P1.63

HWRF PERFORMANCE DIAGNOSTICS FROM THE 2009 ATLANTIC HURRICANE SEASON

Brian McNoldy *[1], Mark DeMaria [1], Vijay Tallapragada [2], and Timothy Marchok [3]

[1] Colorado State University – Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado
[2] NOAA Environmental Modeling Center, Camp Springs, Maryland
[3] NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

1. INTRODUCTION

HWRF, the Hurricane Weather Research and Forecast System, was utilized for a third year as an NCEP operational model in 2009. However, the 2009 Atlantic hurricane season presented a challenge for many dynamical prediction models, particularly HWRF (Franklin, 2010). It was a relatively quiescent season, with nine named storms and three hurricanes, compared to the 1950-2000 average of ten named storms and six hurricanes. The decreased activity was largely attributed to a moderate El Niño event (Klotzbach and Gray, 2009), which is known to increase the vertical wind shear over the tropical Atlantic, and therefore decrease the likelihood of tropical cyclogenesis and limit the intensification of tropical cyclones that do form (Gray, 1984).

The GFDL (Geophysical Fluid Dynamics Laboratory) Coupled Hurricane-Ocean Prediction System is another full-physics, high-resolution dynamical model and has been used operationally by NCEP/NHC since 1995. Although it also struggled with accurately forecasting many of 2009's storms, it frequently out-performed HWRF, particularly in the 3-5 day forecast periods. This of course leads to the question: *why*?

Challenges for prediction models relevant to this study are to accurately initialize a tropical cyclone's structure, accurately forecast the vertical wind shear along the storm's track, and accurately evolve the tropical cyclone (TC) structure based on the vertical wind shear. A poor track forecast compounds the issue by potentially placing the storm over land or unrealistic sea surface temperatures, greatly increasing the intensity error.

HFIP, the Hurricane Forecast Improvement Project, is a ten-year NOAA program aimed at improving TC track and intensity forecasts through model improvements, optimal use of observations, and model diagnostics and verification. HWRF and GFDL are two of the operational regional forecast models in the program; their performance during the 2009 Atlantic hurricane season will be investigated here.

2. DATA

Detailed descriptions of the HWRF and GFDL model equations, inititalizations, parameterizations, grid configurations, *etc.* can be found in Gopalakrishnan *et al.* (2010) and Bender *et al.* (2007), respectively.

Data from the 2009 operational HWRF and GFDL models are processed and summaries are written out in identical formats for easy comparison to validation data (McNoldy, 2010a and 2010b). The storm center is not taken to be the center of the nested grid, but rather is assumed to be co-located with the lowest sea level pressure closest to the center of the nested grid. The sea surface temperature values are calculated from an average of the five model gridpoints directly under the storm center. The vertical wind shear is calculated between 850-200 hPa and from winds averaged in an annulus 200-250 km around the storm center, which removes the effect of the vortex itself and any slight uncertainty in the exact location of the storm center. [Vertical wind shear is calculated in two areas in the verification database: 0-500 km and 200-800 km, however, the limited size of the nested grids in HWRF and GFDL do not allow either of these areas to be used, so 200-250 km was chosen as a compromise.] The intensity for the HWRF model storm is taken to be the maximum 10 m wind speed closest to the center of the nested grid, and the intensity for the GFDL model storm is taken to be the maximum 35 m wind speed closest to the center of the nested grid. Since the GFDL model does not output 10 m winds, the 35 m winds are reduced by a constant factor of 0.89 to closely represent 10 m winds in the eyewall region (Marchok, 2010). Note that this is a slightly larger reduction than the 0.92 factor found by Franklin *et al.* (2003), and is only being claimed to be valid for the operational GFDL data.

The models are validated against operational NCEP global model fields and NHC best track data, both available in the post-season time frame

---

* Corresponding author: Brian McNoldy, Colorado State Univ., Dept. of Atmospheric Science, Fort Collins, CO 80523-1371; e-mail: mcnoldy@atmos.colostate.edu

(DeMaria, 2010). The majority of the fields are self-explanatory, but the sea surface temperature (SST) fields used in this study are Reynolds SST, and the vertical wind shear is calculated between 850-200 hPa and from winds averaged in a 500 km radius circle around the storm center, removing the effect of the vortex itself. The position and intensity are from the NHC best track.

3. PRELIMINARY RESULTS

The full-season average intensity and track errors are shown in figure 1. The 5-year (2004-2008) average NHC errors are plotted in the dotted black line, the 2009 operational NHC errors are plotted in the red line, the 2009 operational HWRF errors are plotted in the purple line, and the 2009 operational GFDL errors are plotted in the green line. The homogeneous sample size at each forecast hour is shown in parentheses, and the values all come from post-season ATCF data.

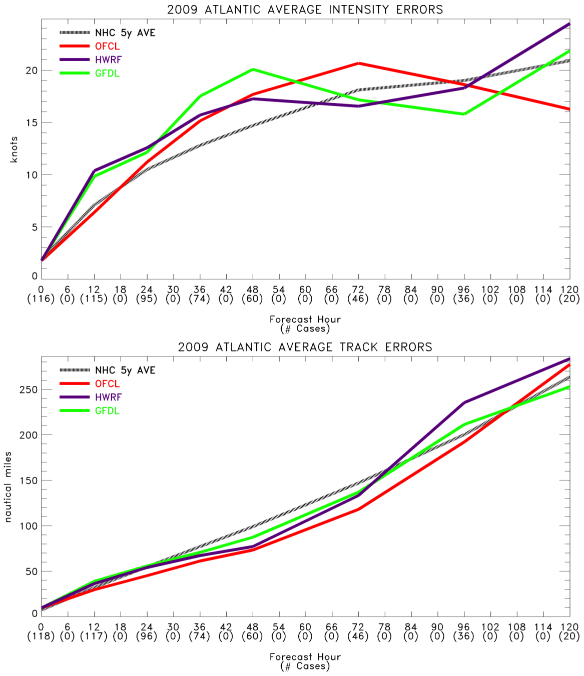

Figure 1. Intensity (top) and track (bottom) errors for NHC (red), HWRF (purple), and GFDL (green) during the 2009 Atlantic hurricane season.

Notice that the 2009 NHC intensity errors are generally larger than their 5-year errors for the 24-96 h forecasts, and slightly smaller for the 120 h forecasts. The HWRF and GFDL models performed similarly through approximately 24 h, then HWRF out-performed GFDL from 24-72 h, and from 72-120 h, GFDL out-performed HWRF. In terms of track errors, HWRF and GFDL performed similarly through 72 h, then GFDL was more accurate from 72-120 h, and neither of them out-performed the NHC (with the exception of GFDL's average 120 h forecast).

The high bias in forecast intensity comes mostly from weak systems in hostile environments. For systems that are intense or that become intense, both HWRF and GFDL performed reasonably well in 2009 (*e.g.*, Bill, Fred, and Ida). However, not all systems develop, whether due to vertical wind shear, cool SST, land, etc. Both dynamical models considered here had a very difficult time forecasting weakening or steady-state intensity (particularly for Ana, Danny, and Erika). Figure 2 shows each individual intensity forecast for Bill (03L), a case where HWRF and GFDL performed reasonably well. The track forecasts were also quite good (not shown).

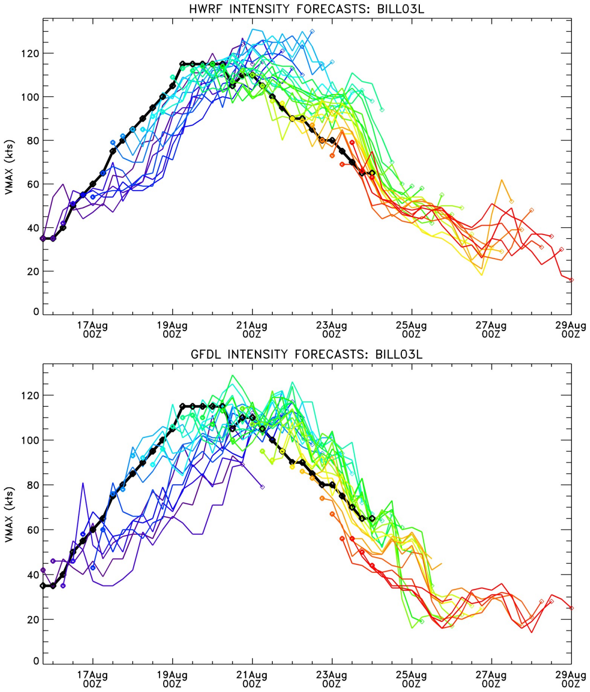

Figure 2. Intensity forecasts for Bill (03L) made by HWRF (top) and GFDL (bottom). The best track intensity is shown by the thick black line.

However, in a case with moderate vertical wind shear, Erika (06L), the results were not as positive. Figure 3 shows the HWRF intensity forecasts; GFDL was qualitatively similar (not shown). The models consistently over-intensified the storm, though HWRF more so.

Examination of individual but representative runs, a potentially critical feature becomes evident.

In figure 4, the HWRF analysis from 2 Sept at 12 UTC is shown for Erika (06L), with the 10 m wind speed and sea level pressure contoured in the top panel, and the along-shear vertical cross-section of wind speed and potential vorticity (PV) contoured in the bottom panel.

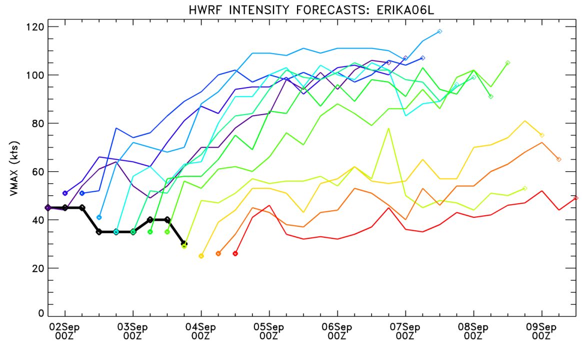

Figure 3. Intensity forecasts for Erika (06L) made by HWRF. The best track intensity is shown by the thick black line.

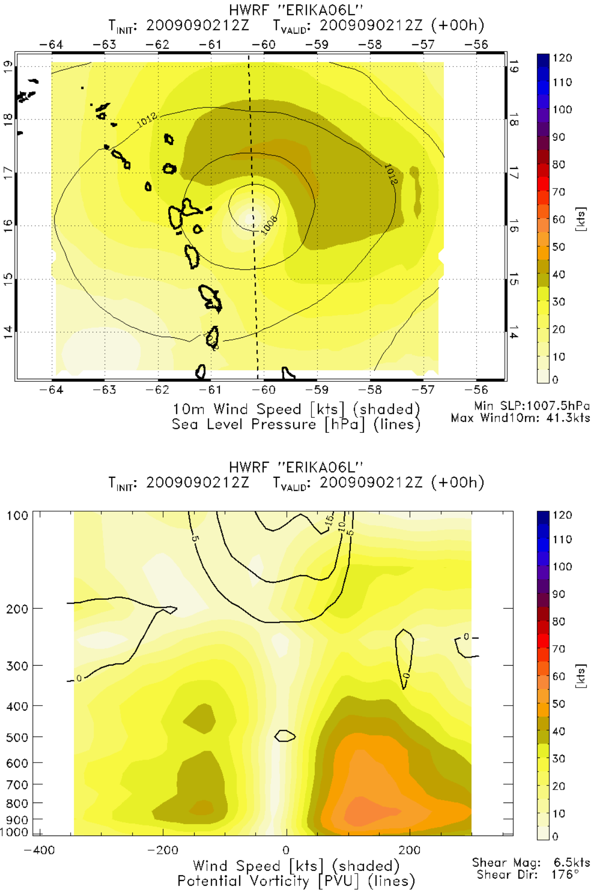

Figure 4. HWRF analysis from 2 Sept 12 UTC for Erika (06L). Wind speed at 10 m and sea level pressure are shown in the top panel, along-shear vertical cross-section of wind speed and PV are shown in the bottom panel.

The same fields from the same time are shown in figure 5, but from the GFDL model. Note that the horizontal domain is slightly smaller (smaller nested grid). Although the surface wind and pressure fields are fairly similar, GFDL produces a much less vertically-coherent vortex. In fact, there is barely a discernible vortex above the surface, which more closely simulates the real storm (not shown). Numerous studies have shown (Shapiro and Montgomery (1993), Flatau *et al.* (1994), *Jones* (1995), DeMaria (1996), Vandermeirsh *et al.* (2002), Reasor *et al.* (2004), and others) that a vertically-coherent vortex is much more likely to maintain itself or even intensify in the face of moderate vertical wind shear. This ability of deeper and/or stronger vortices to resist vertical shear is explained by a variety and combination of processes, including larger Rossby penetration depths, vertical transport of PV from low levels to mid- and upper-levels acting to keep the vortex rigid, co-rotation of the low- and upper-level centers acting to restore the vortex to the vertical, *etc.*

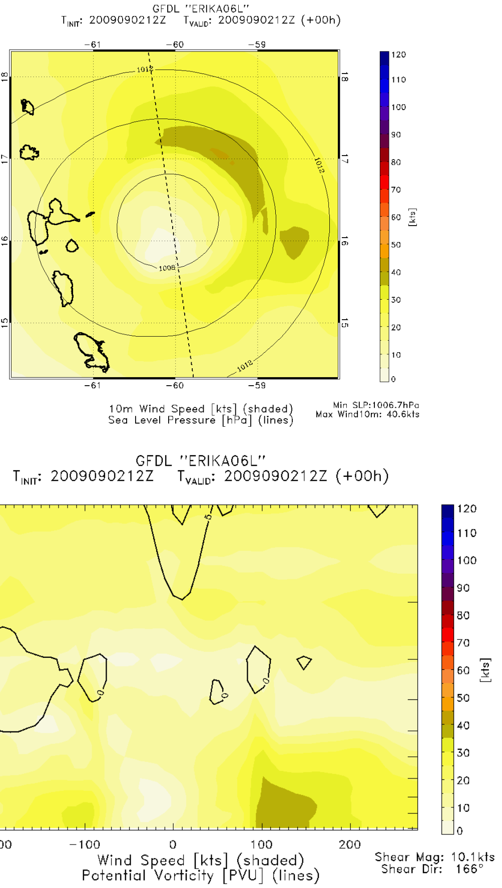

Figure 5. GFDL analysis from 2 Sept 12 UTC for Erika (06L). Wind speed at 10 m and sea level pressure are shown in the top panel, along-shear vertical cross-section of wind speed and PV are shown in the bottom panel.

Extending out in time, the situation is dramatically amplified. Figures 6 and 7 show the same fields as figures 4 and 5, but for 36 h into the model runs. HWRF produces an intensifying 88 kt hurricane with a 964 hPa central pressure, full tropospheric vertical extent, and mature PV "tower" in the eye. GFDL, on the other hand, produces a 55 kt tropical storm with a 1004 hPa central pressure, and barely any vertical extent to speak of.

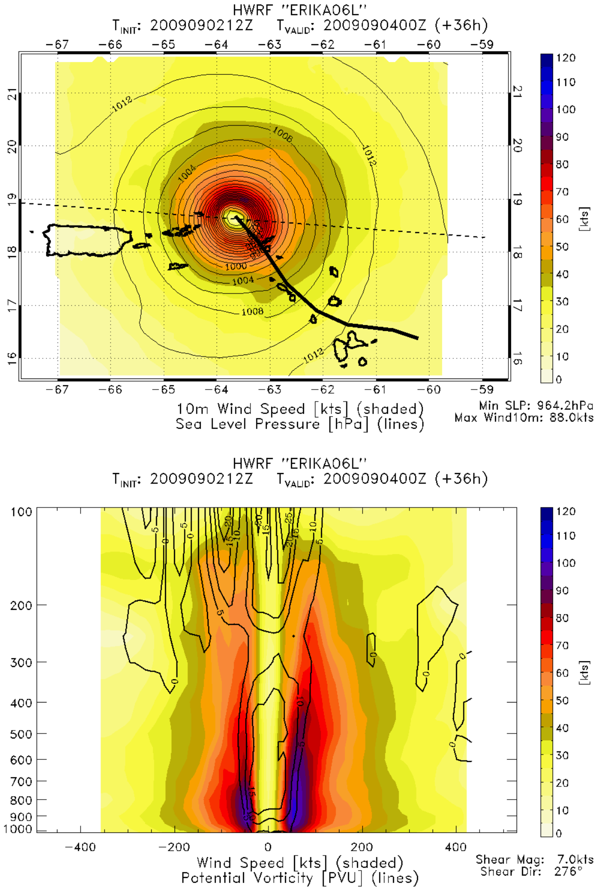

Figure 6. HWRF 36 h forecast from 2 Sept 12 UTC (valid 4 Sept 00 UTC) for Erika (06L). Wind speed at 10 m and sea level pressure are shown in the top panel, along-shear vertical cross-section of wind speed and PV are shown in the bottom panel.

In reality, Erika at this time was classified as a 25 kt remnant low, with "a tight swirl of low clouds accompanied by a small area of deep convection to the southeast" (Avila, 2009). The vertical shear was analyzed to be 18 kts at 267º.

Seeing how the vortex in HWRF remains so rigid and upright, it leads to the question of vertical shear accuracy. Did HWRF under-analyze and under-forecast the vertical wind shear? Using the definition of wind shear described in section 2, and applying it to HWRF and GFDL, the answer is no, not always. In fact, for the full-season average, HWRF actually over-analyzed the vertical shear, by 3 kts at 12 h and by 4 kts at 60 h (see figure 8). The mean absolute error (which removes the sign of the errors, resulting in the total magnitude) at those times is 5 kts and 7 kts. GFDL had a similar mean absolute error (MAE) through 60 h, but smaller average difference (AD). At 12 h, the MAE was 4 kts and the AD was 0.5 kts; at 60 h, the MAE was 7 kts and the AD was 2 kts. After 60 h, HWRF had a smaller MAE but larger AD than GFDL indicating that GFDL had larger random errors while HWRF had smaller but positively biased errors.

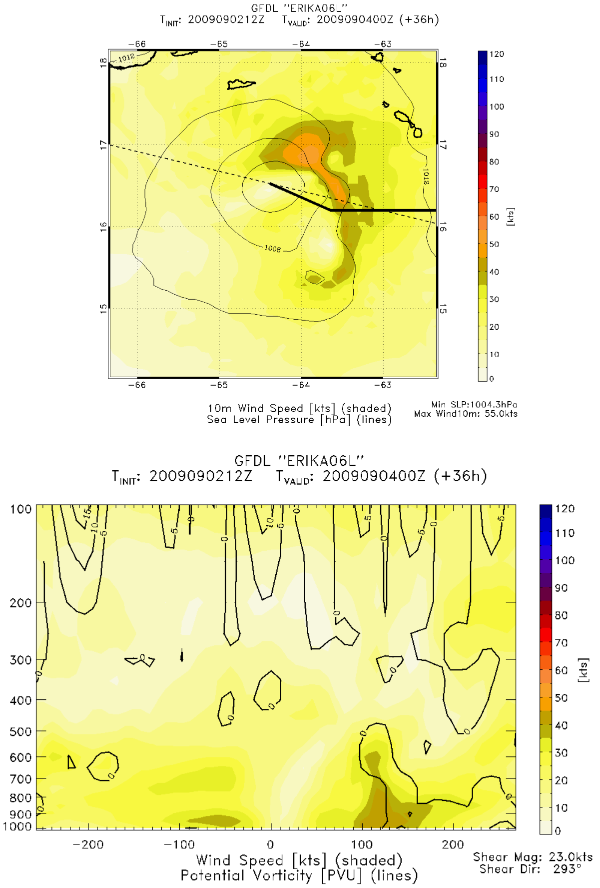

Figure 7. GFDL 36 h forecast from 2 Sept 12 UTC (valid 4 Sept 00 UTC) for Erika (06L). Wind speed at 10 m and sea level pressure are shown in the top panel, along-shear vertical cross-section of wind speed and PV are shown in the bottom panel.

For clarification, if the simple average difference $(\overline{x_1 - x_2})$ at a given time is 0 kts, the mean absolute error $(\overline{|x_1 - x_2|})$ could still be quite large. Consider a sample size of two errors: +15 kts and -15 kts; the mean absolute error is 15 kts, but the average difference is 0 kts. Clearly, the model did not perform well, but there is not a systematic bias.

Figure 8. Mean absolute error (top) and average difference (bottom) of vertical wind shear calculated from HWRF (purple) and GFDL (green) model output. The top panel shows that HWRF and GFDL had similar absolute errors through 60 h, but the bottom panel reveals that HWRF had a high bias at all forecast times.

4. CONCLUSIONS AND FUTURE WORK

The 2009 Atlantic hurricane season presented unusual challenges for sophisticated dynamical forecast models such as HWRF. The season was relatively inactive, but a moderate El Niño generated vertical wind shear over the tropical Atlantic basin, and it is difficult to predict exactly how much influence that shear will have on a particular storm.

This preliminary study shows that the problems HWRF had were a combination of 1) the vortex structure being too deep and too rigid, 2) the vortex's inadequate response to vertical shear, and 3) poor forecasts of vertical shear.

There are several methods that can be used to calculate vertical wind shear. The basic 200-850 hPa area average vector difference is the most common, but those exact layers may not (most likely DO NOT) best represent the vertical wind profile in which the vortex is embedded. We plan on extending this work to calculate a generalized shear from the model output, which uses the deviation of the mass-weighted winds at each pressure level from the mass-weighted deep-layer mean winds.

5. ACKNOWLEDGMENTS

This research is funded by NOAA/HFIP Contract NA09OAR4320074#21. We also would like to thank Wayne Schubert, Paul Ciesielski, and Bob Tuleya for helpful comments.

6. REFERENCES

Avila, L., 2009: Tropical Depression Erika Discussion Number 10. Available online at http://www.nhc.noaa.gov/archive/2009/al06/al062009.discus.010.shtml

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, T. P. Marchok, 2007: The operational GFDL Coupled Hurricane-Ocean Prediction System and a summary of its performance. *Mon. Wea. Rev.*, **135**, 3965-3989.

DeMaria, M., 1996: The effect of vertical shear on tropical cyclone intensity change. *J. Atmos. Sci.*, **53**, 2076-2087.

DeMaria, M., 2010: Statistical tropical cyclone intensity forecast technique development. Available online at http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp

Flatau, M., W. H. Schubert, and D. E. Stevens, 1994: The role of baroclinic processes in tropical cyclone motion: The influence of vertical tilt. *J. Atmos. Sci.*, **51**, 2589-2601.

Franklin, J. L., M. L. Black, and K. Vale, 2003: GPS dropwindsonde wind profiles in hurricanes and their operational implications. *Wea. Forecasting*, **18**, 32-44.

Franklin, J. L., 2010: 2009 National Hurricane Center forecast verification report. Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2009.pdf

Gopalakrishnan, S., Q. Liu, T. Marchok, D. Sheinin, N. Surgi, R. Tuleya, R. Yablonski, and X. Zhang, 2010: Hurricane Weather Research and Forecasting (HWRF) model scientific documentation. Available online at http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRF_final_2-2_cm.pdf

Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part I: El Niño and 30 mb Quasi-Biennial Oscillation influences. *Mon. Wea. Rev.*, **112**, 1649-1668.

Jones, S. C., 1995: The evolution of vortices in vertical shear. Part I: Initially barotropic vortices. *Q. J. R. Meteorol. Soc.*, **121**, 821-851.

Klotzbach, P. J., and W. M. Gray, 2009: Summary of 2009 Atlantic tropical cyclone activity and verification of author's seasonal and 15-day forecasts. Available online at http://hurricane.atmos.colostate.edu/Forecasts/2009/nov2009/nov2009.pdf

Marchok, T. P., 2010: personal communication.

McNoldy, B. D., 2010a: Operational HWRF model fields. Available online at http://einstein.atmos.colostate.edu/~mcnoldy/tropics/hwrf2/

McNoldy, B. D., 2010b: Operational GFDL model fields. Available online at http://einstein.atmos.colostate.edu/~mcnoldy/tropics/gfdl/

Reasor, P. D., M. T. Montgomery, and L. D. Grasso, 2004: A new look at the problem of tropical cyclones in vertical shear flow: Vortex resiliency. *J. Atmos. Sci.*, **61**, 3-22.

Shapiro, L. J., and M. T. Montgomery, 1993: A three-dimensional balance theory for rapidly rotating vortices. *J. Atmos. Sci.*, **50**, 3322-3335.

Vandermeirsh, F., Y. Morel, and G. Sutyrin, 2002: Resistance of a coherent vortex to a vertical shear. *J. Phys. Oceanogr.*, **32**, 3089-3100.