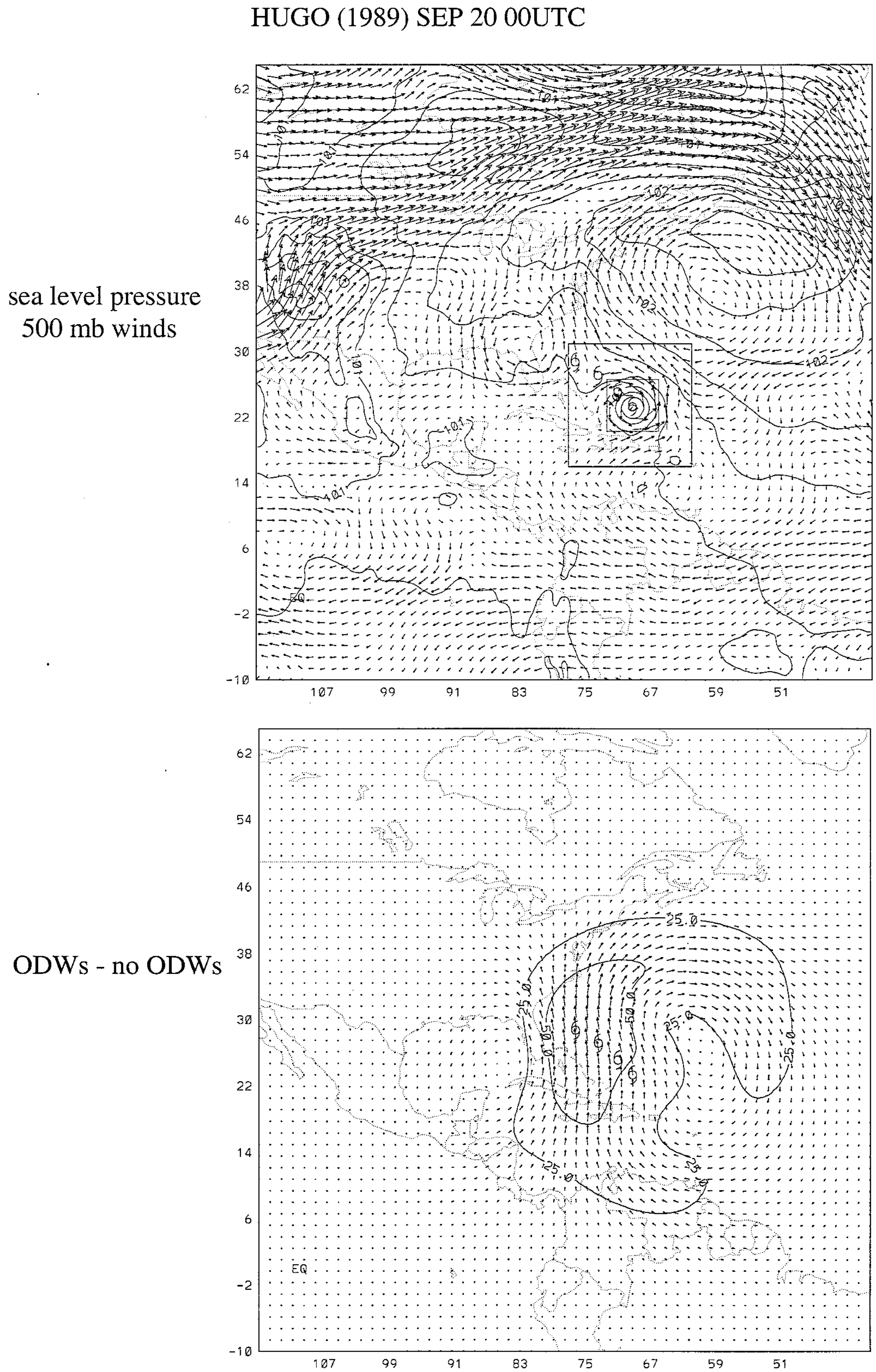

FIG. 12. Sea level pressure (contour interval of 4 mb) and vector winds at model level 10 (~500 mb) at 0000 UTC 20 September 1989 for initial conditions with ODWs (top). The difference (10^{-2} m s^{-1}, 0.25 m s^{-1} contour interval) in wind analyses, ODWs minus no ODWs (bottom). See Fig. 10 for more details.

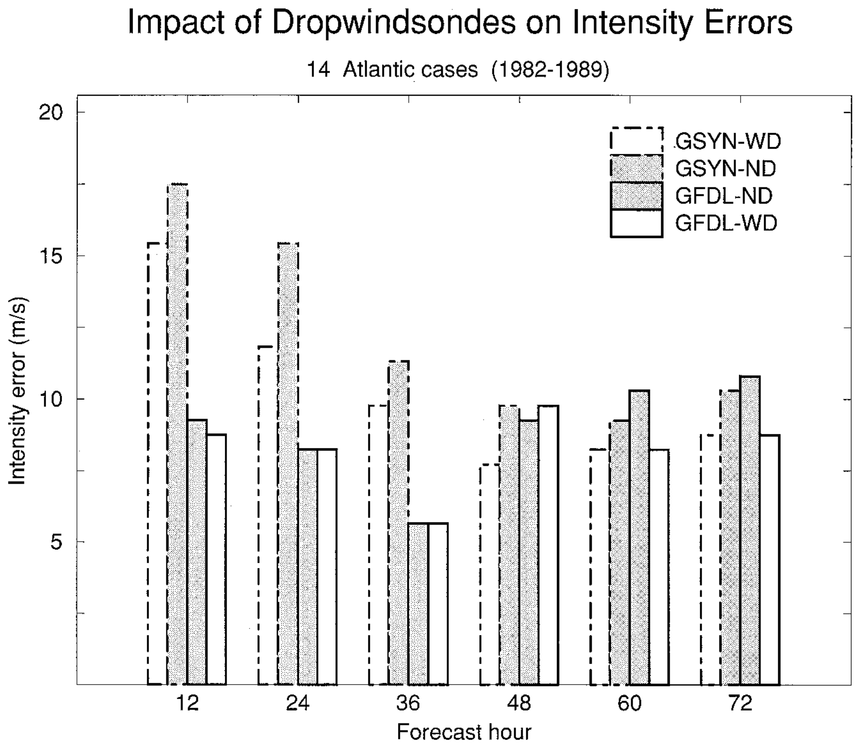

FIG. 13. Mean intensity errors for the GFDL model with and without ODWs for the 14 cases studied. Shaded values at each time period indicate mean values without ODWs, while unshaded values indicate mean values with ODWs. See Fig. 6 for more details.

the early forecast periods. When ODWs are utilized in the GFDL vortex specification system, the mean error shows little significant increase in skill except at 60 and 72 h. For this period 75% of verifying times in the 14-case ensemble show improvement over an already skillful forecast, with a mean error reduction of 20%. Changes due to ODWs may be partially obscured by the known spinup bias at early periods in the GFDL forecast and possibly filtered out by the GFDL vortex specification technique.

Individual forecasts (Fig. 14) show improvement in intensity with ODWs on a case-by-case basis. This is indicated by the majority of solid dots (52 of 75) above the diagonal line for the GSYN forecasts. However as mentioned above, many of the improved cases occur when the intensity forecast is worse than SHIFOR (see Figs. 4 and 13). For the GFDL forecasts, the impacts of ODWs are not as apparent, although the majority (43 of 75 verifying times) indicate improvement. One can look at individual intensity forecasts and their relationship to forecasted large-scale fields. One factor known to be related to storm intensity is vertical shear of the environmental wind (DeMaria and Kaplan 1994). The shear was calculated from 200 to 850 mb for a $10^\circ \times 10^\circ$ area surrounding the forecast storm position for the GSYN forecasts with and without ODWs. This quantity was then compared to both model forecast intensity and observed storm intensity. The results are shown in Table 4 for the 60-h forecast time for the 11 cases when positions were available. Notice that 7 of 11 cases showed improvement in intensity forecasts when ODWs were used and that in 5 of these 7 cases the vertical shear apparently was partially accountable for the improved forecast. Increased (reduced) shear contributed to reduced (increased) intensity including those cases with the three largest improvements in GSYN-WD: error reductions of 60%, 80%, and 15% for the cases of Hurricanes Floyd (11 October), Florence (9 September), and Hugo (21 September), respectively. On the other hand, note the degradation of forecast intensity for Hurricane Emily (24 September) in which GSYN-WD prematurely decayed the storm. One might expect that intensity prediction and track prediction may be correlated. No general relationship was found, however, for these experiments.

4. Summary

We have substantiated the positive impact of ODW observations on hurricane track and intensity forecasts. The GFDL MMM, a sophisticated, state-of-the-art, operational tropical cyclone model, reduced average errors by more than 50 km from 24 to 72 h for 14 cases in the Atlantic. This corresponded to a 20%–30% improvement compared to experiments with no ODWs. Results indicate that differences in the mean environmental

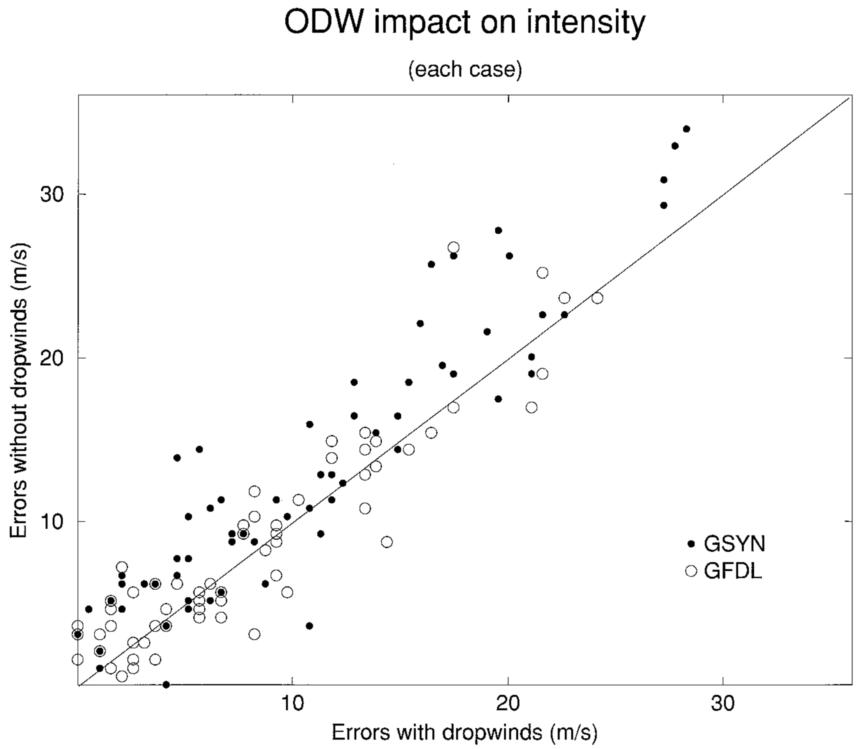

FIG. 14. Plot of individual GFDL model intensity forecast errors using analyses without vs with ODWs at 12, 24, 36, 48, and 72 h. Open circles and closed dots refer to experiments using NCEP synthetic vortex with and without the GFDL vortex specification system, respectively.

steering current of ~1 m s^{-1} were consistent with track improvements with ODWs. These positive impacts were apparent for two different versions of the GFDL forecast system: one utilizing the GFDL vortex specification scheme, the other using only the NCEP synthetic vortex.

TABLE 4. Relation between forecast shear, forecast intensity, and error reduction using ODWs for the 60-h forecast period. The vertical shear was calculated from 200 to 850 mb, averaged between 48 h and 60 h, for a 10° × 10° area surrounding the forecast storm position for the GSYN forecasts with and without ODWs. Bold number indicates where shear decrease (increase) led to forecast intensity increase (decrease) between ODW and no-ODW cases. Initial times were all 0000 UTC.

Storm (year)	Day	Shear difference ND–WD (m s^{-1})	Forecast intensity difference WD–ND (m s^{-1})	Intensity error reduction (%)
Debby (1982)	15 Sep	−1.8	+3	+5
Debby (1982)	16 Sep	**+1.7**	**+1**	+1
Josephine (1984)	10 Oct	+0.5	−1	−2
Josephine (1984)	11 Oct	+1.3	−4	−9
Josephine (1984)	12 Oct	−0.7	+2	+6
Gloria (1985)	25 Sep	**+0.6**	**+1**	+2
Emily (1987)	24 Sep	**−2.7**	**−12**	−35
Floyd (1987)	11 Oct	**−1.4**	**−14**	+60
Florence (1988)	9 Sep	**−3.9**	**−6**	+80
Hugo (1989)	20 Sep	−0.8	−0	+0
Hugo (1989)	21 Sep	**−3.8**	**−3**	+15

This study confirmed, for the most part, results from previous studies including those with a global model and a regional barotropic model for the same cases and initial conditions.

Individual case-by-case analysis was also performed. Track forecast results ranged from a few cases of degradation to improvements greater than 500 km. For approximately 75% of the verifying times of the 14-case ensemble, GFDL model forecasts were improved using analyses with ODWs. Different improvements were observed even between the two versions of the GFDL system for individual cases. Additionally a case-by-case comparison among the NCEP, VICBAR, and GFDL model experiments showed considerable variation, although overall a majority of cases had a similar positive impact. Further study is needed to investigate the variation of ODW impact among cases and to identify reasons why the impact is minimal or negative in some cases, while positive overall. Clearly additional analysis errors remain after ODWs are assimilated and forecast model errors may make further substantial contributions in limiting forecast accuracy. Furthermore research in optimum observing system design is needed for improved tropical cyclone track forecasts.

The GFDL model is the first dynamic model to produce operational intensity forecasts. An evaluation of the impact of ODWs on intensity forecasts found that improvements during the early forecast periods (0–36

h) may be masked by the known spinup bias in the GFDL forecast and GFDL vortex specification system. A combination of ODWs and the GFDL vortex specification system generally produced the most realistic initial wind distribution. For periods greater than 2 days, ODWs have a ~20% positive impact (~2 m s^{-1}) when either the GFDL vortex specification scheme or the NCEP synthetic vortex is used. More work needs to be done to understand the basic physical mechanisms governing storm intensity, but we may speculate that at least in some cases improved prediction of the hurricane environment (e.g., vertical wind shear) produces positive impacts on intensity forecasts. This was supported by the discovery that for the ODW-improved cases at the 60-h forecast period there was a consistency between intensity change and vertical wind shear with and without ODWs.

Acknowledgments. The authors would like to thank J. Mahlman of GFDL and E. Kalnay, R. McPherson, and W. Bonner of NCEP for their continuing support of this project. They would also like to thank M. Bender and Y. Kurihara for aid and discussion in the implementation of the GFDL vortex specification system in these experiments and to J. Gross of the National Hurricane Center for supplying the SHIFOR intensity model. Thanks also are due J. Franklin of the Hurricane Research Division of AOML for ODW processing, providing individual VICBAR results, and discussions concerning ODW improvements in general. Global observations for the Hurricane Debby and Josephine cases were kindly provided by R. Jenne of NCAR.

REFERENCES

Aberson, S. D., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). *Mon. Wea. Rev.,* **122,** 2804–2815.

Bender, M. A., R. J. Ross, R. E. Tuleya, and Y. Kurihara, 1993: Improvements in tropical cyclone track and intensity forecasts using the GFDL initialization system. *Mon. Wea. Rev.,* **121,** 2046–2061.

Burpee, R. W., D. G. Marks, and R. T. Merrill, 1984: An assessment of Omega dropwindsonde data in track forecasts of Hurricane Debby (1982). *Bull. Amer. Meteor. Soc.,* **65,** 1050–1058.

——, 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.

Carr, L. E., and R. L. Elsberry, 1990: Observational evidence for predictions of tropical cyclone propagation relative to environmental steering. *J. Atmos. Sci.,* **47,** 542–546.

Chan, J. C., and W. M. Gray, 1982: Tropical cyclone movement and surrounding flow relationship. *Mon. Wea. Rev.,* **110,** 1354–1374.

DeMaria, M., and J. Kaplan, 1994: A statistical hurricane intensity predictive scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting,* **9,** 209–220.

——, S. D. Aberson, K. V. Ooyama, and S. J. Lord, 1992: A nested spectral model for hurricane track forecasting. *Mon. Wea. Rev.,* **120,** 1628–1643.

Dong, K., and C. J. Neuman, 1986: The relationship between tropical cyclone motion and environmental geostrophic flows. *Mon. Wea. Rev.,* **114,** 115–122.

Franklin, J. L., and M. DeMaria, 1992: The impact of Omega dropwindsonde observations on barotropic hurricane track forecasts. *Mon. Wea. Rev.,* **120,** 381–391.

——, S. E. Feuer, J. Kaplan, and S. Aberson, 1996: Tropical cyclone motion and surrounding flow relationships: Searching for beta-gyres in a Omega dropwindsonde dataset. *Mon. Wea. Rev.,* **124,** 64–84.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 335–342.

——, and Coauthors, 1991: Recent changes implemented into the global forecast system at NMC. *Wea. Forecasting,* **6,** 424–435.

Kurihara, Y., and R. E. Tuleya, 1974: Structure of a tropical cyclone developed in a three-dimensional numerical simulation model. *J. Atmos. Sci.,* **31,** 893–919.

——, and M. A. Bender, 1980: Use of a movable nested mesh model for tracking a small vortex. *Mon. Wea. Rev.,* **108,** 1792–1809.

——, ——, R. E. Tuleya, and R. J. Ross, 1990: Prediction experiments of Hurricane Gloria (1985) using a multiply nested movable mesh model. *Mon. Wea. Rev.,* **118,** 2185–2198.

——, ——, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.,* **121,** 2030–2045.

——, ——, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.,* **123,** 2791–2801.

Lord, S. J., 1991: A bogusing system for vortex circulations in the National Meteorological Center global forecast model. Preprints, *19th Conf. on Hurricanes and Tropical Meteorology,* San Antonio, TX, Amer. Meteor. Soc., 328–330.

——, 1993: Recent developments in tropical cyclone track forecasting with the NMC global analysis and forecasting system. Preprints, *20th Conf. on Hurricanes and Tropical Meteorology,* Miami, FL, Amer. Meteor. Soc., 290–291.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus,* **17,** 321–333.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.