

# Thinning Satellite Data Using Wavelets for Weather Forecasting

Ross N. Hoffman, Christian Alcala and S. Mark Leidner  
Atmospheric and Environmental Research, Inc.  
131 Hartwell Avenue, Lexington, MA 02421

**Abstract**—Operational weather prediction centers use only a fraction of observations of the atmosphere and the earth’s surface that are made by satellite, *in situ*, and ground-based instruments. We are investigating the use of wavelet analysis to develop an adaptable selection method based on the local information content in a satellite data scene to determine the density of observations to use. This investigation supports and enhances Earth science capability by 1) improving the selection and impact of the vast, information-rich and valuable satellite observations of the Earth system, 2) combining mature technologies (atmospheric data assimilation and wavelet analysis) for a novel and practical use, and 3) raising the technology readiness level (TRL) of this technique to a working prototype in a realistic setting.

Our results to date show that wavelet-based selection is roughly equivalent to regular decimation to every 8<sup>th</sup> or 10<sup>th</sup> datum. Extracting information at the smallest spatial scales (25 and 50 km) required the development of a new noise thresholding approach because the signal-to-noise ratio is small. This new method is described and demonstrated.

## I. INTRODUCTION

Current practice at operational weather centers reduces today’s enormous volume of satellite data to practical levels by regular decimation, i.e., keeping every 2<sup>nd</sup> or 4<sup>th</sup> or 6<sup>th</sup> point along and across the satellite’s data swath. Our work investigates an adaptable approach to select data. We are using the continuous wavelet transform (CWT) [1] in two dimensions ( $x,y$ ) to identify features of interest for more informed data thinning. Retrieved wind speeds from NASA’s SeaWinds scatterometer on QuikSCAT [2] provide the input data for the thinning technique in our study. The amplitudes of the wavelet coefficients from two passes of the CWT (top-to-bottom and left-to-right) are summed to identify edges and gradients in the satellite-observed wind speed field. The CWT provides information on six spatial scales (25, 50, 100, 200, 400, 800 km), so features of interest can be identified on each of these scales. Once features are identified at each spatial scale, the satellite data are decimated to a density appropriate to the associated spatial scale. The final data selection is the union of all points selected at every scale. We test our thinning technique by data assimilation in atmospheric models with a 2d-variational method [3]. The baseline case assimilates all available data (ALL). This provides a “best” analysis since it uses all available data. But it is also computationally expensive. Experiments assimilating thinned subsets of the data by regular decimation and wavelet-based selection are evaluated for information content. Our goal is to retain as much information as possible in the data assimilation

analysis but by only using 3-5% of the data through wavelet-based selection.

## II. WAVELET ANALYSIS APPLIED TO SATELITE WIND SPEEDS

### A. Wavelet-based selection

Our data-thinning algorithm uses the CWT to obtain a measure of the local information content of the data. This approach was first developed in the context of edge detection [4]. In particular, the CWT has several important advantages over the more commonly used discrete wavelet transform (DWT). Both transforms provide a complete and invertible representation of the data. However, while the DWT uses a set of orthogonal wavelet bases to obtain the most compact representation of the image, useful for data compression, the CWT uses a set of nonorthogonal wavelet frames to provide a highly redundant representation. The consequence of this redundancy is that the CWT gives a wavelet coefficient at each analysis scale for each pixel in the image, allowing us to characterize the local information content. In addition, this redundancy improved the stability of the reconstruction (inverse transform) in the presence of noise. Using a wavelet based on the conventional Canny edge detector allows us to simultaneously detect, localize, and characterize the edges in the observation data. Following the approach described by Mallat [1], we first apply the CWT using a wavelet oriented across the satellite track and then repeat using a wavelet oriented along the satellite track. We then take the coefficients from these two-orthogonal directions to produce a wavelet amplitude and phase for each pixel at each scale.

After analyzing the data with the CWT, we have developed the following technique to identify the features of interest and select the data points along these features:

1. Identify the wavelet transform modulus maxima (WTMM) for each scale using the wavelet magnitude and phase.
2. Create wavelet maxima chains from the WTMM for each scale by comparing nearest neighbors.
3. Perform wavelet noise reduction to eliminate extraneous WTMM chains.
4. Connect remaining wavelet maxima chains through scales to create wavelet ridge (signal skeleton).
5. Select data points along wavelet ridge lines.

Because noise introduces multiple small-scale edges within the data, we have developed a technique for distinguishing between WTMM produced by signal and WTMM produced by noise at these scales. Simple thresholding does not work since both types of WTMM are of equal magnitude.

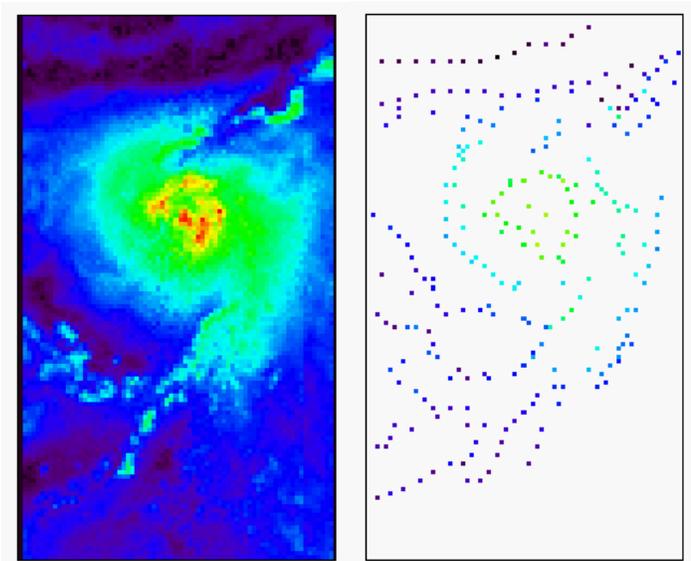


Fig. 1. SeaWinds wind speeds (left) and selected wind speeds (right) for Typhoon Meranti at 1800 UTC 7 August 2005. Wind speeds from 0 to 30 m/s are shown as colors from purple to red.

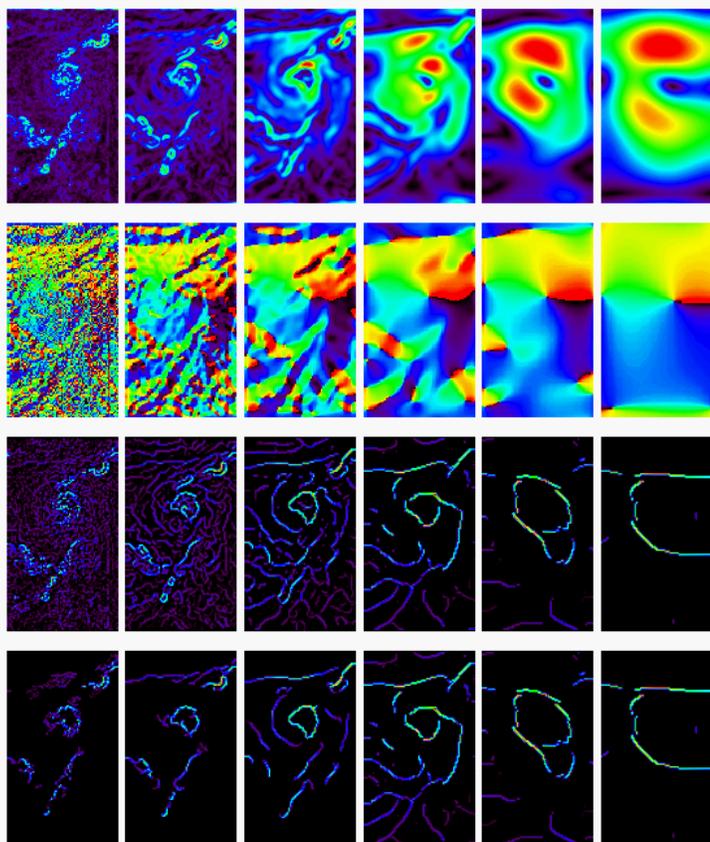


Fig. 2. Two-dimensional wavelet analysis of Typhoon Meranti ocean surface wind speeds (from Fig. 1). See text for detailed description.

However, the WTMM associated with actual geophysical structure will track down from larger to smaller scales. Therefore, we first used an approach similar to the scale multiplication method of Zhang and Bao [5] to partition the signal from the noise. We have developed a more robust technique for partitioning signal and noise using a Bayesian method (described below in Section II.B).

In Figs. 1 and 2, we apply our data-thinning algorithm to an example of satellite-observed ocean-surface wind speeds obtained during Typhoon Meranti, 1800 UTC 7 August, 2004. Figure 1 shows the satellite image in the left panel and the results of our analysis in the right panel. The number of data values selected by the algorithm for this example corresponds to about 2.5% of the total number of data points. Figure 2 shows the intermediate results generated by the wavelet analysis for the image in Fig. 1. The top two rows in Fig. 2 plot the magnitude and phase of the two-pass CWT operation for the scales of 25, 50, 100, 200, 400, and 800 km. The bottom two rows show the results of the data-thinning algorithm before data location selection both without and with the additional noise-partitioning step for all of the scales. By comparing these two rows we see a significant reduction in number of WTMM ridges in the smallest three scales (25, 50, and 100 km).

After noise thresholding, we compare the remaining WTMM chains for each scale with those from the next largest scale. We retain only those WTMMs that track down from the larger scale. In order to track down, the WTMM must be located to within a scale size neighborhood of the position of the larger scale WTMM. The complete set of WTMM cascading from large scales to small scales are referred to as the wavelet ridges and form the basic skeleton of the signal. We then select points by either using a scale-based regular selection method or by using the maxima in the WTMM following the technique of Arneodo[6]. The regular selection technique retains points spaced a scale-size distance apart along the WTMM chains for each scale size.

### B. Bayesian partitioning of signal and noise

We found that the multiplication method of partitioning signal and noise requires scale specific thresholds and is not adaptive. This is problematic for the smaller spatial scales where signal and noise are often have similar magnitudes. To accomplish noise removal on all scales without specifying thresholds, we developed a Bayesian approach. First, we calculate square modulus from wavelet components for each scale,

$$M_j^2[m, n] = |d_j^H[m, n]|^2 + |d_j^V[m, n]|^2$$

where  $d_j^H$  and  $d_j^V$  are the wavelet coefficients for the  $j^{\text{th}}$  scale at each point in a domain of dimensions  $[m, n]$  from horizontal ( $H$ ) and vertical ( $V$ ) analyses of the image. Next, we model

the squared modulus as a two-component Gamma distribution using the Nelder-Mead simplex method,

$$p_{M_j^2}(x) = w_0 \frac{1}{\lambda_0} e^{-x/\lambda_0} + w_1 \frac{1}{\lambda_1} e^{-x/\lambda_1}$$

where  $\lambda_0$ ,  $\lambda_1$  and  $w_0$ ,  $w_1$  are computed parameters of the fitted distributions. Finally, we calculate posterior probability that each wavelet coefficient is significant using Bayes Method. Coefficients that fall within the “signal” distribution are retained.

Figure 3 shows an example of this noise partitioning for two spatial scales. Wavelet analysis of the image of Typhoon Meranti wind speeds (left) generates the squared modulus images at scales of  $2^3$  and  $2^4$  (middle panels). The distribution of the squared modulus values are modelled by two-component Gamma distributions (right panels). At each scale and for any point in the image, that point either falls within the “signal” distribution (left sides of the histograms) or the “noise” distribution (tail of the histograms). Bayes method determines the posterior probability that each point falls within one or the other distribution.

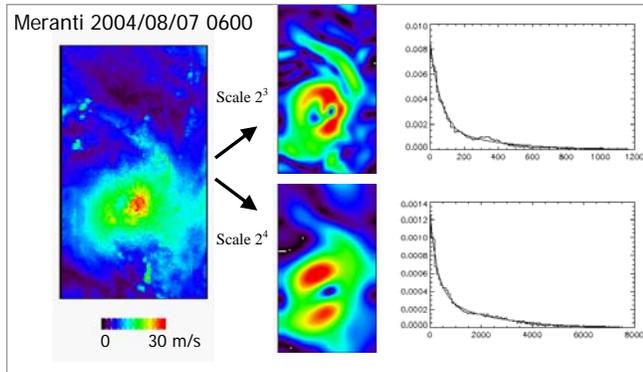


Fig. 3. Noise thresholding on two spatial scales using fitted two-component distributions.

Figure 4 shows an example of noise thresholding on all scales using this technique for the Meranti image in Fig. 3.

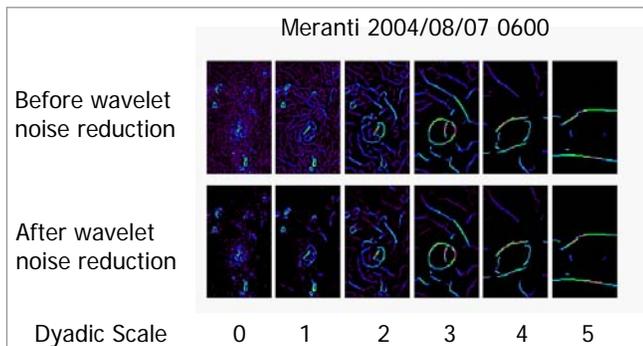


Fig. 4. Wavelet noise reduction for the image in Fig. 3.

### III. EVALUATION OF THINNED AND DECIMATED SWS DATA

We use atmospheric data assimilation as a tool to evaluate the efficacy of wavelet-selected points versus other thinning techniques. In this study we compare wavelet-thinning to regular decimation (i.e., every 2<sup>nd</sup>, 4<sup>th</sup>, 8<sup>th</sup>, etc. observation). We use the following test: if an atmospheric analysis using wavelet-selected observations is improved compared to an atmospheric analysis using regularly decimated data (given a comparable fraction of observations are retained by each technique), the wavelet-selected points are superior. In the future, we will also conduct forecast experiments to evaluate the effects of data thinning strategies on forecast skill. In this paper, we report only on the results of data assimilation experiments.

We conducted a search for appropriate cases for data assimilation and forecast experiments. We want to isolate the impacts of data thinning from other contaminating influences as much as possible, so we have adopted a case selection strategy that only considers satellite data within ~10 minutes of synoptic times (i.e., 00, 06, 12, and 18 UTC). This minimizes the time difference between satellite observations and available global analyses of the surface wind field so that meteorological features of interest are aligned as closely as possible. We selected three cases that represent a wide range of meteorological conditions and have sufficient observations at subsequent times for validating the mesoscale forecasts. One of the cases has already been presented in Fig. 1 (Typhoon Meranti, August 2004). The other two are cases are an anticyclone in the South Indian Ocean (October 2004) and a very light wind case in the tropical Pacific (March 2005). In this paper, we will only report on results for Typhoon Meranti.

The design of data assimilation experiments includes twelve treatments. The names of the treatments indicate which data are included in the assimilation:

- 1) ALL
- 2) THIN2 every 2<sup>nd</sup> datum
- 3) THIN4 every 4<sup>th</sup> datum
- 4) THIN6 etc.
- 5) THIN8 ...
- 6) THIN10 ...
- 7) WAVELET scale-based, regular selection along WTMM chains
- 8) WAVELET4 = WAVELET + THIN4 points
- 9) WAVELET6 = WAVELET + THIN6 points
- 10) WAVELET8 etc.
- 11) WAVELET10 ...
- 12) WTMMM points selected using WTMM maxima

We face one difficult problem in conducting data assimilation experiments: the true atmospheric state is not known. The atmospheric state from global models and measured by satellites both contain errors that are likely to be correlated in space and time. These errors cannot be known completely, since we do not know the true atmospheric state. To eliminate the uncertainties associated with real

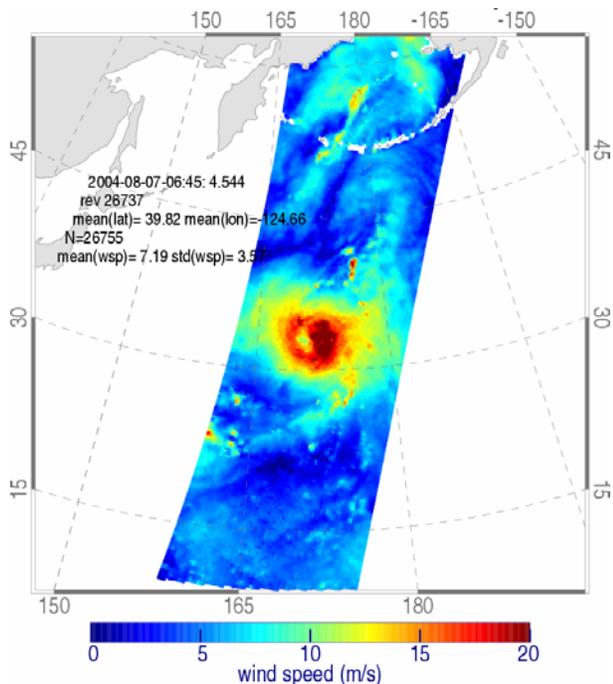


Fig. 5. Ocean surface wind speed around Typhoon Meranti as seen by QuikSCAT.

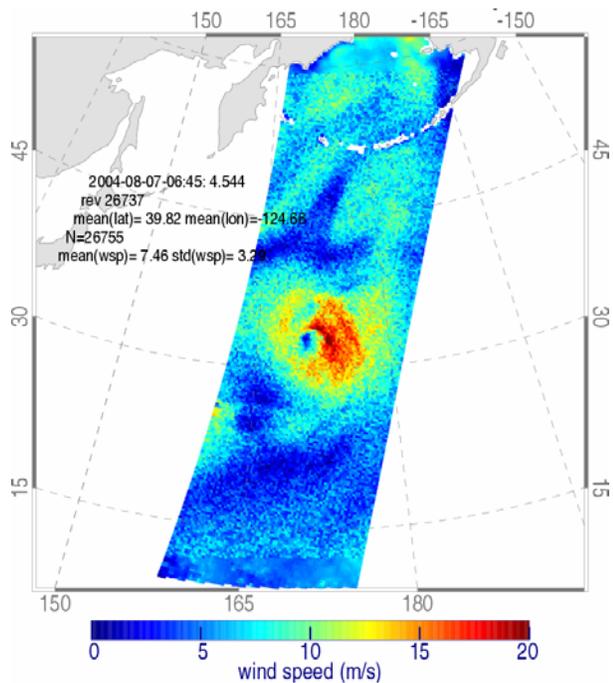


Fig. 6. Simulated ocean surface wind speed from a weather forecast model nature run (6-hour forecast, in this case), sampled at QuikSCAT data locations, plus 2 m/s of white noise to account for observation noise.

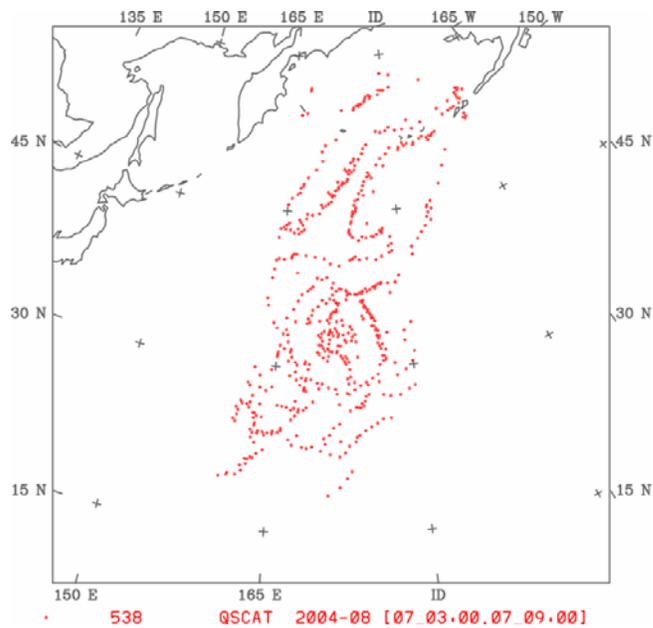


Fig. 7. The location of simulated observations selected by scale-based, regular point selection along WTMM chains.

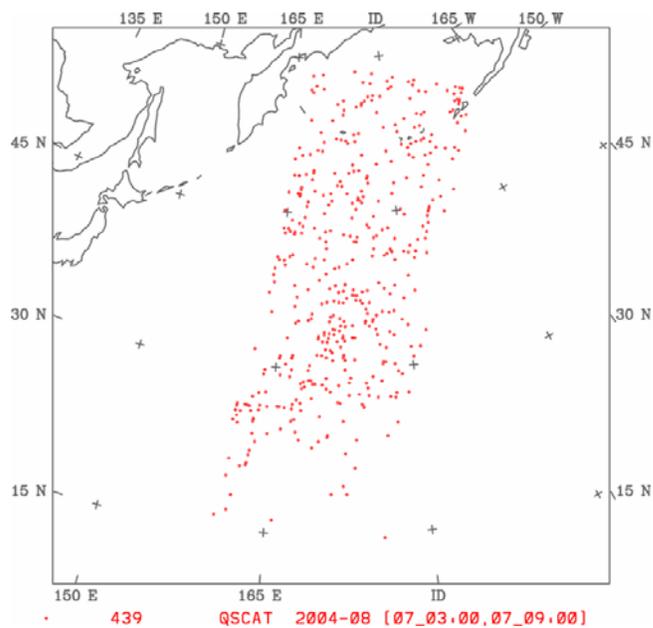


Fig. 8. As in Fig. 7 but points are selected using the WTMM maxima (see description at the end of Section II.A).

observations and global model estimates of the atmospheric state, and to assess the impact of each treatment as precisely as possible, we generate a “true” atmospheric state (or “nature run”) using a weather forecast model, and simulate “true” observations by sampling from our true atmosphere. To make the simulated observations more realistic, we add random noise, since real observations have noise from a variety of sources. Figure 5 shows the observed QuikSCAT wind speeds and Fig. 6 shows the simulated observations (with added noise) generated from a 6-hour nature run forecast. We use the Weather Research and Forecasting (WRF) model to generate the “true” atmosphere. The model domain is 201 x 201 x 30 grid points with a horizontal spacing of 27 km and a model top of 50 hPa. Figures 7 and 8 show the locations of observations selected by wavelet-based methods WAVELET and WTMMM, respectively. Notice that both wavelet point selection methods choose about 3.5% of all observations (N=15187).

For our data assimilation experiments, we use the WRF 3-dimensional variational (3dVAR) assimilation system, described by Barker et al. [7]. We are using background error covariances developed by Wu et al. [8]. The characteristic horizontal and vertical spatial scales of the assimilation system have been demonstrated using single observation tests (not shown) of surface winds and 500 hPa temperatures. The system’s response to these single observations is appropriate for our 27 km-resolution model domain.

Figure 9 shows the verification of the twelve treatments for the 0600 UTC 7 August 2004 QuikSCAT overpass of Typhoon Meranti. An atmospheric analysis is generated for each treatment using the appropriate subset of observations, plus a WRF forecast field 6 hours away from the analysis time. We use an off-time background field so the simulated observations and background field have a spatial mismatch that 3dVAR attempts to correct. Note that the number of observations assimilated for each treatment is shown in the legend of Fig. 9. Then each analysis is compared to the full set of simulated observations (N=15187) to evaluate how closely the analysis fits the observations.

Using all of the data (ALL) produces the best analysis, as expected. Also, the accuracy of the THINx analyses generally degrades as more data are thinned, as expected. The analyses from the two treatments that rely solely on wavelet-selected data, however, do not have nearly the accuracy that one might expect based on the number of observations used. Given that the two wavelet selection methods, WAVELET and WTMMM, retain ~450-550 of the 15,000+ observations, one would expect that the atmospheric analyses using the wavelet-selected data should be at least as accurate as the THIN6 analysis (~450 observations). Yet the wavelet-only analyses are comparable to THIN8 or THIN10 analysis accuracies.

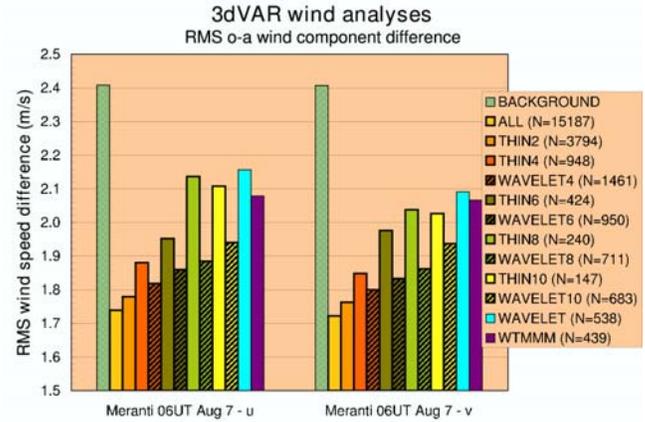


Fig. 9. Verification of data assimilation experiments. u- and v-components of the wind are evaluated separately.

#### IV. SUMMARY

We have developed and tested wavelet-based data thinning methodologies. While the wavelet techniques can clearly identify features of meteorological interest, we are not yet realizing improved efficiency or analysis accuracy by using the wavelet-selected data. The current basis functions used in our wavelet techniques are designed to identify gradients. Perhaps another basis function that identifies local maxima/minima could be used to augment or replace the current basis function.

#### ACKNOWLEDGMENT

We thank NASA’s Earth Science Technology Office (ESTO) for supporting this research (award # AIST-QRS-04-3019).

#### REFERENCES

- [1] Mallat, S., 1998: *A wavelet tour of signal processing*, 2<sup>nd</sup> edition, Academic Press.
- [2] Spencer, M. W., C. Wu, and D. G. Long, 2000: Improved resolution backscatter measurements with the SeaWinds pencil-beam scatterometer. *IEEE Trans. Geosci. Remote Sens.*, **38**, 89–104.
- [3] Hoffman, R. N., S. M. Leidner, J. M. Henderson, R. Atlas, J. V. Ardizzone, and S. C. Bloom, 2003: A two-dimensional variational analysis method for NSCAT ambiguity removal: Methodology, sensitivity, and tuning. *J. Atmos. Oceanic Technol.*, **20**, 585–605.
- [4] Mallat, S. and S. Zhong, 1992: Characterization of signals from multiscale edges, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, 710–732.
- [5] Zhang, L. and P. Bao, 2002: Edge detection by scale multiplication in the wavelet domain, *Pattern Recognition Letters*, **23**, 1771–1784.
- [6] Arneodo, A. and N. Decoster and S.G. Roux, 2000: A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and

anisotropic random rough surfaces, *Eur. Phys. J. B*, **15**, 567–600.

- [7] Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results, *MWR*, **132**(4), 897–914.

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