Stochastic Multiresolution Models for Turbulence

Brandon Whitcher* Jeffrey B. Weiss† Douglas W. Nychka* Timothy J. Hoar*

August 6, 2002

Abstract

The efficient and accurate representation of two-dimensional turbulent fields is of interest in the geosciences because the fundamental equations that describe turbulence are difficult to handle directly. Rather than extract the coherent portion of the image from the background variation, as in the classical signal plus noise model, we present a statistical model for individual vortices using the non-decimated discrete wavelet transform. A template image, supplied by the user, provides the features we want to extract from the observed field. By transforming the vortex template into the wavelet domain specific characteristics present in the template, such as size and symmetry, are broken down into components associated with spatial frequencies. Multivariate multiple linear regression is used to fit the vortex template to the observed vorticity field in the wavelet domain.

Key words. multivariate multiple linear regression, non-decimated discrete wavelet transform, penalized likelihood, turbulence, vorticity.

*Geophysical Statistics Project, National Center for Atmospheric Research, Boulder, Colorado 80307-3000. E-mail: *whitcher@ucar.edu*, *nychka@ucar.edu*, *thoar@ucar.edu*.

†Program in Atmospheric and Oceanic Sciences, Department of Astrophysical, Planetary and Atmospheric Sciences, University of Colorado at Boulder, Boulder, Colorado 80309. E-mail: *jweiss@colorado.edu*.

1 Introduction

The efficient and accurate representation of two-dimensional turbulent fields is of interest in the geosciences because the fundamental equations that describe turbulence are difficult to handle directly. Recent advances in computing allow scientists to produce realistic simulations of turbulent flows. The new emphasis of research in turbulence is on the individual coherent structures in the flow instead of the signal versus background noise. Our approach to this problem is to develop a flexible model for a single coherent structure (vortex) using a multiresolution analysis and then efficiently represent the entire vorticity field through a collection of these structures. Defining the vortex utilizes specific information from the scientist via a template function, but is flexible enough to capture a broad range of features associated with observed coherent structures. Modeling the vorticity field is done using classic regression methodology, allowing the information in the observed data to dictate what is and what is not a vortex. This paper concentrates on the computational algorithms used in our methodology and is a companion to Whitcher *et al.* (2002).

Wavelet-based analysis of two-dimensional turbulence has been performed previously; see, for example, Farge *et al.* (1992), Wickerhauser *et al.* (1994), Wickerhauser *et al.* (1997) and the summary article by Farge *et al.* (1996). The focus of the methodology used in these papers has a signal processing orientation; that is, the wavelet transform discriminates between "signal" and "noise" using compression rate as a classifier. The idea being that the coherent structures (vortices) observed in the turbulent flow will be well represented by a few wavelet coefficients and therefore most coefficients may be discarded. From one vorticity field, two fields are produced from such a procedure: one based on the largest wavelet coefficients that is meant to capture the coherent structures in the original field, and another based on the remaining wavelet coefficients which captures the filament and background structure. Our methodology differs significantly from the signal-plus-noise model, in that we formulate a statistical model for an individual coherent structure and extract a fixed number of isolated vortices from the original field. From these estimates a completely different set of summary statistics may be calculated; for example, instead of global summary statistics like the enstrophy spectrum we are able to look at local statistics for each structure such as the average size, amplitude, circulation and enstrophy.

Previous work in identifying vortices in a turbulent flow has produced a wide variety of methods. We summarize two of them here. Let $\zeta(\mathbf{x})$ define the vorticity at pixel $\mathbf{x} \in \mathbb{R}^2$. A simple approach to isolating individual structures within the vorticity field is to define a vorticity threshold δ. Any

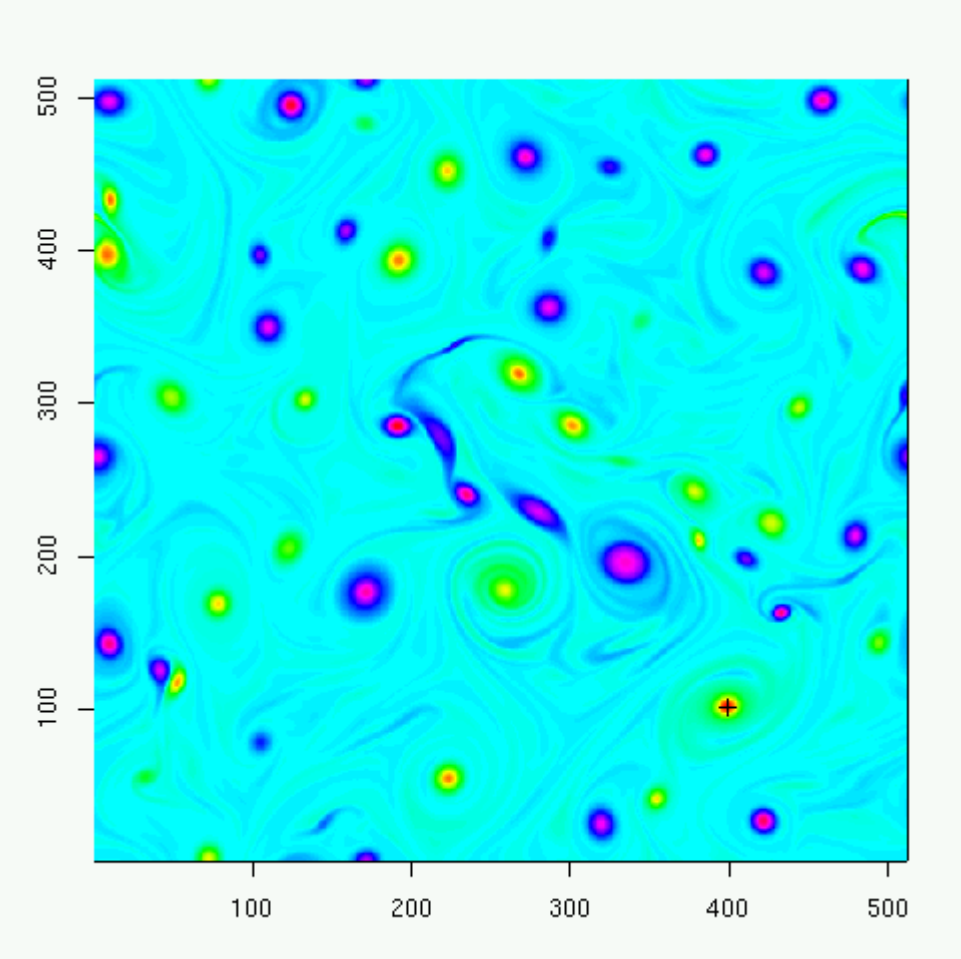

Figure 1: Simulated vortex field (512×512) at time $t = 15$. The gray scales are artificial and were converted from the spectrum of visible light. Negative vortices are seen as white with dark gray centers; while positive vortices are seen as black with light gray centers. Zero vorticity is a uniform light gray.

observed values such that $\zeta(\mathbf{x}) > |\delta|$ would describe a region of "anomalously high vorticity" and, hence, a coherent structure.

There are several problems with a thresholding approach. Coherent structures are stripped of any vorticity beneath the threshold thus introducing bias and reduced variability into vortex statistics, the threshold must be adjusted *a posteriori* from time step to time step whereas we would prefer a method that adapts to the set of coherent structures in each individual image, and lastly there is no clear way to generalize the technique to more complex problems (e.g., multipole or noisy structures).

An interesting method developed in the late 1980s is the *objective observer* approach of McWilliams

(1990). The method is as follows: determine the pixel with the largest (in magnitude) vorticity, descend/ascend the peak in vorticity in the positive x-direction until a vortex boundary is encountered, record this point and search for additional vortex boundary points at the base of this coherent structure until a closed curve is produced. The set of pixels found in this way define a boundary for the vortex. Repeat this procedure for all local minima/maxima in the observed vorticity field. The limitations of this procedure is that it depends too heavily on properly defining what is and what is not a coherent structure.

A wavelet-packet census algorithm was developed by Siegel and Weiss (1997) where the discrete wavelet packet transform is applied to simulated turbulent flow. The first step of their procedure is an iterative denoising method that partitions the wavelet packet coefficients into two groups – large and small coefficients. The large wavelet packet coefficients correspond to the coherent portion of the observed vorticity field and the small wavelet coefficients correspond to the disorganized, background information. On the coherent portion, wavelet coefficients are grouped together through their ideal support in the spatial domain to form individual coherent structures. The wavelet-based method presented here is an attempt to replace the *ad hoc* and iterative schemes of Siegel and Weiss (1997) with sound statistical procedures.

Section 2 introduces the two-dimensional multiresolution analysis of an image and its decomposition into unique spatial scales and directions. Section 3 outlines our modeling strategy by discussing vortex template functions and models for both single and multiple coherent structures. Section 4 provides the methodology we use to estimate a single coherent structure and also multiple structures in the vorticity field. Section 5 examines how well the technique performs at estimating multiple coherent structures in a sample vorticity field.

2 The Two-Dimensional Discrete Wavelet Transform

2.1 Multiresolution Analysis

To simplify the vortex dynamics within a given realization of a vorticity field, we rely on the two-dimensional discrete wavelet transform (2D DWT) and assume the vorticity field $\zeta(\mathbf{x})$ has finite energy; i.e., $\int \zeta^2(\mathbf{x})d\mathbf{x} < \infty$. Let ϕ be a scaling function and ψ be the corresponding wavelet generating an orthonormal basis on $L^2(\mathbb{R})$ and define the three separable wavelets

$$\psi^{\mathrm{h}}(x_1, x_2) = \phi(x_1)\psi(x_2), \quad \psi^{\mathrm{v}}(x_1, x_2) = \psi(x_1)\phi(x_2), \quad \psi^{\mathrm{d}}(x_1, x_2) = \psi(x_1)\psi(x_2), \tag{1}$$

corresponding to horizontal, vertical and diagonal directions, respectively. More information concerning the interpretation of "spatial directions" captured by the separable wavelet transform may be found in Section 2.2. The separable scaling function $\phi(x_1, x_2) = \phi(x_1)\phi(x_2)$ is associated with the approximation space. For $m \in \{\mathrm{h}, \mathrm{v}, \mathrm{d}\}$, let

$$\psi^m_{j,k,l}(x_1, x_2) = \frac{1}{2^j}\psi^m\left(\frac{x_1 - 2^j k}{2^j}, \frac{x_2 - 2^j l}{2^j}\right),$$

such that

$$\left\{\psi^{\mathrm{h}}_{j,k,l}(x_1, x_2), \psi^{\mathrm{v}}_{j,k,l}(x_1, x_2), \psi^{\mathrm{d}}_{j,k,l}(x_1, x_2)\right\}_{(j,k,l)\in\mathbb{Z}^3}$$

is an orthonormal basis of $L^2(\mathbb{R}^2)$. The triplet (j, k, l) denotes the jth wavelet scale at the spatial position (k, l). In practice the depth of the 2D DWT is determined by the finite dimension $M \times N$ of $\zeta(\mathbf{x})$, such that $1 \le j \le J = \lfloor \log_2 \min\{M, N\} \rfloor$. Efficient implementation of the separable 2D DWT is performed via a series of convolutions and subsampling (Mallat 1998, Sec. 7.7).

Whereas it would be most efficient to represent the vorticity field using a decimated transform such as the 2D DWT, we find it advantageous to utilize the 2D maximal overlap DWT (2D MODWT). Unlike the orthonormal 2D DWT, the 2D MODWT produces a redundant non-orthogonal transform. The reason for this discrepancy is that the 2D MODWT does not subsample in either dimension, it only filters the original image. The advantages are that the transform is translation invariant to integer shifts in space and it reduces artifacts caused by the choice of a specific wavelet filter.

Let $\bar{\phi} = \phi/\sqrt{2}$ and $\bar{\psi} = \psi/\sqrt{2}$ be rescaled versions of the scaling and wavelet functions. The three separable wavelet functions associated with spatial directions $m \in \{\mathrm{h}, \mathrm{v}, \mathrm{d}\}$ are now

$$\bar{\psi}^m_{j,k,l}(x_1, x_2) = \frac{1}{2^{2j}}\psi^m\left(\frac{x_1 - k}{2^j}, \frac{x_2 - l}{2^j}\right).$$

Hence, each level in the transform will have the same spatial dimension as the original field ($M \times N$) and represent a redundant set of wavelet coefficients. The 2D MODWT begins with the original vorticity field $\zeta(\mathbf{x})$, and at all scales we denote $\bar{v}_j(k, l) = \langle \zeta, \phi_{j,k,l}\rangle$ and $\bar{w}_j(k, l) = \langle \zeta, \psi^m_{j,k,l}\rangle$ for $m \in \{\mathrm{h}, \mathrm{v}, \mathrm{d}\}$, where $\langle x, y\rangle$ is the two-dimensional inner product of $x(\cdot, \cdot)$ and $y(\cdot, \cdot)$. The vorticity field $\zeta(\mathbf{x})$ may now be decomposed into $3J + 1$ sub-fields,

$$\left[\left\{\bar{w}^{\mathrm{h}}_j, \bar{w}^{\mathrm{v}}_j, \bar{w}^{\mathrm{d}}_j\right\}_{1\le j\le J}, \bar{v}_J\right]; \tag{2}$$

three fields of wavelet coefficients corresponding to distinct spatial directions and one field containing the scaling coefficients at the final level. The scaling (approximation) field for level j may be

obtained from the four fields at level $j+1$ via

$$\bar{v}_j(k,l) = \langle \bar{v}_{j+1}, \phi^*_{j,k,l} \rangle + \sum_m \langle \bar{w}^m_{j+1}, \psi^{m*}_{j,k,l} \rangle. \tag{3}$$

The 2D MODWT *multiresolution analysis* (MRA) of the vorticity field is an additive decomposition given by recursively applying (3) over all j; i.e.,

$$\zeta(x_1,x_2) = \langle \bar{v}_J, \phi^*_{J,x_1,x_2} \rangle + \sum_{j=1}^{J} \sum_m \langle \bar{w}^m_j, \psi^{m*}_{j,x_1,x_2} \rangle = \bar{a}_J(x_1,x_2) + \sum_{j=1}^{J} \sum_m \bar{d}^m_j(x_1,x_2),$$

where $\bar{a}_J(x_1,x_2)$ is the wavelet approximation field and $\bar{d}^m_j(x_1,x_2)$ is the wavelet detail field associated with the spatial direction $m \in \{\mathrm{h},\mathrm{v},\mathrm{d}\}$. The MRA of $\zeta(\mathbf{x})$ provides a convenient way of isolating features at different scales and directions with coefficients in the spatial domain versus the wavelet domain. This is advantageous since reconstruction is now reduced from the full inverse 2D MODWT to simple addition.

2.2 Spatial Directions of the 2D MRA

Estimating the coefficient associated with a basis function is equivalent to a filtering operation on the original image. The qualitative features of these filters help to interpret the wavelet coefficients. Two-dimensional wavelets (denoted by ψ^{h}, ψ^{v}, and ψ^{d} in (1)) are associated with horizontal, vertical and diagonal features in the vorticity field. This follows from the fact that the two-dimensional filters are the outer product of two one-dimensional filters, where the scaling filter averages (smooths) across its spatial direction and the wavelet filter differences across its spatial direction. The two-dimensional wavelet $\psi^{\mathrm{h}}(x_1,x_2) = \phi(x_1)\psi(x_2)$ will therefore smooth across the first dimension (x_1-axis) and difference across the second dimension (x_2-axis), thus favoring horizontal features. The two-dimensional filter $\psi^{\mathrm{v}}(x_1,x_2) = \psi(x_1)\phi(x_2)$ differences across the x_1-axis and smooths across the x_2-axis, thus favoring vertical features. Finally, the $\psi^{\mathrm{d}}(x_1,x_2) = \psi(x_1)\psi(x_2)$ differences across both directions and favors non-vertical/non-horizontal (i.e., diagonal) features.

Figures 2 and 3 display the six scales from a 2D MRA of the sample vorticity field in Figure 1, only a portion is shown centered at $(x_1 = 399, x_2 = 101)$. The location of this structure is indicated by a cross in the vorticity field in Figure 1. Each row displays the wavelet detail fields associated with the three spatial directions: horizontal, vertical and diagonal. It is clear that each of the two-dimensional wavelet filters captures distinct spatial directions at a fixed spatial scale. Given

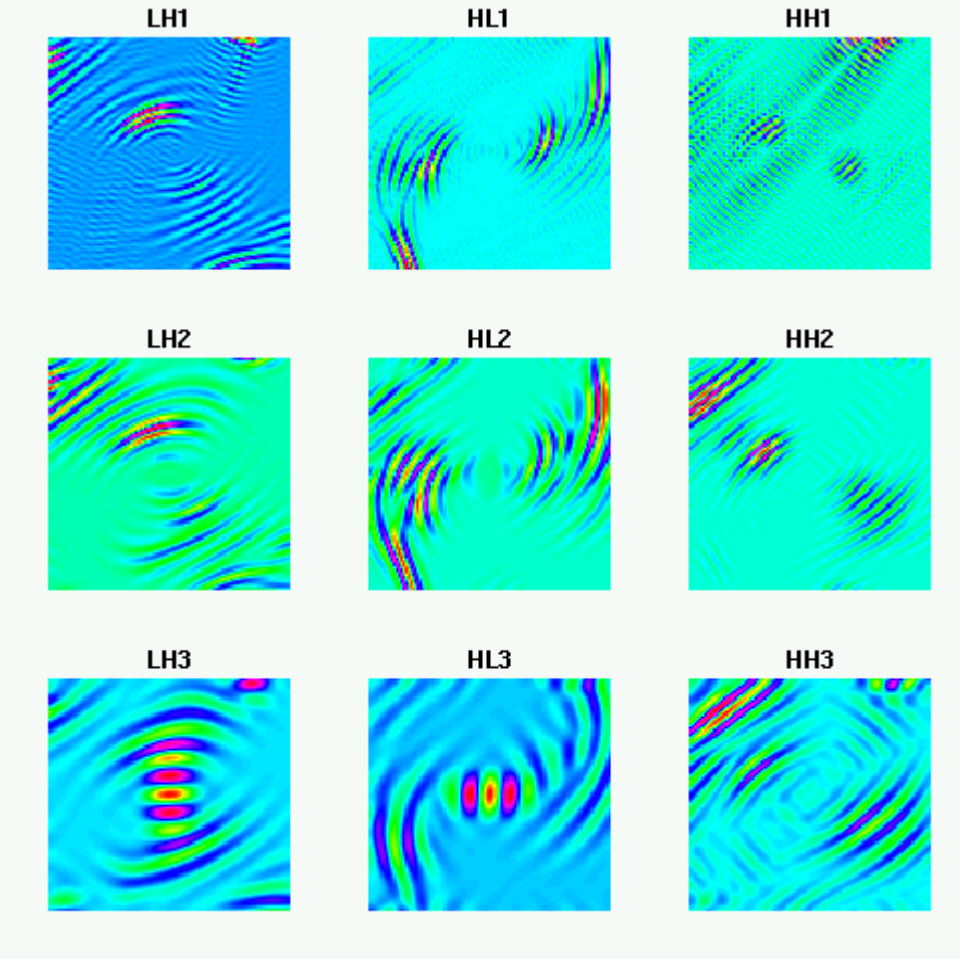

Figure 2: 2D multiresolution analysis ($j = 1, 2, 3$) of vorticity field centered at ($x_1 = 399, x_2 = 101$). The Daubechies least asymmetric wavelet filter ($L = 8$) was used in the computation.

the filaments from this particular vortex are elliptical in shape, it is not surprising to see the detail coefficients of the filament structures strongest in the northeast-southwest directions.

It is interesting to note that the coherent structure is not seen until the third scale (third row in Figure 2) and then only in the horizontal and vertical directions. At higher scales, corresponding to larger spatial areas and lower spatial frequencies, the coherent structure is apparent in all directions. These figures indicate the usefulness of a multiscale representation in separating stylized features of a coherent structure.

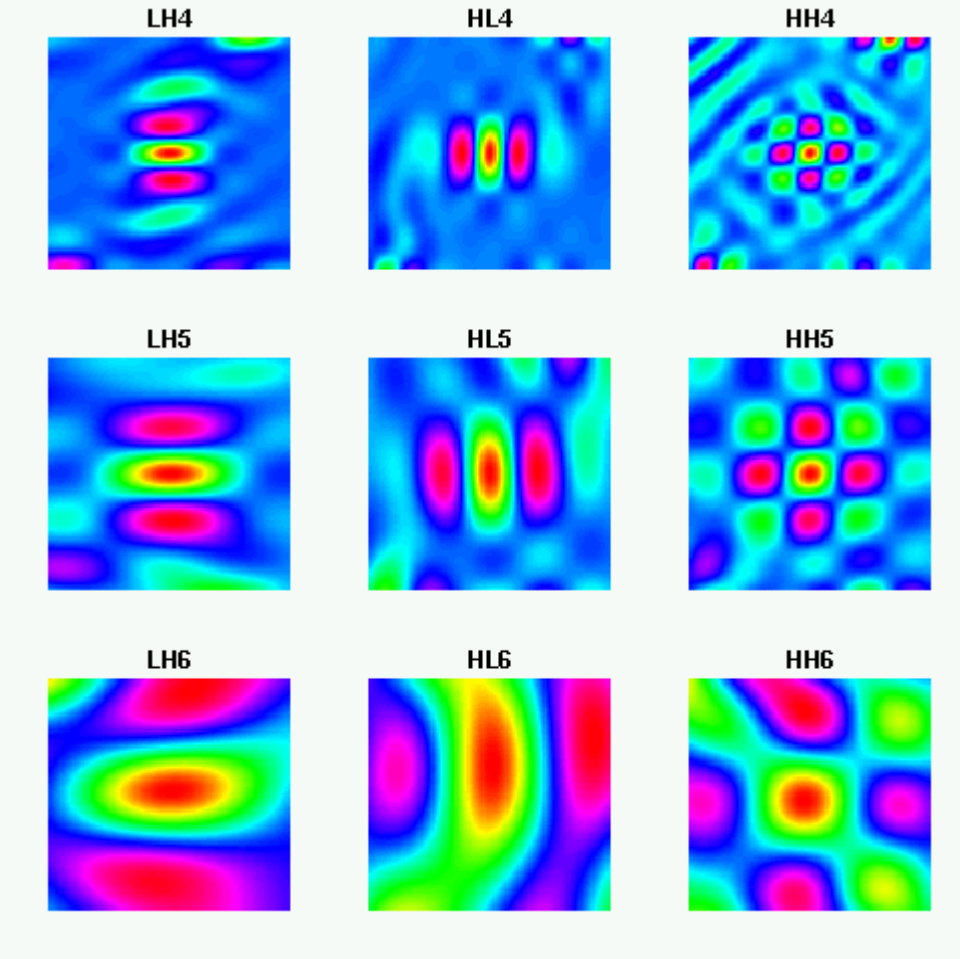

Figure 3: 2D multiresolution analysis ($j = 4, 5, 6$) of vorticity field centered at ($x_1 = 399, x_2 = 101$). The Daubechies least asymmetric wavelet filter ($L = 8$) was used in the computation.

3 Models for Vorticity Fields

Let $\zeta_t(\mathbf{x})$ denote an observed vorticity field at time t. The spatial locations $\{\mathbf{x}\}$ are assumed to fall on a regular grid for both the x_1 and x_2 dimensions, denoted by M rows and N columns, being dyadic in length (M and N are not required to be equal).

3.1 Selecting the Vortex Template

Our goal in this work is to identify and accurately model individual structures in an observed vorticity field. Arguably there must be some user-defined information provided in order to attain this goal, specifically the definition of what a coherent structure should look like from the observed vorticity field. This provides focus for the regression model and yields scientifically interesting

results. By framing our modeling technique in terms of a MRA, the influence of this user-defined information is important but not too restrictive.

One choice of a template function $\tau(\mathbf{x})$ is the outer product of two identical Gaussian curves. This appeared to capture the relevant features of an ideal vortex even if the observed vortices decay back to zero at a faster rate than a Gaussian curve. Although this template is too simple or naive for modeling an observed vortex directly, its MRA is strikingly similar to the MRA of individual structures in the observed vorticity field (Whitcher *et al.* 2002).

There are obvious differences between the template function and coherent structures in the observed vorticity field, some of them are handled automatically by the MRA while others require further discussion. Limitations of the template function that are taken care of by framing our linear regression model in terms of a MRA include: fixed spatial size, amplitude, and radial symmetry. The fixed spatial size is taken care of by levels of the MRA being associated with different spatial ranges. The multivariate multiple linear regression is allowed to pick any spatial size of the template function and associate it with any spatial size in the observed data. The amplitude of $\tau(\mathbf{x})$ is similarly handled through the magnitude of the regression coefficients. The radial symmetry of $\tau(\mathbf{x})$ is another obvious drawback. This is dealt with through the MRA since within each level of the decomposition there are three spatial directions (Section 2.2).

3.2 Single Vortex Model

Our statistical model for a single coherent structure, or *vortex*, in the vorticity field $\zeta(\mathbf{x})$ (we drop the time index to simplify notation) is based on a multivariate multiple linear regression model. The observational model for a single coherent structure is given by

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{4}$$

where $\boldsymbol{\beta}$ is the $(3J+1)\times(3J+1)$ matrix of regression coefficients and $\boldsymbol{\epsilon}$ is the length $3J+1$ vector of errors. The data matrix $\mathbf{Y}$, of dimension $(MN)\times(3J+1)$, is based on a multiresolution analysis (MRA) of $\zeta(\mathbf{x})$; see Section 2.1. The $3J+1$ columns of $\mathbf{Y}$ are generated by column-scanning the 2D wavelet detail fields and concatenating each column. The $(MN)\times(3J+1)$ matrix of covariates (design matrix) $\mathbf{Z}$ is generated by column-scanning and concatenating the user-defined template function $\tau(\mathbf{x})$. The template function is the same dimension as the observed vorticity field.

The implied regression model for each response vector Y_i (column of $\mathbf{Y}$) is

$$Y_i = \beta_{1,i}Z_1 + \beta_{2,i}Z_2 + \beta_{3,i}Z_3 + \cdots + \beta_{3J+1,i}Z_{3J+1} + \epsilon_i,$$

for $i = 1, \ldots, 3J+1$. By transforming the template function via a MRA, a function that is idealized and simplistic in nature provides the foundation of our semi-parametric fitting procedure.

3.3 Vortex Field Model

The previous section proposed modeling an isolated coherent structure in the vorticity field $\zeta(\mathbf{x})$. To more accurately model multiple coherent structures that may be spatially correlated, we propose a vortex field model where the single vortex design matrix is replicated k times (k is assumed to be known)

$$\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\beta}_1 + \mathbf{Z}_2\boldsymbol{\beta}_2 + \cdots + \mathbf{Z}_k\boldsymbol{\beta}_k + \boldsymbol{\epsilon}, \tag{5}$$

where each $\mathbf{Z}_i$ is of dimension $(MN)\times(3J+1)$, each regression coefficient matrix $\boldsymbol{\beta}_i$ is $(3J+1)\times(3J+1)$, and $\boldsymbol{\epsilon}$ is defined as before. The significance of having k distinct design matrices is that each one is centered at the spatial location $(x_{1,k}, x_{2,k})$. Thus, the model in (5) allows for the simultaneous estimation of multiple coherent structures that interact in space.

4 Model Estimation

4.1 Multiresolution-Based Census Algorithm

The following algorithm is our implementation of a vortex extraction procedure for an observed vorticity field of arbitrary finite dimension. The procedure requires at input: a wavelet filter, the maximum depth of the 2D MODWT, and the template function.

1. Compute single vortex models at each spatial location (x_1, x_2) using the efficient implementation in Section 4.2.

2. From the single vortex fits in [1], we compute the distance measure Λ for each spatial location (x_1, x_2) and obtain the local maxima (Section 4.3). The pixels associated with these local maxima constitute our collection of *candidate points* $\mathcal{M}$.

3. Compute the vortex field model, via backfitting, for spatial locations $k \in \mathcal{M}$. We order these candidate points via their distance measure Λ_k from [2], and use forward stepwise selection to include coherent structures into the model (Section 4.4).

4. Statistics may be computed from the individual models $\zeta_t(\mathbf{x})$ of the vortex field model, such as:

- # vortices
- Circulation: $\Gamma_t = \int \zeta_t(\mathbf{x})\, d\mathbf{x}$
- Enstrophy: $\Omega_t = \int \zeta_t^2(\mathbf{x})\, d\mathbf{x}$
- Peak amplitude: $\alpha_t = \max_{\mathbf{x}} \zeta_t(\mathbf{x})$
- Size: the area associated with a 2D Gaussian curve-fitting procedure.

These calculations are not provided in this paper, see Whitcher *et al.* (2002) for more details.

4.2 Fitting the Single-Vortex Model

From the definition of $\mathbf{Z}$ in (4), the template does not shift in the horizontal or vertical and a MRA does not change this fact. Parameters of interest include the spatial position (x_1, x_2) of each coherent structure, along with its matrix $\boldsymbol{\beta}$ of regression coefficients. In order to capture all possible spatial locations and prepare for the eventual fitting of multiple coherent structures, the estimated regression matrix

$$\hat{\boldsymbol{\beta}}^{(x_1,x_2)} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}, \qquad 1 \le x_1 \le M, \;\; 1 \le x_2 \le N, \tag{6}$$

must be performed $M \cdot N$ times; that is, centered at every spatial location. If a sample vorticity field is 512×512, that means (6) must be evaluated approximately 2.6 million times. Even using popular methods for least-squares fitting, such as the QR or singular value decomposition, this is a daunting computational task.

To fit a multivariate multiple linear regression at each pixel in an observed vorticity field, we make use of the fact that regression may be viewed as smoothing, and utilize the Fourier transform. First, let $\mathbf{UDV}' = \mathbf{Z}$ be the singular value decomposition of the $(MN) \times (3J+1)$ covariate matrix in (4) and replace $\mathbf{Z}$ with the orthogonal matrix $\mathbf{U}$ in (6) to obtain the matrix of regression coefficients $\hat{\boldsymbol{\beta}}_{\text{SVD}}^{(x_1,x_2)} = \mathbf{U}'\mathbf{Y}$. The regression matrix of interest in (6) is recovered by undoing the SVD via $\hat{\boldsymbol{\beta}}^{(x_1,x_2)} = \mathbf{V}\mathbf{D}^{-1}\hat{\boldsymbol{\beta}}_{\text{SVD}}^{(x_1,x_2)}$. This eliminates the explicit matrix inversion, but does not replace iterating over MN pixels in the observed vorticity field. If we regard regressing Y_i on U_k at each pixel as just "smoothing," then the discrete Fourier transform $\mathcal{F}$ may be utilized to perform this convolution; i.e.,

$$\widehat{\mathcal{B}}_{k,i} = \mathcal{F}^{-1}\left\{\mathcal{F}[\text{image}(U_k)] \cdot \mathcal{F}^*[\text{image}(Y_i)]\right\}, \qquad 1 \le i, k \le 3J+1, \tag{7}$$

where image$(\cdot)$ reconstructs a matrix from a vector, $\widehat{\mathcal{B}}_{k,i}$ is of dimension $M \times N$ and * denotes the complex conjugate. Since both dimensions were defined to be dyadic, the two-dimensional fast

Fourier transform may be used to greatly increase computational efficiency. Perform (7) $(3J+1)\times(3J+1)$ times to obtain the alternative regression matrix $\hat{\boldsymbol{\beta}}_{\text{SVD}}^{(x_1,x_2)}$ for all spatial positions (x_1, x_2); e.g., letting $J = 6$ implies iterating (7) $19^2 = 361$ times versus 2.6 million by direct computation.

4.3 Candidate Point Selection

For an observed vorticity field with $M = N = 512$ and $J = 6$, the above procedure will produce approximately $95,000,000$ regression coefficients. We would like to compute a regression model where multiple coherent structures are involved. This is not possible when the number of possible coherent structures is $M{\cdot}N$, so we introduce a simple procedure for discarding a large number of spatial locations. From the single vortex fits we obtained in (7), we compute

$$\Lambda_k = \min\left\{\text{tr}[(\hat{\boldsymbol{\beta}}_k - \mathbf{I})^2], \text{tr}[(\hat{\boldsymbol{\beta}}_k + \mathbf{I})^2]\right\} \tag{8}$$

for each spatial location (x_1, x_2). The criterion Λ_k is small when the regression matrix $\hat{\boldsymbol{\beta}}_k$ is similar to the identity matrix, and thus, the single-vortex model is similar to the template function. Two comparisons are made, one to the positive identity matrix and one to the negative identity matrix, so that positive and negative spinning vortices are favored equally. For specific vorticity fields, the set of candidate points was well partitioned into coherent and non-coherent structures using Λ_k.

We denote the eight-member neighborhood around the pixel (x_1, x_2) by $\mathcal{N}$ and look for local maxima in the Λ_k field. The collection of *candidate points* is defined via $\mathcal{M} = \{(a,b) : \|\hat{\boldsymbol{\beta}}^{(a,b)}\|^2 > \|\hat{\boldsymbol{\beta}}^{(x_1,x_2)}\|^2, \text{ for all } (x_1,x_2) \in \mathcal{N}\}$. When applying this criterion on observed vorticity fields, the collection of candidate points $\mathcal{M}$ not only includes almost all coherent structures in the field but also additional points. The point being that the size of $\mathcal{M}$ is much less, by at least two orders of magnitude, than the number of pixels in the original field.

4.4 Fitting the Vortex Field Model

Instead of directly calculating $\hat{\boldsymbol{\beta}}$ via linear algebra for (5), we utilize backfitting to solve the k simultaneous linear equations

$$\begin{aligned}
\mathbf{Y}_1 - (\cdot + \widehat{\mathbf{Y}}_2 + \widehat{\mathbf{Y}}_3 + \cdots + \widehat{\mathbf{Y}}_{k-1} + \widehat{\mathbf{Y}}_k) &= \mathbf{Z}_1\boldsymbol{\beta}_1 \\
\mathbf{Y}_2 - (\widehat{\mathbf{Y}}_1 + \cdot + \widehat{\mathbf{Y}}_3 + \cdots + \widehat{\mathbf{Y}}_{k-1} + \widehat{\mathbf{Y}}_k) &= \mathbf{Z}_2\boldsymbol{\beta}_2 \\
&\vdots \\
\mathbf{Y}_k - (\widehat{\mathbf{Y}}_1 + \widehat{\mathbf{Y}}_2 + \cdots + \widehat{\mathbf{Y}}_{k-1} + \cdot) &= \mathbf{Z}_k\boldsymbol{\beta}_k.
\end{aligned}$$

The kth vortex model is associated with the spatial location $(x_{1,k}, x_{2,k})$ in $\zeta(\mathbf{x})$, where $k \in \mathcal{M}$. Convergence for this algorithm is achieved by looking at the absolute difference between each regression coefficient matrix $\hat{\boldsymbol{\beta}}_k$ from iteration $i-1$ to i. For any k, if $|\sum_{m,n} \hat{\beta}^{(i)}_{m,n,k} - \hat{\beta}^{(i-1)}_{m,n,k}| > \delta$, for some $\delta > 0$, then the backfitting method was applied again. The number of iterations was found to be small across a wide range of k, usually two or four iterations sufficed. We attribute the small number of iterations to the fact that the coherent structures in an observed vorticity field are relatively isolated (Figure 1) and therefore estimation is not seriously affected by fitting them simultaneously.

To determine the number K of coherent structures in the field we utilize a penalized likelihood criterion, specifically the generalized cross validation (GCV) function. We would like to search the entire model space for the best combination of coherent structures (all subset regression), but this would be too computationally intensive for several hundred candidate points. Instead, an ordering is imposed on the candidate points and a model for the vorticity field is found by only searching along the ordering. After single-vortex models are obtained for all candidate points, they are ranked by their similarity to the identity matrix of dimension $(3J+1)\times(3J+1)$ via Λ_k (8). After the candidate points are ordered according to Λ_k, additional coherent structures are added to the vortex-field model one by one when the GCV is reduced. The number of free parameters associated with a single-vortex model is simply $(3J+1)\times(3J+1)$. If the added vortex model does not reduce the GCV it is discarded and the next model is added. This procedure is iterated until no more candidate models exist.

5 Application to Two-dimensional Turbulent Flows

We return to the observed vorticity field in Figure 1 and compute the vortex statistics using our multiresolution census algorithm outlined in Section 4.1. This vorticity field provides a reasonable example of multiple coherent structures embedded within relatively quiescent background structure. A total number of 53 coherent structures were identified. To illustrate the ability of our procedure to find coherent structures of varying sizes and shapes, Figure 4 shows the fitted vortex field model $\widehat{\mathbf{Y}}$ and the residuals for the vorticity field. Although the template function is of fixed spatial size and amplitude, the fitted coherent structures are quite different from the template function and from each other. Comparing the residual field to the original data (Figure 1), it appears that all visually obvious structures have been identified. In addition, one spurious structure was included

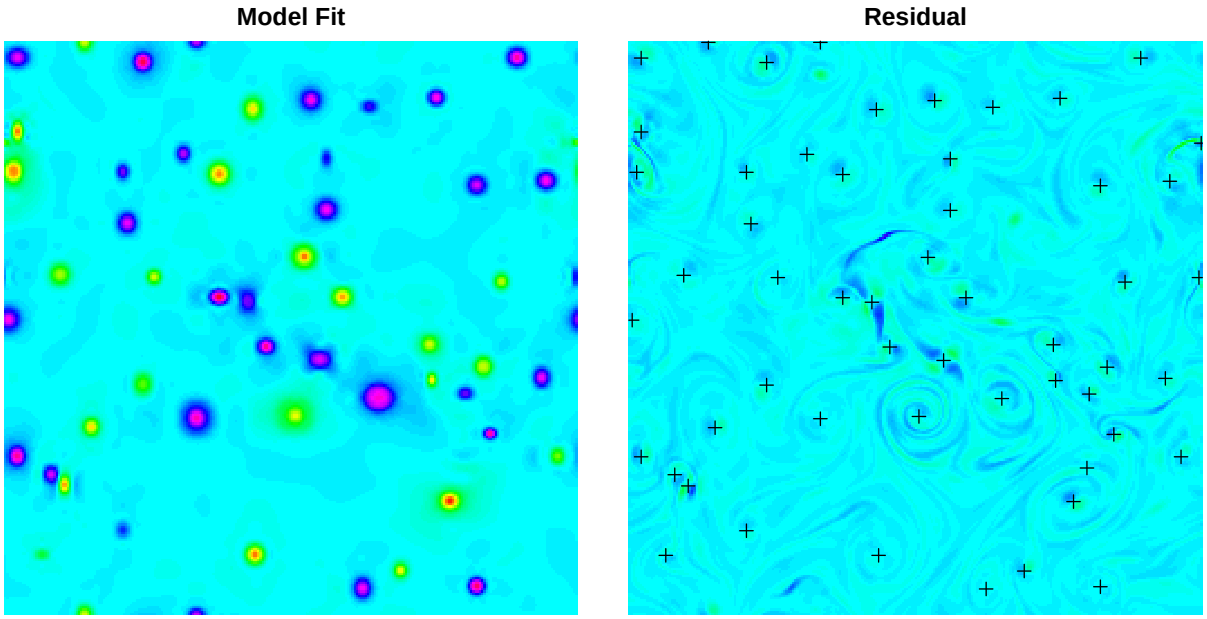

Figure 4: Fitted vortex field model $\widehat{\mathbf{Y}}$ and residual field for the vorticity field at time $t = 15$, plotted on the same color map. The spatial locations of the estimated coherent structures are plotted on the residual field for ease of comparison.

in the model.

6 Discussion

Using a multiresolution regression model, we have been able to accurately model and extract coherent structures from a simulated two-dimensional fluid flow. This technique allows for a fast, objective analysis of observed turbulent fields. Although these results have been achieved by previous algorithms, our method provides flexibility to the researcher through the choice of a template function and modular implementation of the algorithm while stability and precision of the results are guaranteed by using sound statistical techniques. We plan on applying this method to a variety of data sets including more realistic simulations, climate model output, and satellite imagery. The method may also be applied to simulations of three-dimensional quasi-geostrophic turbulence.

Acknowledgments

Support for this research was provided by the National Center for Atmospheric Research (NCAR) Geophysical Statistics Project, sponsored by the National Science Foundation (NSF) under Grants DMS98-15344 and DMS93-12686.

References

Farge, M., E. Goirand, Y. Meyer, F. Pascal, and M. V. Wickerhauser (1992). Improved predictability of two-dimensional turbulent flows using wavelet packet compression. *Fluid Dynamics Research 10*, 229–250.

Farge, M., N. Kevlahan, V. Perrier, and E. Goirand (1996). Wavelets and turbulence. *Proceedings of the IEEE 84*(4), 639–669.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing* (1st ed.). San Diego: Academic Press.

McWilliams, J. C. (1990). The vortices of two-dimensional turbulence. *Journal of Fluid Mechanics 219*, 361–385.

Siegel, A. and J. B. Weiss (1997). A wavelet-packet census algorithm for calculating vortex statistics. *Physics of Fluids 9*(7), 1988–1999.

Whitcher, B., J. B. Weiss, D. W. Nychka, and T. J. Hoar (2002). A multiresolution census algorithm for calculating vortex statistics. Geophysical Statistics Project, National Center for Atmospheric Research.

Wickerhauser, M. V., M. Farge, and E. Goirand (1997). Theoretical dimension and the complexity of simulated turbulence. In W. Dahmen, P. Oswald, and A. J. Kurdila (Eds.), *Multiscale Wavlet Methods for Partial Differential Equations*, Volume 6 of *Wavelet Analysis and Applications*, pp. 473–492. Boston: Academic Press.

Wickerhauser, M. V., M. Farge, E. Goirand, E. Wesfreid, and E. Cubillo (1994). Efficiency comparison of wavelet packet and adapted local cosine bases for compression of a two-dimensional turbulent flow. In C. K. Chui, L. Montefusco, and L. Puccio (Eds.), *Wavelets: Theory, Algorithms, and Applications*, pp. 509–531. San Diego, California: Academic Press.