7B.2 FORECAST VERIFICATION USING METEOROLOGICAL EVENT COMPOSITES

Jason E. Nachamkin*

Naval Research Laboratory, Monterey, CA

1. INTRODUCTION

Intense wind events over the open ocean are often a concern to Navy operations. Since high-resolution numerical models are frequently used to explicitly forecast these events, event-based evaluation would be a useful measure of model performance. Ebert and McBride (2000), and Baldwin et al (2001) describe methods of varying complexity in which pattern shifting or recognition can be used in this type of validation. These methods are quite useful, especially when reliable and consistent observational data are widely available. Unfortunately, most oceanic wind observations consist of scattered ship observations or satellite data. The satellite data are quite promising for their relatively high resolution, but coverage is limited to finite swaths. Rain and very high winds, often associated with oceanic wind events, further limit the coverage. In many cases, the full extent, intensity, and coverage of a specific event are only partially known. This causes considerable problems for techniques based on direct model-to-observation comparisons of specific events.

Composite techniques offer an attractive and relatively simple approach to this problem. The philosophy is similar to that applied by Gray and Frank (1977), who used sounding composites to gain insight on hurricane structure. Applying a similar philosophy to event verification relaxes the information requirements on any one event. If enough quasi-random observations of a distribution of similar events exist, bulk properties of the forecasts and the observations can be reliably estimated. This allows partial observations to be smoothly incorporated into a coherent, statistically meaningful comparison.

In this work, the composite method is applied to forecasts of the Mistral. Mistrals are well-defined regions of strong northeasterly, northerly, or northwesterly winds that occur in the northern Mediterranean Sea (Fig. 1). The flow is often characterized by wind funneling through gaps in several mountain ranges between the Alps and Pyrenees. Anecdotal evidence indicates that these events are well predicted by the models, thus the event-driven statistics should show some palpable measure of skill.

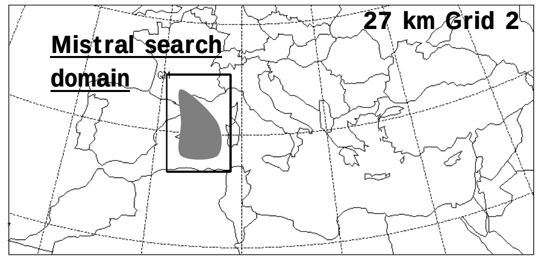

Fig. 1. The COAMPS™ Europe 27 km Grid domain. The Mistral event collection sub-domain is indicated by the box. Shading indicates the coverage of a typical Mistral.

2. MODEL AND OBSERVATIONAL DATA

Mistral wind events were selected from approximately one year of forecasts from the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS™)[1] operational Europe (Fig. 1) forecasts generated at the Fleet Numerical Oceanography and Meteorology Center (FNMOC). Ten-meter wind data were collected from all forecasts initialized at 0000 UTC and 1200 UTC from 1 November 2000 through 30 October 2001. Each forecast was initialized with the multivariate optimum interpolation (MVOI) analysis (Barker, 1992), and boundary conditions were obtained from the Navy Operational Global Atmospheric Prediction System (NOGAPS). Two one-way nested grids were used in both areas with horizontal spacings of 81 and 27 km, respectively. The domain had 30 sigma levels in the vertical, with the lowest level at 10 m AGL. The forecasts were all run to 72 hours.

The forecasts were verified against the Special Sensor Microwave/Imager (SSM/I) winds retrieved using the Goodberlet et al. (1990) regression. In the absence of rain, Goodberlet et al. (1990) found the retrieved winds estimated the in-situ buoy and ship winds at the 19.5 m level to within 2 m s^{-1}. For this study, All rain-flagged data were discarded, and the SSM/I speeds were adjusted to 10 m, for comparison to the model data, using the logarithmic wind profile. Typical adjustments were less than 1 m s^{-1}.

* *Corresponding author address:* Jason E. Nachamkin
Naval Research Laboratory, Monterey, CA 93943-5502
email address: nachamkin@nrlmry.navy.mil

[1] COAMPS is a registered trademark of the Naval Research Laboratory.

The SSM/I data were bilinearly interpolated to the model 27 km grid, which is quite compatible with the 25 km SSM/I footprint. The typical swath was about 1400 km wide, and all satellite passes that occurred within one hour of the verification time were included. The higher-density data passes over Europe occurred at approximately nine-hour intervals, starting from 0600-0900 UTC.

3. VERIFICATION TECHNIQUE

A rules-based algorithm was used to define unique, contiguous Mistral events in each forecast. A sub-domain was defined on the 27 km grid (Fig 1) in which the algorithm searched for Mistrals. All contiguous points with winds greater than 12 m s^{-1} and directions between 270 and 070 degrees were grouped as an event. Those events with portions that extended beyond the sub-domain were kept in the composite if the center of the event was still within the sub-domain.

Due to the lack of contiguous observations, all Mistral events were identified using the model forecasts alone. This implies, at least initially, that only the conditional distribution of Mistrals given that a Mistral is predicted can be obtained. Once this distribution is known, there are ways to estimate the full distribution of observed events, but that will be discussed later in this work.

Once an event was identified, all surrounding data were transferred to a 31X31 point relative grid with the same grid spacing as the model. The center of the forecast event as defined by its "center of mass" was positioned at the center of the relative grid. At that point, all available observational data were also positioned on the relative grid. Model data were then templated by the available observations, such that all forecasts outside of the SSM/I swath were cut from the set (Fig. 2).

Transferring the events to the relative grid effectively synchronizes them about a common central point. The primary sources of variance within the event distribution are now size, shape, and intensity. Although size criteria could easily be used to further refine the distribution, high-quality, meaningful results were obtained by compositing all events with a total forecast size of 75 or more grid points. Smaller events did exist, but they were often not well-defined Mistrals.

Although the Mistral is relatively common, the observations were not frequent enough to derive composites at every forecast hour. Instead model-SSM/I samples were collected in six-hour intervals starting with the three-hour forecast. Thus, statistics associated with the six-hour forecast actually contain SSM/I-model comparisons for all forecasts between three and nine hours. The intervals between zero and three hours and 69-72 hours were not used as almost no SSM/I overpasses occurred at those times. The increased interval slightly broadens the distributions, but the compromise is necessary to maintain statistical significance.

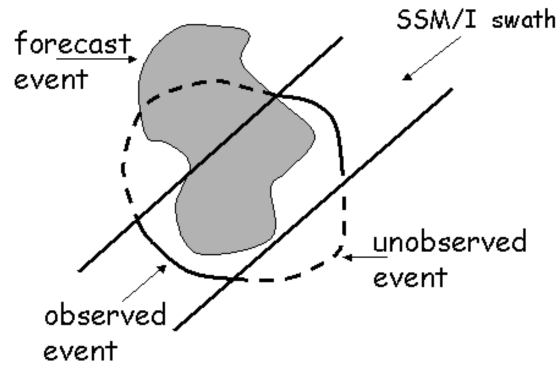

Fig. 2. Schematic depicting the data collection strategy for the event composites.

4. RESULTS

Since compositing is a means of sampling a larger distribution, some degree of error will occur. The magnitude of this error can be estimated by comparing speeds from the model data collected at valid SSM/I points with those from the full distribution of model speeds. In general, the sampled speeds underestimated the true speeds by an average of 2.28 m s^{-1}. However if only those points with 20 or more samples are used, the underestimate drops to -0.35 m s^{-1}. Thus the speed-related statistics in this work are only derived for areas with 20 or more SSM/I samples. Note the $N< 20$ mask (Fig. 3) reveals minima associated with Sardinia, Corsica, the Balearic islands, and French/Spanish coastline.

Other simple statistics reveal several characteristics of the simulated and observed Mistrals. The average speed near the center of a given 18-hr Mistral forecast approaches 16 m s^{-1} (Fig. 3a), while the bias (Fig. 3b) indicates that average predicted speeds are 2 m s^{-1} higher than the observations here. On either side of the Mistral, model winds tend to be 1-3 m s^{-1} low. These numbers must be carefully interpreted as the forecasts and the observations are not quite synchronized. The existence of a predicted Mistral is a given fact, but multiple false alarms may be compensated by extreme underestimates during a few major events. Fortunately such a scenario would incur very high RMS errors. The RMS values in this case are relatively low, especially within the Mistral suggesting that the bias is more systematic. The largest RMS and bias errors exist on the eastern and western sides of the Mistral, in regions where the predicted winds are relatively weak. Variance plots (not shown) indicate higher inter-event variability in the observations than the forecasts in these areas. This suggests that high winds occasionally occur in these areas, but the model does not typically predict them.

The exact degree of synchronization between the model and the observations is difficult to know given the nature of the data. However, occurrence distributions, defined by the number of points the exceeding 12 m s^{-1} threshold, (Fig. 4) can be cross-checked with the bias/RMS statistics for consistency. The distributions in Figure 4a indicate that model speeds near the center of the average Mistral crossed

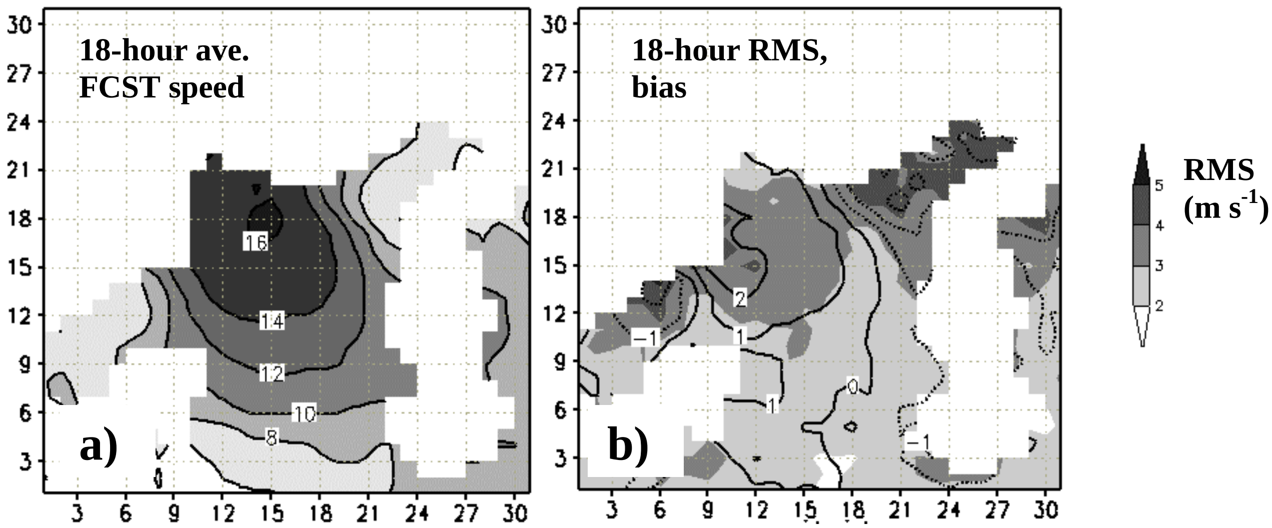

Fig. 3. Average speed in m s^{-1} for all 18-hour Mistral forecasts at valid SSM/I points is shaded in (a). In (b), the 18-hour RMS is shaded and the bias (FCST-SSM/I) is contoured, both statistics are in m s^{-1}. All plots are on the relative grid, each tick interval represents three grid points or 51 km. Statistics were only derived at data points with 20 or more observations.

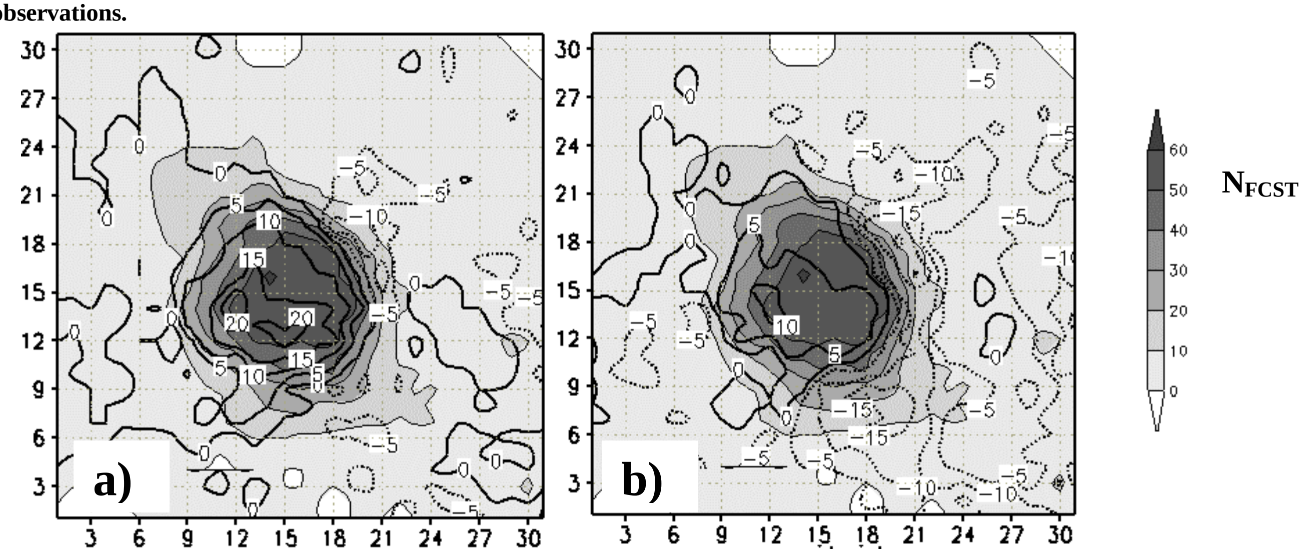

Fig. 4. The number occurrence of 18-hour predicted winds exceeding 12 m s^{-1} (N_{FCST}) is shaded in both (a) and (b). Contours in (a) represent the difference between the SSM/I and predicted occurrence distributions (N_{FCST}-$N_{SSM/I}$). Contours in (b) are the same as (a) except 1.5 m s^{-1} was added to all SSM/I data. Axis tick intervals are as in Fig. 3.

the 12 m s^{-1} threshold up to 30% less frequently than the SSM/I did. Initially, this seems like a problem, however the numbers are quite sensitive. Adding 1.5 m s^{-1} to the SSM/I values significantly shifts the structure of the difference field (Fig. 4b), leaving higher observed values near the Mistral center. This reveals a peril of relying on knife-edge thresholds to discretize a continuous field. It also indicates that the observed and predicted Mistral occurrence distributions are similar to one another. Direct correlations between the model and observed speed distributions (Fig.5) are also quite high through the forecast period.

The bias and RMS errors, though steadily increasing, stay relatively low through 66 hours, indicating that the distributions are relatively well synchronized. The smallest overall bias errors occur between 30 and 42 hours, however the RMS and correlation scores are slightly worse than at earlier forecasts. The spatial error patters maintain a structure similar to those at 18-hours (Figs 3,4) with an increased amplitude. Overall, these statistics indicate that COAMPS predicts the Mistral well out to three days.

Finally, as mentioned in Section 3, the conditional distribution of the predicted events given that a Mistral is *observed* is difficult to estimate from the SSM/I data.

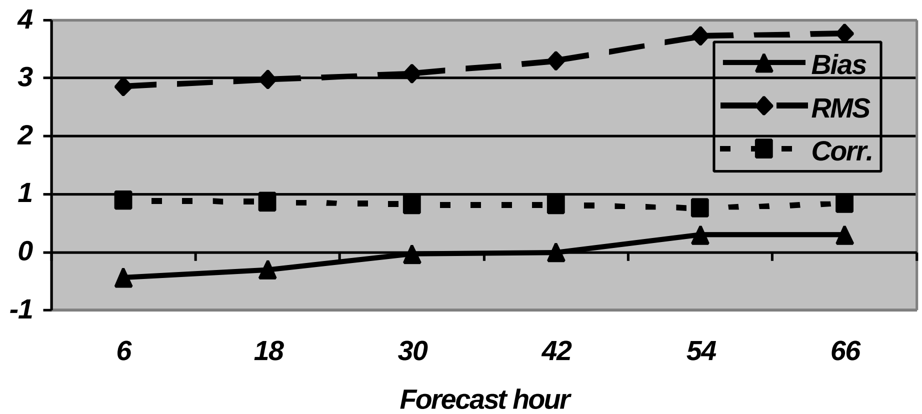

Fig. 5. Relative grid-total values of RMS and bias (FCST-SSM/I) in m s^{-1} are plotted for each forecast hour as dashed and solid lines, and the correlation coefficient between the FCST and SSM/I speed distributions is plotted as the short-dashed line.

However, if a subset of forecasts is found to have a very high probability of being correct, one may be able to assume that the partially observed events in that subset are indeed true events. These could then be used to validate the conditional probabilities for subsequent forecasts in a composite sense. An experiment of this nature is planned in the near future. It should be noted that this approach would have limited use for certain phenomena, such as precipitation, but a similar philosophy could at least be applied to handle precipitation observations of varying quality.

5. CONCLUSIONS

Deterministic predictions of specific events are among the primary uses of a mesoscale model. Verification of these forecasts is a difficult problem, as the accuracy of this information must often be diagnosed using observational data with inadequate precision. As grid spacing collapses towards 1 kilometer and below, full deterministic knowledge of any relevant event at that scale may forever be beyond reach. This problem is similar to the conundrum faced by quantum physicists. At quantum scales the exact nature of matter is unknown, and statistical methods are sometimes the only way to reveal information about the system and its behavior. Though hardly quantum physics, compositing evokes a similar philosophy. Bulk properties of a distribution of precise forecasts can be diagnosed using observations of limited precision in conjunction with relatively simple statistics. Although the nature of the distributions is ambiguous at times, enough information is often available to consistently diagnose forecast quality. Even when, or if, precise measurements are available, some precision is lost when combining multiple event occurrences for statistical purposes. This issue will likely play a role in the development of new mesoscale verification tools.

6. REFERENCES

Baldwin, M. E., S. Lakshmivarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. *Preprints, 9th Conference on Mesoscale Processes,* 30 July-2 August, Ft. Lauderdale, FL, American Meteorological society, 255-258.

Barker, E. H., 1992: Design of the Navy's multivariate optimum interpolation analysis system. *Wea. Forecasting,* **7,** 220-231.

Ebert, E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *Journal of Hydrology*, **239**, 179-202.

Goodberlet, M. A., C. T. Swift and J. C. Wilkerson, 1990: Ocean surface wind speed measurements of the Special Sensor Microwave/Imager (SSM/I). *IEEE trans. Geoscience Remote Sens*., **28**, 823-827.

Gray, W. M., and W. M. Frank, 1977: Tropical cyclone research by data compositing. NEPRF Technical Report TR-177-01, Naval Environmental Prediction Research Facility, Monterey, CA, 70 pp.

6. ACKNOWLEDGEMENTS

This research is supported by the Office of Naval Research (ONR) through program element 0602435N, and the Space and Naval Warfare Systems Command (SPAWAR) through program element 603207N. Computing time was supported in part by a grant of HPC time from the Department of Defense Shared Resource Center, Stennis Space Center, MS, and performed on a Cray T-90. Special thanks go to Kim Richardson (NRL), the NRL satellite data group, and FNMOC for providing the SSM/I data.