DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

Automated Verification of Mesoscale Forecasts using Spatial Statistics

David W. Jones
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 543-3236 fax: (206) 545-4404 email: dwjones@apl.washington.edu

Scott A. Sandgathe
35574 Camp Creek Rd.
Springfield, OR 97478
phone: (541) 988-0289 fax: (541) 988-0289 email: sandgathe@apl.washington.edu

Caren Marzban
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 221-4361 fax: (206) 545-4404 email: marzban@stat.washington.edu

Award Number: N00014-01-G-0460/0049
http://eis.apl.washington.edu/
http://coamps.apl.washington.edu/COAMPS-bin/COAMPSOS_homepage.cgi

INTRODUCTION

The proposed work had two major thrusts, listed as two main tasks. Task 1 proposed to continue the development of three verification techniques, and to employ them to compare GFS-MM5, GFS-ARW, and NOGAPS-COAMPS on a data set consisting of several meteorological parameters. Task 2 put forth the idea of developing a statistical model representing the relationship between model parameters (e.g., diffusion rate) and macroscopic features in meteorological quantities (e.g., frontal speed). That work straddles the line between s*ensitivity analysis* and what is often called *model emulation*.

The previous report discussed the application of three verification techniques (Optical Flow, Cluster Analysis, and Variograms) to a data set spanning March 2008 – March 2009. In this report, we will present our latest work on the Optical Flow method applied to a longer data set (April 2008 – November 2009). Our initial article prepared to describe the Optical Flow work has now appeared on the Early Release website of Weather and Forecasting (Marzban and Sandgathe 2010), after drastic revisions brought about by the review process. A manuscript describing our work on Task 2 was submitted, but was rejected. We have critically examined the reviews and are in the process of revising and resubmitting the work. Some of the criticisms and our responses are mentioned in this report.

TASK 1

OPTICAL FLOW (OF): This method provides a diagnostic assessment of forecast error by decomposing it into intensity, and displacement errors, with the latter further decomposed into errors in magnitude and angle.

The quantity being verified here is the 24h sea-level forecasts from the three mesoscale models at 36km horizontal resolution. The MM5 and ARW forecasts are from the UW Atmos ensemble (http://www.atmos.washington.edu/mm5rt/) run on the UW Atmos computer systems and using the 1-degree resolution National Weather Service GFS global model for initial and boundary conditions. The COAMPS forecasts are generated at UW APL using the COAMPS-OS™ system run on APL computer systems and using the Navy NOGAPS global forecast system for initial and boundary conditions. Since different boundary conditions are employed for MM5 and ARW vs COAMPS, it is difficult to separate entirely the role of the global models from the mesoscale models; however, the difference between GFS-MM5 and GFS-ARW gives and indication of the impact of the global models on a short (24h) forecast. The 417 forecasts span the dates April 2008 – November 2009.

Figures 1, 2, and 3 all have the same format, and concisely capture the quality of the forecasts. The top panel shows the outline of the US, and a series of arrows (placed on every fourth grid point), suggesting the manner in which the forecast must be transformed to yield the analysis. In other words, the arrows display the optical flow field mapping the forecast to the observed field. The next three rows of figures show the three components of the forecast error: The intensity error, and the magnitude and angle of the displacement error (i.e. of the OF arrows). The left column shows the spatial distribution of the errors, and the right column shows the histogram of the corresponding error; the distribution of the angles is presented with Rose plots. Ideally, good forecasts would produce a spatial field of errors with no particular spatial pattern; the histogram of the intensity errors would be narrow and centered at zero, and that of the displacement (magnitude) would be narrow and clustered on zero. The distribution of the angles in the Rose plot would be uniform.

Optical flow patterns: The conclusions based on the analysis of this larger data set (417 cases) are mostly the same as those based on the smaller number of cases (277) reported previously. Over water, both ARW and COAMPS display predominately west or more correctly, random direction indicating no significant errors. ARW indicates something going on in the SW corner where the direction is SW. Since the plot is from forecast to observation, that implies the mean forecast is NE of the analyzed location. MM5 seems to have a much stronger S to SW bias over the ocean areas or mean forecast positions to the N or NE generally from the analyzed positions.

Over land, ARW and COAMPS have much more variability. COAMPS especially seems to exhibit a convergence toward the central Rockies. Since this is an average over most of a year, this indicates a substantial bias in the mountainous regions. MM5, while more variable over the ocean, exhibits less variability over land than either ARW or COAMPS.

Intensity error distribution: MM5 exhibits organized intensity error patterns indicating a systematic bias either in GFS or MM5. The errors are greatest along the coast of the US, predominately over land. The errors seem to indicate a terrain bias or land-sea bias as they decrease further inland. Errors in the NW and SW corners of the ocean region also indicate a possible systematic issue with assimilating boundary conditions. (Note: the outer five rows of the model predictions are not displayed due to interpolation across the model nests.) COAMPS exhibits two locations with significant systematic

errors – over the US Rockies centered on Colorado and over Alaska. This pattern is similar in some ways to the MM5 (and to be discussed) ARW errors; however, these seem to have specific maximum points instead of following the coastal topography. The ARW error distribution follows the terrain/coastal edge similar to MM5; however, without any issues on the western boundary of the region.

Intensity histograms: The intensity histograms are a little misleading since Forecast minus Observation is interpreted differently for a low pressure system (plus values mean under forecast or not as intense as observed) and a high pressure system (plus values mean forecast to high or more intense than observed), however, useful information is still presented. MM5 seems to have almost a normal distribution of error with a slight skew toward underforecasting and a range of -1.1 to 0.9. COAMPS is definitely skewed toward low values, and from the distribution plot, likely an issue with sea level reduction over mountains or with terrain. The range is -2.2 to 0.2 with the majority between -0.5 and 0.1. ARW also appears to be skewed; however, the range of values is much smaller: -0.6 to 0.6 with far fewer values in the positive range, again indicating a terrain/SLP reduction issue.

Displacement error: Reviewing the displacement error plots, COAMP and ARW have significantly less displacement error than MM5. The peak value on the histogram is approximately 0.3 for COAMPS and 0.5 for ARW yet both have similar histogram distributions and fairly random geographic distributions. Error distributions for MM5 indicate consistently higher values over land.

Angular error: Both MM5 and ARW exhibit a peak in the error histogram at approximately 230 degrees or, in polar coordinates, SW. This means the forecast is generally to the NE of the observation. What is curious about these values is that for both models, these peaks are very dominant and for ARW, the range of values is only from 170 to 280, but COAMPS displays a lesser degree of this error.

VARIABILITY OF DAILY ERRORS

An additional analysis was performed on the daily variability of the forecast errors. To that end, the distribution across the forecast domain, of each of the three components of error, for each of the 417 forecasts is summarized as a boxplot. Ideally, for perfect forecasts such a boxplot should be centered on zero. The spread of the boxplot reflects variability in the error due to the spatial variation of the errors, and so for perfect forecasts ought to be zero. The results for real forecasts are shown in Figures 4, 5, and 6, for the three models, respectively.

Intensity error distribution: Examining the intensity errors (top panels in Figures 4-6), it is apparent that there is no systematic error for any of the three models, because the boxplots are mostly centered around zero. There is a noticeable increase in the spread of the boxplots, in the latter portion of Autumn 2008 and early Winter 2009.

Displacement error: The aforementioned increase in the error spread is accompanied by an increase in the magnitude of the displacement errors (middle panels in Figures 4-6), but mostly for MM5. This is evident in the manner in which the boxplots in the mid-portion of the figure do not overlap zero. COAMPS displays this error more consistently across the entire span of 417 days. WRF displays this type of error as well, but only for the first half of the time series. It is possible that changes made to WRF around Winter 2009 corrected this error.

Angular error: The boxplots showing the distribution of these errors for each of the 417 forecasts are all centered around 180, consistent with the hypothesis that there is no systematic angular error on any given day, for any of the three models.

TASK 2

Sensitivity Analysis and Model Emulation: One reason for performing verification is to discover sources of error which can be corrected. Although most verification methods do offer some explanation regarding the errors, rarely (if ever) they propose means of correcting the errors. Often model parameters are utilized to correct some of the errors, but only to give rise to other undesirable and unforeseen errors. This portion of the project aims to develop a statistical model representing the relationship between model parameters and forecast features/errors. Such statistical models are often referred to as *emulators*. The project also aims to develop the inverse emulators, mapping forecast errors to model parameters.

Forward and inverse emulators were developed for COAMPS, relating two model parameters (sponge layer timing, and frequency of executing the radiation package) to two forecast quantities (sea-level pressure and boundary-layer height). The Lorenz 1963 model was also employed to demonstrate the methodology, all of which were submitted for publication in 2008. The paper was rejected mostly for reasons that we believe reflect misunderstanding on the part of the Reviewers. For example, the Reviewers claim that the proposed work is no different from performing sensitivity analysis. However, this is not true, for at least two reasons: 1) sensitivity analysis (in any of its multitude of forms) assesses **only the effect** of model parameters on forecasts. It is analogous to computing the correlation coefficient for assessing the correlation between two quantities. One cannot use the correlation coefficient for predicting one variable from the other. That requires a statistical model relating the two quantities, i.e., an emulator. Our proposal involves developing the emulator itself, not just assessing the effect of model parameters on forecast errors. 2) sensitivity analysis, as understood in meteorological circles is only a **local** technique. This is evidenced in the use of adjoint models, or other derivative-based models, for performing the analysis. There exists, however, an entirely different class – referred to as **global** sensitivity analysis, which is much more inline with the emulation approach. It examines the amount of variance in the forecast errors, which can be attributed to the various model parameters. For this reason, these techniques are also called variance-based; they do not require derivatives. As a consequence of the revisions to the first paper, we have also submitted a paper to the upcoming AMS meeting (in Seattle) which explains the difference between local and global sensitivity analysis; that talk has been accepted.

We intend to spend the remainder of the funded period to 1) revise and resubmit the initial paper, and 2) proceed with the development of the emulator for COAMPS. In parallel with the statistical ongoing work, an updated version of COAMPS is also being installed, because we hope it will allow us to vary more of the model parameters.

REFERENCES

Marzban, C. and S. Sandgathe 2010: Optical Flow for verification. Weather and Forecasting, 25(8), 1479-1494.

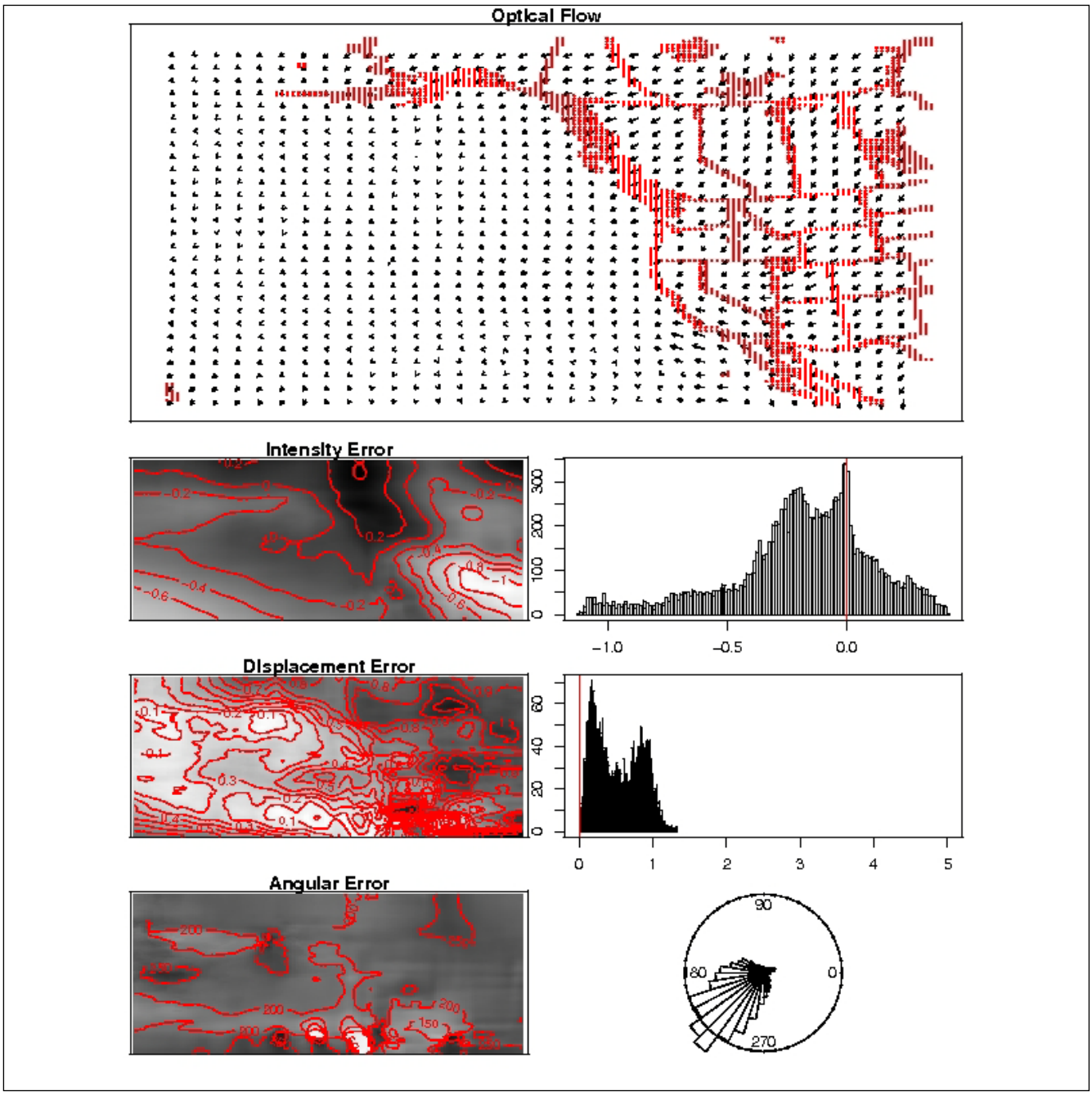

Figure 1: The Optical flow results for MM5.

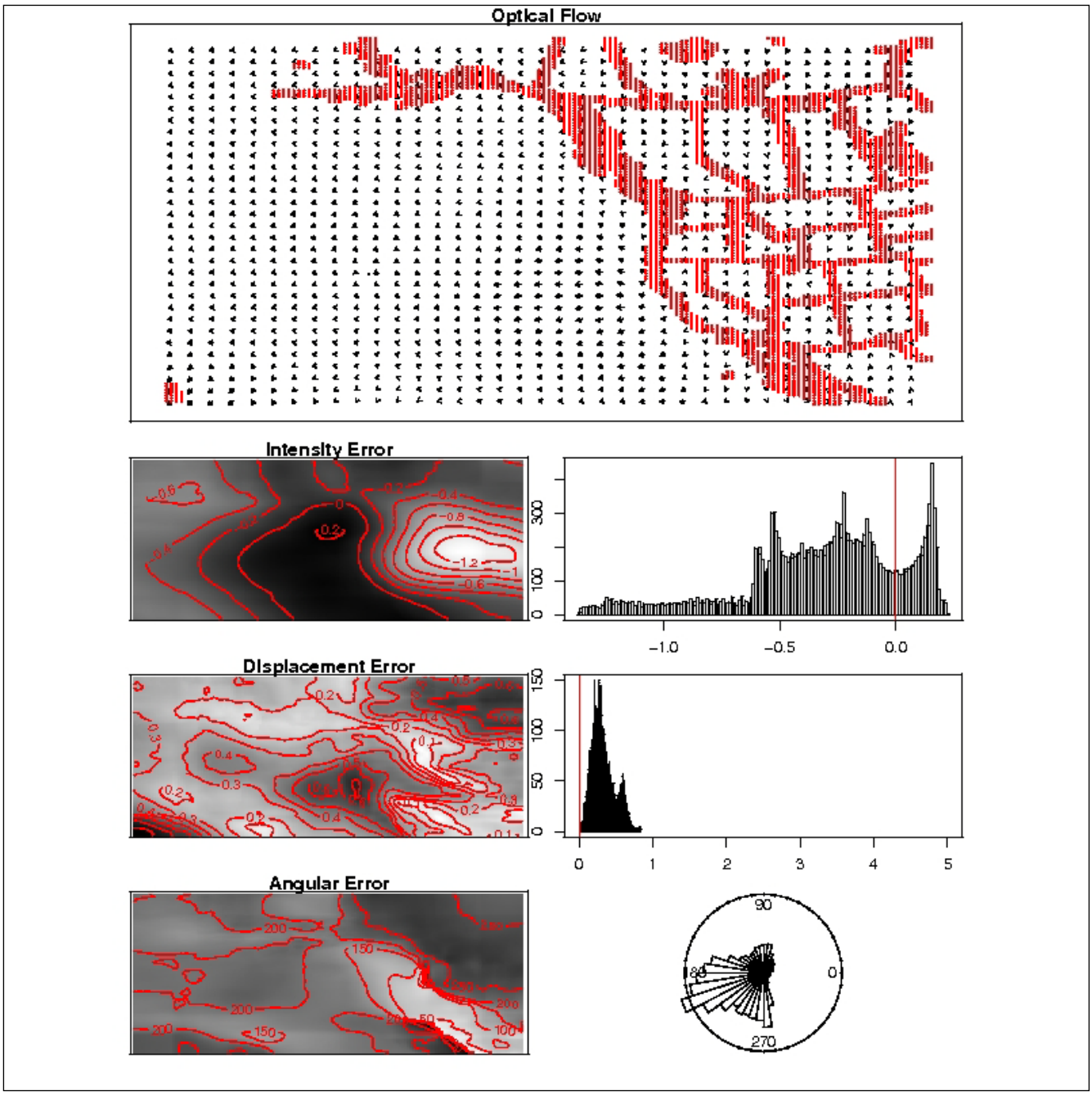

Figure 2. Optical flow results for COAMPS.

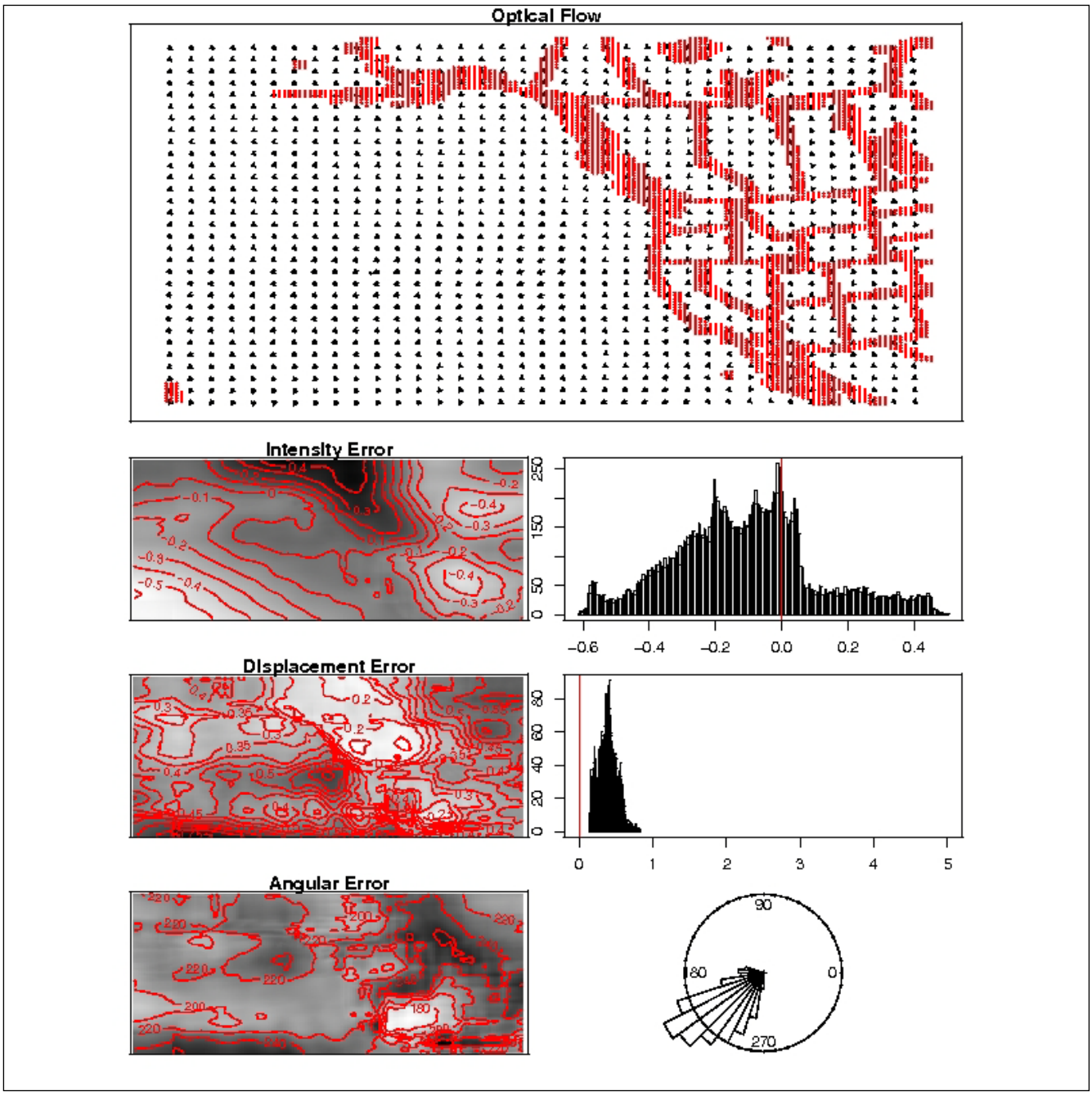

Figure 3. Optical flow results for ARW (WRF).

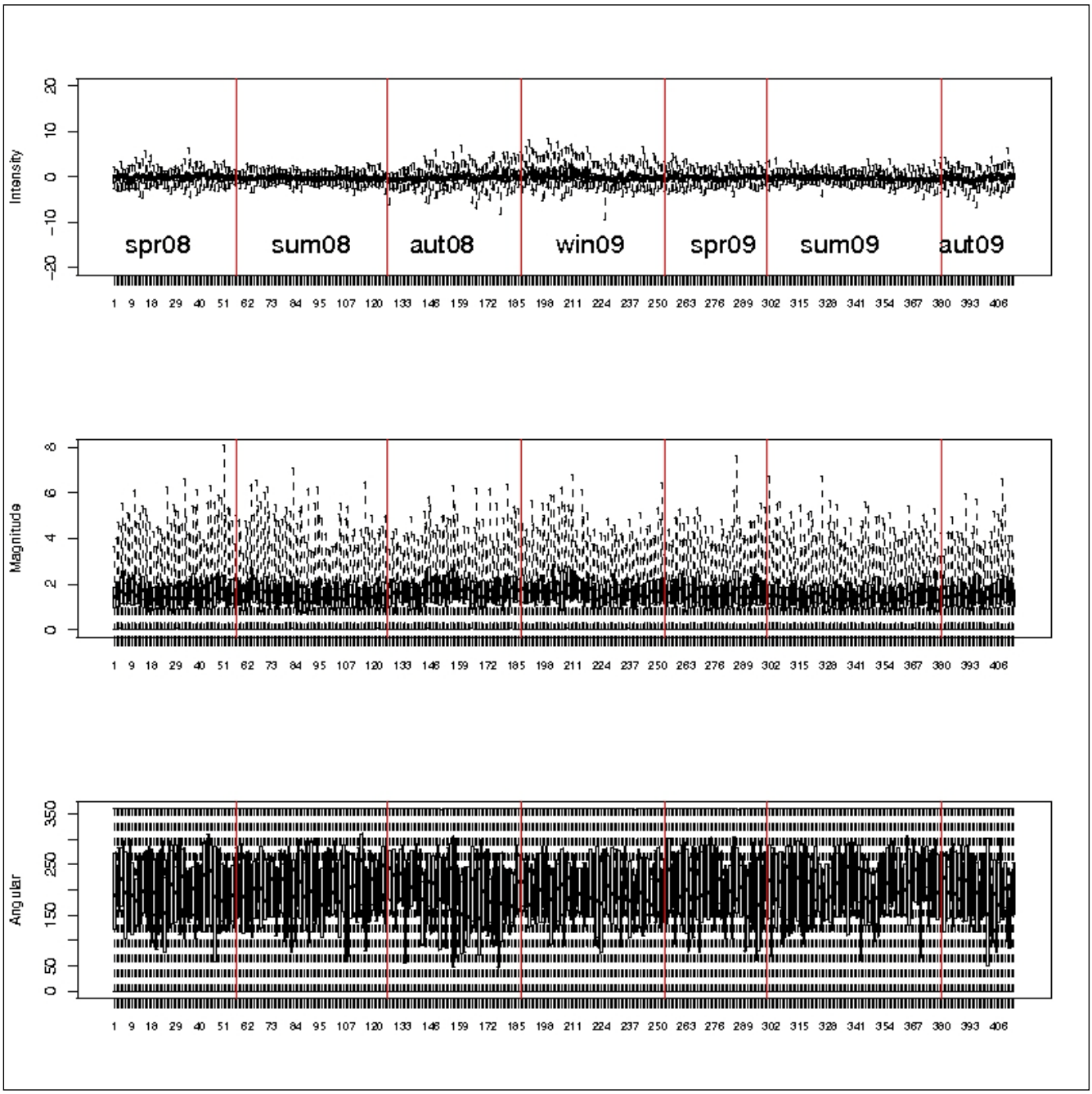

Figure 4. Distribution of three error components for each of 417 forecasts from MM5 (i.e., within-day variability). Also shown are the seasons.

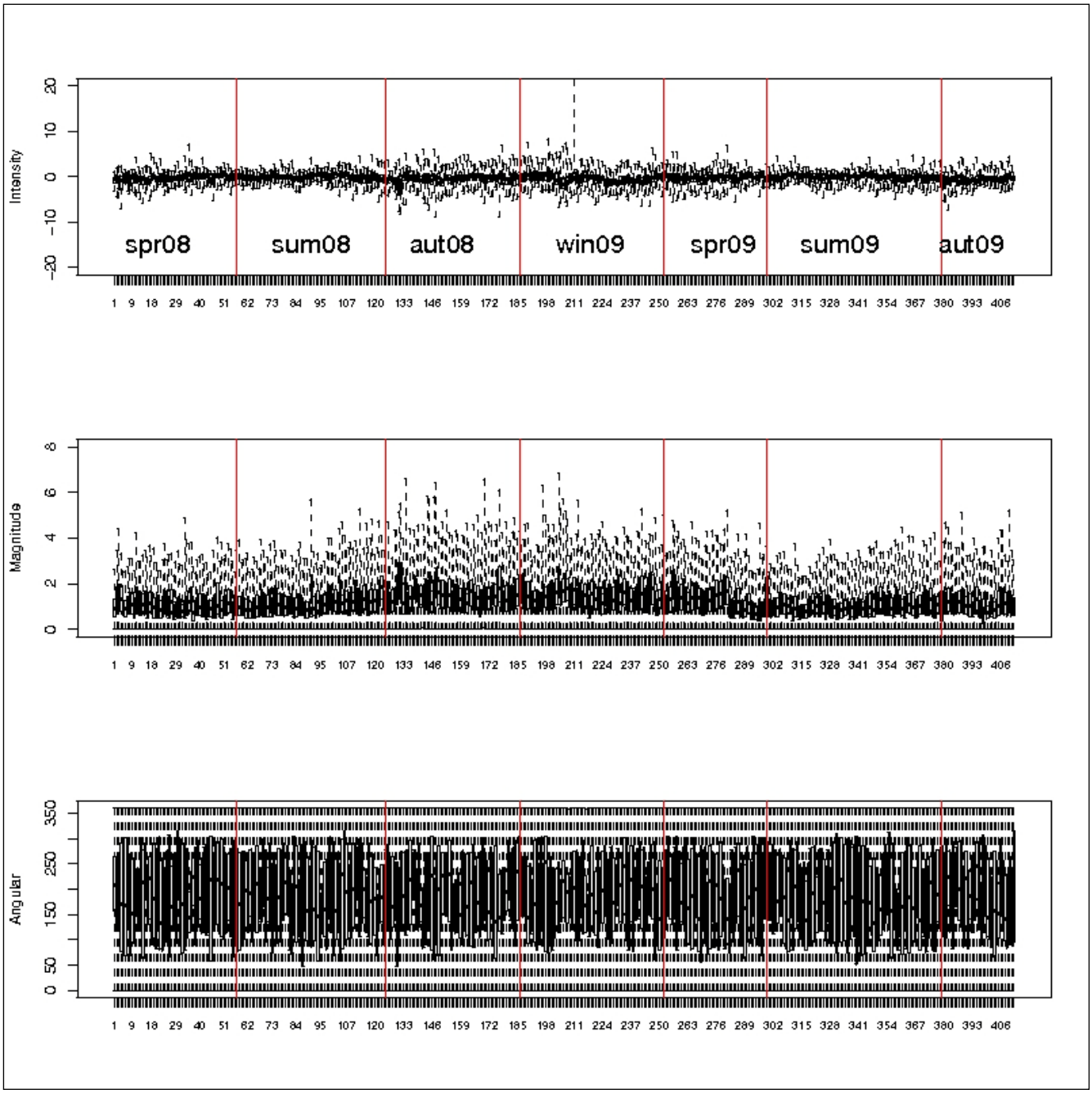

Figure 5. Distribution of three error components for each of 417 forecasts from COAMPS (i.e., within-day variability). Also shown are the seasons.

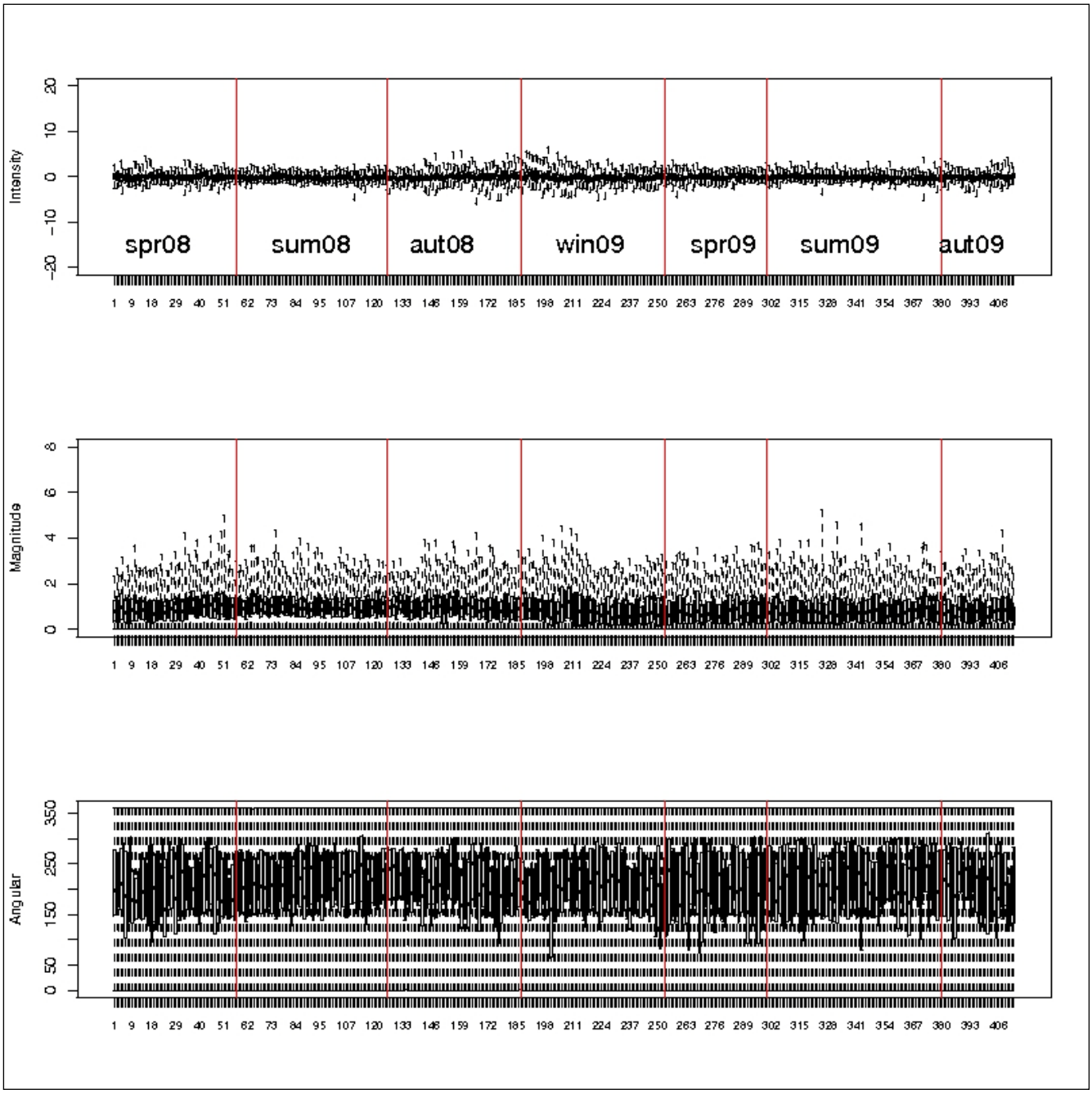

Figure 6. Distribution of three error components for each of 417 forecasts from WRF (i.e., within-day variability). Also shown are the seasons.