An Adaptive Multiresolution Data Filter: Applications to Turbulence and Climatic Time Series

J. F. HOWELL AND L. MAHRT

College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon

(Manuscript received 6 July 1993, in final form 3 January 1994)

ABSTRACT

To remove small-scale variance and noise, time series of data are generally filtered using a moving window with a specified distribution of weights. Such filters unfortunately smooth sharp changes associated with larger-scale structures. In this study, an adaptive low-pass filter is developed that not only effectively removes random small-scale variations but also retains sudden changes or sharp edges that are part of the large-scale features. These sudden changes include fronts, abrupt shifts in climate, sharp changes associated with a heterogeneous surface, or any jump in conditions associated with change on a larger scale.

To construct the filters, gradients on different scales and at different positions in the time series are computed using a multiresolution representation of the data. The low-pass filter adapts to include smaller-scale variations at positions in the time series where the small-scale gradient is steep and represents change on a larger scale. The action of the filter is to apply a more concentrated distribution of weights at locations in the original time series where the signal is rapidly varying.

As application examples, the filter is applied to turbulence data observed under strong wind conditions and climate data corresponding to 52 years of a Southern Oscillation index.

1. Introduction

Low-pass filters such as running means computed over a constant window width not only remove noise and small-scale fluctuations but also smooth sharp changes associated with large amplitude coherent structures, an undesirable property for some applications. Examples of such sharp changes include synoptic-scale fronts, wind gusts and edges of thermals, nearly discontinuous surface features, relatively abrupt changes in climate, and sudden changes between modes in a dynamical system. The goal of this study is to develop an adaptive low-pass filter that retains these sharp features, yet removes other small-scale variance and noise.

Adaptive filtering schemes that partially address this issue have been developed and applied successfully (Burt and Adelson 1983; Witkin 1983; Vautard and Ghil 1989; Penland et al. 1991; Mallat and Hwang 1992; Keppenne and Ghil 1992). The low-pass characteristics of the adaptive filters applied in these studies, however, are not formally discussed. In this study an adaptive filter is formulated explicitly as a low-pass filter. The present filter is constructed in terms of a multiresolution decomposition/reconstruction of the data contained in a moving window. Specifying the "window width" (or spatially constant cutoff scale) in conventional running means is augmented in the present filter by specification of the within-window variance to be retained in the low-pass signal. The spatially varying "scale cutoff" is locally determined by the data and the specified variance. The window width serves as an upper bound for the small-scale cutoff and a lower bound for the total record variance to be retained. This data filter will be referred to as a variance-conserving multiresolution (VCM) filter.

From another point of view, the filter acts as a running mean with a weighting distribution that varies in such a manner that small amplitude, small-scale variations are removed while rapid changes associated with large-scale structure are retained. The VCM filter is characterized by a spatially concentrated distribution of weights near a sharp change in the signal and weights that are spatially spread out where the signal is slowly varying. The precise shape of the weighting distribution is determined by the local data.

The analysis in this study could equivalently be posed in terms of the Haar basis (Fig. 1; also Haar 1910), which in turn can be interpreted in terms of wavelets (Daubechies 1988; Chui 1992). Decompositions of geophysical data in terms of the Haar or other local basis sets can be found in Ganage (1990), Mahrt (1991), Gambis (1992), Ganage and Hagelberg (1993); Kumar and Foufoula (1993), Collineau and Brunet (1992), Turner and Leclerc (1994), Turner et al. (1994), and Howell and Mahrt (1994). The anal-

Corresponding author address: Dr. Larry Mahrt, Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331-2209.

© 1994 American Meteorological Society
ysis in this study, however, is formulated so that it is not necessary to define explicitly a basis set in the decomposition/reconstruction algorithm (sections 2 and 3). Instead the filter is constructed in terms of the formalism of multiresolution analysis (Burt 1984; Mallat 1989). As examples, the filter is applied to a turbulence time series and a Southern Oscillation index.

2. The decomposition/reconstruction

The present technique decomposes a window of data in terms of differences between averages computed over window subsegments. A record of data is defined to consist of points sampled at a constant rate between the times $t = 0$ and $t = N\delta$. The time series is denoted as

$$f_i; \quad i = 1, 2, \ldots, N,$$

where the sampling interval $\delta$ is the time between data points and $N$ is the total number of points in the time series. A data point $f_i$ is identified with the time $t_i = (i - \frac{1}{2})\delta$, which is the geometric center of the $i$th sampling interval. A decomposition/reconstruction window is defined as a segment of the time series consisting of $L = 2^m$ consecutive data points, where $M$ is an integer $\geq 1$ to be specified. The data in each such window are decomposed and then reconstructed.

In order that the low-pass information be retained, it is necessary first to determine the window average denoted as

$$\bar{f}_k(a_M; 1) = \frac{1}{L} \sum_{j=1}^{L} f_{k+j},$$

where the general notation $f_k(a_m; n)$ represents the data contained in the time interval $[k\delta + a_m(n-1), k\delta + a_m(n)]$, which is of length $a_m = 2^n\delta$. The values $m = M$ and $n = 1$ correspond to an entire window of data, including $L$ data points starting at the time $k\delta$. The overbar in $\bar{f}_k(a_M; 1)$ is used to indicate the arithmetic average of the data contained in the corresponding time interval. For future use, (2) is alternatively expressed as

$$\bar{f}_k(a_M; 1) = \frac{1}{2} [\bar{f}_k(a_{M-1}; 1) + \bar{f}_k(a_{M-1}; 2)],$$

corresponding to one-half of the sum of the half-window averages.

One-half of the difference in the half-window averages is expressed as

$$\Delta f_k(a_M; 1) = \frac{1}{2} [\bar{f}_k(a_{M-1}; 2) - \bar{f}_k(a_{M-1}; 1)],$$

so the averages of the first and second half of the window can then be written as

$$\bar{f}_k(a_{M-1}; 1) = \bar{f}_k(a_M; 1) - \Delta f_k(a_M; 1)$$

and

$$\bar{f}_k(a_{M-1}; 2) = \bar{f}_k(a_M; 1) + \Delta f_k(a_M; 1),$$

respectively. If $M = 1$, that is, if there are only two points in the reconstruction window, then the quantities in (5) and (6) equal $f_{k-1}$ and $f_{k+2}$, respectively.

Generally, the window is sequentially halved into smaller and smaller segments, thus increasing the resolution of the decomposition. As the segment size decreases, more translated segments are needed to cover the window. If a segment is of length $a_m = 2^n\delta$, then $2^{M-m}$ nonoverlapping segments are required to completely cover a window of length $a_M = 2^M\delta$. The difference operation applied to the first segment of length $a_m = 1$ (first half of the window) is

$$\Delta f_k(a_{M-1}; 1) = \frac{1}{2} [\bar{f}_k(a_{M-2}; 2) - \bar{f}_k(a_{M-2}; 1)],$$

and the operation on the second segment (second half of the window) is

$$\Delta f_k(a_{M-1}; 2) = \frac{1}{2} [\bar{f}_k(a_{M-2}; 4) - \bar{f}_k(a_{M-2}; 3)].$$

These two quantities can be used to reconstruct the quarter-window averages by starting with the half-window averages represented in (5) and (6) and then adding or subtracting either (7) or (8) to account for the differences between adjacent quarter-window averages.

For example, the average value of the data contained in the third segment of length $a_m = 2$ is

$$\bar{f}_k(a_{M-2}; 3) = \bar{f}_k(a_{M-2}; 2) - \Delta f_k(a_{M-1}; 2).$$

In general, the window is subdivided in half a total of $M$ times in order to transform the original data into gradients computed on different scales and at different within-window positions. A point in the window is reconstructed in terms of the transforms at the different scales that include the point. In Fig. 1 this corresponds to drawing a vertical line at a fixed horizontal position, and the transform intervals crossed by that line are used...
to reconstruct the given point. One transform term at each scale is used. A transform term at a given scale is either added or subtracted depending on whether that given point was added or subtracted in the associated difference operation.

Summarizing the decomposition, a datum located at the $j$th within-window position in a window containing $2^M$ data points is reconstructed as

$$f_{k+j} = \bar{f}_k + \sum_{m=1}^{M} (-1)^l \Delta f_k(a_m; n)$$

$$n = 1 + \text{int} \left( \frac{j - 1}{2^m} \right) ; \quad l = 1 + \text{int} \left( \frac{j - 1}{2^{m-1}} \right), \quad (10)$$

where $\text{int}(\cdot)$ is the integer part of $\cdot$, the window mean $\bar{f}_k$ is expressed in (2), and $\Delta f_k(a_m; n)$ represents the local gradient in $f$ on the scale $a_m = 2^m \delta$ across the transform segment covering the given datum. Nonoverlapping window subsegments of length $a_m$ and $a_{m-1}$ are enumerated, leading to the auxiliary translation numbers $n$ and $l$, respectively.

The total sampled variance contained in the window of data is decomposed as

$$\frac{1}{2^M} \sum_{j=1}^{2^M} (f_{k+j} - \bar{f}_k)^2 = \sum_{m=1}^{M} \frac{1}{2^{m-1}} \sum_{n=1}^{2^{m-1}} [\Delta f_k(a_m; n)]^2. \quad (11)$$

Equations (10) and (11) can be directly verified by substituting in the expression for $\bar{f}_k$ from (2) and expressing $\Delta f_k(a_m; n)$ in terms of $f$ as

$$\frac{1}{2^{m-1}} \sum_{j=1}^{2^{m-1}} (f_{k+2^{m-1}(n-1/2)+j} - f_{k+2^{m-1}(n-1)+j}) \quad (12)$$

and carrying out the algebra.

The sum in (10) represents the total decomposition of the within-window variations. Truncating this sum, so that the sum from $m = 1$ to $m = M$ is replaced by the sum from $m = m_0$ to $m = M$ where $m_0 \geq 1$, leads to a small-scale cutoff $a_{m_0}$ that may depend on the position within the window. A technique for selecting the scale cutoff is presented in the next section.

3. Truncating the reconstruction

The previous section formulated the representation of a datum point in terms of a window average and smaller subwindow scale variations according to (10). In this section, a method of truncating this representation is introduced. The method relies on specifying a fraction $\alpha$ of the total sampled within-window variance to be reconstructed. Within a given window, the variance is decomposed (11) in terms of the differences $\Delta f_k(a_m; n)$. The construction of the filtered signal within each window proceeds in incremental steps until the prescribed fraction of total within-window variance is restored. The step-by-step reconstruction algorithm is now described for an individual window starting at the time $k_0$. A graphical presentation of this procedure is presented in section 4a.

The decomposition term with the largest magnitude, say $\Delta f_k(a_{m^*}; n^*)$, is identified in order to initiate the first reconstruction step. If $m^* < M$, then all larger-scale terms that cover the within-window transform interval $[a_{m^*}, n^* - 1], a_{m^*}, n^*]$ are also included in the first step. The low-pass filter requires inclusion of the larger-scale terms.

Of the decomposition terms not used in the first step, the term with the largest magnitude is used to initiate the second step. All larger-scale terms covering the associated transform interval are then included as well, provided they were not used in the first step. Each reconstruction step proceeds in this manner.

At the end of each reconstruction step, the total amount of variance reconstructed is updated. If this quantity is greater than the amount specified, then the window reconstruction is terminated; otherwise the next reconstruction step is taken. For example, the variance reconstructed after the first step is

$$\sum_{m=m_0}^{M} \frac{1}{2^{m-m^*}} [\Delta f_k(a_m; n')]^2; \quad n' = 1 + \text{int} \left( \frac{n^* - 1}{2^{m-m^*}} \right), \quad (13)$$

where $m^*$ and $n^*$ represent the maximum difference $|\Delta f_k(a_m; n^*)|$ among all the differences $|\Delta f_k(a_m; n)|$ for a fixed $k$. If the sum in (13) is greater than the specified fraction of the total within-window variance to be reconstructed, $\alpha \times (11)$, then the window reconstruction is complete. Otherwise, the largest coefficient not previously included is found in order to begin the next reconstruction step. This process is repeated until the specified variance is obtained. Upon satisfying the specified variance, the smallest scale included in the window reconstruction that covers the $j$th within-window position corresponds to the index $m_0(j, k; \alpha)$.

In other words, the filter is formulated by modifying (10), so the summation begins at $m_0(j, k; \alpha)$ where $1 \leq m_0(j, k; \alpha) \leq M$. The truncated reconstruction of a given point in a window is then written as

$$\hat{f}_{k+j} = \bar{f}_k + \sum_{m=m_0(j,k;\alpha)}^{M} (-1)^l \Delta f_k(a_m; n)$$

$$n = 1 + \text{int} \left( \frac{j - 1}{2^m} \right) ; \quad l = 1 + \text{int} \left( \frac{j - 1}{2^{m-1}} \right), \quad (14)$$

which is the truncated form of (10) resulting from the introduction of the cutoff scale index $m_0(j, k; \alpha)$. The relation between the reconstruction steps and the cutoff
scales \(a_{md}(j,k)\) is described in terms of a specific atmospheric example in section 4a. If \(m_0(j,k;\alpha) = 1\), then the original datum is reconstructed exactly (no truncation). If \(m_0(j,k;\alpha) = M\), then the reconstructed value is simply the half-window average. Specifically, the value of \(\hat{f}_{k+j}\) in (14) is equal to an average of \(2m_0(j,k;\alpha) - 1\) consecutive data points.

A low-pass filter (adaptive running mean) for the entire record is implemented by moving the reconstruction window. In the present study, all possible starting positions for the reconstruction windows are used, corresponding to maximum overlap of windows. Translating the window through the record one point at a time leads to decomposition/reconstruction windows of length \(a_M = 2^M\delta\) starting at times \(t = k\delta\) for \(k = 0, \ldots, N - 2^M\). Each point in the time series (except for the endpoints) is reconstructed more than once, leading to a redundant representation. The reconstructions of a given point corresponding to the different windows that cover that point are averaged to yield the final filtered output, written as

\[
\hat{f}_t = \frac{1}{K_t} \sum_{k=-K_t}^{K_t} \hat{f}_{t+k},
\]

where \(j = i - k\) and the sum is over the \(K_t\) windows that cover the datum \(f_i\), specifically,

\[
k_t^- = \max(i - 2^M, 0), \quad k_t^+ = \min(i - 1, N - 2^M),
\]

and

\[
K_t = k_t^+ - k_t^- + 1 = \min(2^M, i, N - i - 1, N - 2^M + 1).
\]

The averaging of window reconstructions reduces possible phase problems, as discussed by Vautard and Ghil (1989). Characteristics of the filter are now illustrated in terms of a specific atmospheric application.

4. Turbulence example

In this first example, a 5-min time series of the fluctuating longitudinal and vertical wind components measured 45 m above flat terrain in near-neutral conditions are analyzed (Kristensen et al. 1989). The wind was sampled at a rate of 16 Hz, so the 5-min segment contains 4800 data points. The mean longitudinal wind speed in the segment is about 16 m s\(^{-1}\), so a conversion from time to pseudo distance (Taylor’s hypothesis) is defined to be 1 km per 1 min. On the horizontal axes of Fig. 2, time increases toward the left in order for the spatial structure to be visualized.

The truncation of the observed signal, summarized in the previous section, depends on the location of the largest values of the transform terms (12) generated from the decomposition. The decomposition terms represent the “average change” over the interval \([b - a/2, b + a/2]\), where \(a = a_m = 2^m\delta\) is the scale (width) of the transform interval, and \(b = b_{m+1} = k\delta + a_m(n - 1)\) is the central position of the interval. Locations of rapid changes in the signal are reconstructed first by starting with the largest decomposition terms as detailed in section 3. This low-pass reconstruction is terminated when a specified fraction of total variance is restored.

a. Step-by-step window reconstruction

This example begins with an individual window reconstruction. Figure 2 shows the final reconstruction of an individual window of data using \(\alpha = 0.8\) in the VCM filtering algorithm, summarized in section 3. The filtered signal (solid line) in Fig. 2 consists of variable-width piecewise constants where the value of a constant is the arithmetic average of the data contained in the corresponding segment (14). Segments are narrower where the wind speed is varying rapidly. For example, the sharpness of the gust front at approximately 2.1 km is resolved.

The reconstruction steps associated with the solid line in Fig. 2 are shown in an \(a-b\) phase plane where the transform segments used in the reconstruction are indicated by the solid horizontal lines (Fig. 3). Step 1 in the window reconstruction starts with the largest transform quantity, which in this case is on a scale \(a = 64\) m and is located at the position \(b = 2.1\) km (Fig. 3a). The reconstruction at this position includes all decomposition terms on larger scales that cover the position \(b = 2.1\) km. These larger-scale terms are required to restore the 32-m averages around the point \(b = 2.1\) km and to form a low-pass signal.

The largest transform term not used in the first step is used to initiate step 2 in the window reconstruction.
This term is on the scale $a = 0.5$ km at the spatial position $b = 1.7$ km (Fig. 3b) and represents eddy substructure. The larger-scale terms that cover the position $b = 1.7$ km were already used in the first step, so no additional terms are required at this stage of the reconstruction. The new term in step 2 adds 10% to the reconstructed within-window variance. Reconstruction steps 3 through 5 account for another 27% of the within-window variance (Fig. 3c). After ten reconstruction steps, 80% of the within-window variance is reconstructed (Fig. 3d), resulting in the final within-window reconstructed signal (solid line, Fig. 2). The window reconstruction has adapted to include relatively large local gradients. The ordering of the steps indicates that the largest gradients in the turbulence signal are at positions $b < 2.5$ km. Overall, the large eddy structure and associated substructure in the turbulent signal are reconstructed in detail, while the smaller more random fluctuations are truncated.

These small-scale fluctuations correspond to 20% of the total within-window variance. Specifying 80% of the within-window variance to be reconstructed is then effective for resolving the large-scale signal. The filter output is less sensitive to the choice of a window width. The specified fraction $\alpha$ of within-window variance to be reconstructed is the key input parameter for the VCM filter. If $\alpha$ is sufficiently close to unity, variations associated with the large-scale coherent structures are essentially recovered by the VCM filter. Relatively small amplitude, small-scale fluctuations are generally removed, while all other sampled variations are retained. A strategy for choosing $\alpha$ and applying the filter iteratively is discussed in the next subsection.

b. Iterative approach

If an insufficient amount of small-scale variance is removed after the first application of the filter, the filter can be applied to the output of the first application using the same $\alpha$. This corresponds to one iteration, as opposed to reducing the value of $\alpha$ and applying the filter to the original data. An iterative application of the filter will skim off additional small-scale variance, which has random phase with respect to the large-scale structure. The ratio of large-scale variance to random phase small-scale variance is thus increased. To remove
still more small-scale variance, the filter is applied to the output of the second application, and so forth, until a satisfactory filtered output is obtained. This "iterative approach" generally yields different results from those obtained by simply decreasing the specified variance $\alpha$ and reapplying the VCM filter to the original data. In particular, if $\alpha$ is selected to be too small then a significant portion of the large-scale coherent signal may be truncated upon the initial application of the filter. The reasoning is as follows.

The average within-window variance reconstructed by the VCM filter will always exceed the specified variance due to the fact that the reconstruction steps add discrete amounts of variance. If this excess variance is relatively large (say $>5\%$ over the specified fraction), then the final reconstruction steps in each window are, on average, adding significant variance. Changes in $\alpha$, in this case, can eliminate or add reconstruction steps that lead to significant changes in the coherent part of the filtered output. On the other hand, when $\alpha$ is sufficiently close to unity, the final reconstruction steps are, on average, adding only small amounts of localized small-scale variance, and the characteristics of the filtered output vary more continuously with small changes in $\alpha$. The iterative approach is a more conservative technique for removing additional small-scale variance since it can be initiated with an $\alpha$ large enough to begin filtering in this more continuous regime. Elimination of significant variance in the coherent signal is then avoided.

In summary, the current approach is to specify $\alpha$ just large enough so that after the first filter application, large-scale structures are resolved in detail. If excessive small-scale variance is still present, the VCM filter is then applied iteratively. In the present application, the iterations were terminated when the primary effect of further iterations was to decrease the amplitude of the large-scale structures.

The reconstruction window size is specified to be 1 km for application to the turbulence data. Because the window averages are reconstructed by the filter, approximately 2-km large eddy structures are automatically restored. By specifying $\alpha = 0.8$, the coherent structures are retained with insignificant smoothing or amplitude reduction (Fig. 4b). After the first iteration, most of the random phase small-scale variance is removed (Fig. 4c). Gusts in the longitudinal wind speed are indicated by the arrows in Fig. 4. The smallest amplitude gust (near 1 km) is eliminated by the first iteration (second application) of the filter (Fig. 4c), though the sharpness of the other gusts is generally retained.

In Fig. 4d, the 250-m running mean is plotted for comparison. While the running mean is everywhere
smooth, the VCM filtered data consist of steep small-scale gradients at the positions of the sharp wind gusts (sudden changes in the larger-scale $u$ motions). The running mean smooths these gusts, leading to small-scale gradients that are spatially more uniform. Despite smoothing most of the small-scale detail, the running mean contains approximately the same total variance as the once-iterated VCM filtered signal in Fig. 4c. The total variances in Fig. 4 indicate that the output of the initial filter application contains $5.6/6.7 = 84\%$ of the total variance in the original signal. In turn, the once-iterated filtered signal (Fig. 4c) contains $4.8/5.6 = 86\%$ of the variance in the output of the initial filter application (Fig. 4b).

To better understand where the total variance in the VCM filtered signals is gained or lost, consider a single application of the filter, with specified $\alpha$, to a time series that could be the output of a previous filter application. The fraction of total variance in the output signal relative to the total variance in the input signal is not necessarily greater than or less than the specified fraction $\alpha$ of the within-window variance to be reconstructed. There are several reasons for the lack of a one-to-one correspondence. The low-pass VCM filter automatically restores energy on scales larger than the window width (variance gain), the filter exceeds the specified variance in each reconstruction window (variance gain), and the windows overlap, which leads to some smoothing (variance loss). An important feature of the VCM filter, however, is that the total variance in the output signal is a nondecreasing function of $\alpha$ and when $\alpha = 1$, the input signal is exactly reproduced.

The filter in the present application is specified in terms of the longitudinal wind $u$. Using the same filtering operations defined by $u$, the filter is also applied to the vertical wind $w$ (Fig. 5). The result suggests how the $w$ motions are organized on scales associated with coherent structures in $u$. Because the $w$ motions tend to vary on smaller scales, the fraction of total variance reconstructed is less than that for the longitudinal wind $u$. Specifying the filter in terms of $w$ leads to signals with more small-scale variance since significant structure in $w$ occurs on smaller ($<100$ m) scales. An alternative approach is to base the filtering on the largest magnitudes of the product $\Delta u(a_m; n) \times \Delta w(a_m; n)$ in order to isolate the motions associated with momentum flux. This approach is taken in Howell and Mahrt (1994).

c. Physical interpretation

The filtered signals isolate several features of the large eddies. The largest organization included in the 5-km record is approximately 2-km ramp structures where the longitudinal wind experiences a sharp increase followed by a slower decay. Examples of the
gust fronts associated with these ramp structures occur at 0.3 km, 2.1 km, and 4.7 km horizontal positions in Fig. 4. The organization on this scale occurs with little vertical motion. These motions may be large boundary-layer eddies (Eting and Brown 1993) whose circulation at the 45-m observation level is forced to be more horizontal by the presence of the ground surface. Such a circulation pattern may contribute to the larger-scale horizontal velocity variance, but less significantly to the flux. Because they do not contribute significantly to the flux, these motions are referred to as inactive eddies in Högström (1990) and Mahrt and Gibson (1994). The vertical flux of momentum (not shown) is associated mainly with the smaller scale substructure of these large eddies, roughly on the 500-m scale (see arrows, Fig. 4). This 500-m substructure is associated with short lulls in the wind speed occurring with rising motion followed by a gust with sinking motion, as can be seen by comparing Figs. 4 and 5.

On still smaller scales, the vertical wind component contains more variance than the longitudinal wind component. The ratio $w/u$ of the structure functions for small separation distances (say < 50 m) is approximately 4/3 (Howell 1993), which is consistent with similarity scaling for high Reynolds number isotropic turbulence (Kolmogorov 1941). Thus, this example indicates that the turbulence signal is divided into at least three physical scales: 1) large eddies, 2) main transporting eddies (large eddy substructure) that dominate the vertical flux of momentum, and 3) small-scale nearly isotropic motions. A filter with one specified cutoff scale could therefore lead to a physically ambiguous partitioning since the range of scales associated with a given mode varies with record position. The low-pass VCM filter partially captures the spatial variability of the scales associated with a given mode. For example, the low-pass filter captures smaller scales at locations where microfronts occur leading to a more complete representation of the large eddy substructure.

d. Distribution of weights

The filter output at a given point is equivalent to a weighted average of the original signal, though such weights are not explicitly applied. The distribution of weights generally changes depending on the position within the time series because the VCM filter is adaptive. If the signal is rapidly varying at a point, then the distribution of weights will be more concentrated about that point and more small-scale variance is retained. Conversely, in regions where the signal is slowly varying, the distribution of weights will be more spread out, leading to a smoother filtered signal. Two such points are marked in the time series of the longitudinal wind component as “rap var” (rapidly varying) and “slow var” (slowly varying). Corresponding distributions of filtering weights are shown in Fig. 6.

The width of significantly nonzero weights will always decrease (increase) as the specified variance increases (decreases) since increasing the specified variance forces the filter to capture smaller scales. The half-window averages are always reconstructed unless $\alpha = 0$. The distribution of weights, as a result, are bounded by a triangle distribution (Bartlett window; Oppenheim and Schafer 1975, section 5.5) that includes $2^M - 1$ data points (solid lines in Fig. 6). This triangle distribution of the filtering weights effectively corresponds to a running mean over the width of the reconstruction window. An exact reconstruction of a given point corresponds to a delta distribution. The final filtering weights always lie inclusively between the triangle and delta distributions. The reason is that if each window reconstruction is exact, the delta distribution results, and reconstructing only the half-window averages in each window results in the triangle distribution.

e. Spatial distribution of cutoff scale

The spatially varying cutoff scale in the VCM filter depends on the largest transform quantities (12). The local maxima of these quantities at each dyadic scale are determined in order to verify the “action” of the
filter as well as to further study the characteristics of the observed data. Gambis (1992) also emphasized the scales and positions of the largest values of wavelet transform coefficients in the $a-b$ phase plane in order to help interpret data.

In Fig. 7, the $(a, b)$ coordinates of the filled circles correspond to local maxima at dyadic scales of the transform quantities (12) for the raw turbulence data plotted in Fig. 4a. Figure 7 shows only local maxima that exceed two standard deviations of all transform quantities on dyadic scales. The radius of a circle in Fig. 7 is linearly proportional to the magnitude of the corresponding transform. The largest transform quantities occur on the scale of a few hundred meters at the location of the relatively large-scale gusts. Such $a-b$ phase plane information provides statistics on the spatial size and spacing of the principal microfronts.

Except for those near the boundaries, each transform quantity at a given dyadic scale enters into an equal number of overlapping window decompositions. The largest decomposition terms are included first in the truncated reconstruction (section 3). Local maxima of the transform quantities with respect to spatial position $b$ on a given scale $a$ will thus be included in the overlapping reconstructions more often compared to spatially nearby terms.

An estimate for the spatially varying cutoff scale is defined here as the logarithmic average of the small-scale cutoffs used in the overlapping reconstructions of the $i$th record position. This quantity is written as

$$a_{\text{shd}(i; \alpha)} = 2^\alpha d_{\text{shd}(i; \alpha)} \delta_i,$$

$$m_0(i; \alpha) = \frac{1}{K_1} \sum_{k=1}^{K_1} m_0(j, k; \alpha),$$

where $j = i - k$, and $k_1$, $k_2$, $k_3$, $K_1$ are defined in (16). The spatially varying cutoff scale associated with one application of the VCM filter to the turbulence signal (Fig. 4b) is indicated by the solid curve in Fig. 7. This curve shows how the VCM filter includes smaller scales in regions of sharp coherent changes associated with wind gusts and illustrates the adaptive nature of the filter.

5. Climatic example

As an alternative example, the VCM filter with overlapping windows is now applied to 52 years (July 1941 through December 1992) of a monthly recorded Southern Oscillation index (SOI). For comparison, the SOI is also filtered according to eigenvectors (EOFs) of overlapping decomposition/reconstruction windows.

Keppenne and Ghil (1992) also applied an EOF-based singular spectrum analysis (SSA) filter to SOI data, though the data in their study are somewhat different from that of the Climate Analysis Center (1993) used in this study. The SOI data are described in the following subsection. The SSA filtering approach is then summarized (section 5b). Both VCM and SSA filters adapt to the data and in the current application lead to filtered signals containing nearly an identical amount of variance (section 5c). These two filtered signals have distinctly different properties, however, as discussed in section 5d.

a. Southern Oscillation index (SOI)

To measure the state of the El Niño–Southern Oscillation (ENSO), a variety of indices have been constructed. These indices include variables such as sea surface temperatures, rainfall amounts, and atmospheric sea level pressures (Wright 1989). This study examines a conventional SOI defined according to differences between the eastern and western tropical Pacific (Tahiti – Darwin) sea level pressure (SLP) anomalies. The standardized SLP anomaly for a given month is obtained by first subtracting out the mean SLP for that month and then dividing by the corresponding standard deviation. This is done for each month at each of the Tahiti (17.3°S, 149.3°W) and Darwin (12.3°S, 130.5°E) stations. Next, the Tahiti – Darwin difference of the resulting SLP anomalies is taken. The differences are then standardized by the mean annual standard deviation. These standardized differences make up the Southern Oscillation index reported by the Climate Analysis Center (1993) and analyzed in this study.

Variability in the SOI is related to the ENSO phenomenon. The ENSO has been associated with planetary-scale oscillations in the ocean–atmosphere system (Bjerknes 1969; Horel and Wallace 1981). These os-
cillations can influence regional weather patterns at remote locations (Kousky et al. 1984). Reviews and models of the ENSO can be found in Schopf and Suarez (1988), Battisti and Hirst (1989), Barnett et al. (1991), Cane (1992), and Chao and Philander (1993). Natural variability in the ENSO is suggested by the studies of Quinn et al. (1987) and Stahle and Cleaveland (1993).

Anomalous sea level pressures at Tahiti and Darwin described by the SOI time series regularly persist up to a year or more. These (spatially fixed) anomalies often appear and disappear suddenly relative to their duration. The raw monthly SOI data contain steep, small-scale (interseasonal) gradients that lead to changes in the average data computed on annual timescales. To isolate interannual variability, conventional filtering techniques tend to preclude steep, interseasonal gradients. Interannual and shorter timescale variability in the SOI time series is described with a 5-month running mean and applying by the VCM and EOF-based filters (section 5c). The results are then characterized in section 5d.

b. Eigenvector decomposition

This eigenvector decomposition relies on the proper orthogonal decomposition theorem (Loève 1963, section 34), which leads to a Karhunen–Loève expansion or expansion into empirical orthogonal functions (EOFs). The EOFs are a statistically optimum global basis set for describing the variance in the data, so the leading-order EOFs capture the dominant variations. For more general discussions and geophysical applications, see North (1984), Mahrt and Frank (1988), Pfeffer et al. (1990), Chang and Mak (1993), and Fraedrich et al. (1993). The present eigenanalysis closely follows the treatment of Keppenne and Ghil (1992), who applied an adaptive data filter to an SOI time series based on an eigenvector decomposition. Their filtering approach relies on a singular spectrum analysis (SSA), yet another variant of the proper orthogonal decomposition theorem. Vautard et al. (1992) provide a current review of SSA and its applications.

For the 52-year SOI record, consider decomposition/reconstruction windows containing L data points. Relative to the starting position of the record (July 1941), windows starting at times k × (δ = 1 month) for k = 0, 1, · · · , N − L are used to generate an L × L lagged covariance matrix C, which is real and symmetric. Here L = 60 is selected so the window width is small enough to achieve reasonable sampling, yet large enough so that variability associated with the ENSO is sufficiently included within a given window (Keppenne and Ghil 1992). An element of the matrix C is defined as

\[ c_{ij} = c_{ji} = \frac{1}{N - L} \sum_{k=0}^{N-L} (f_{k+i} - \bar{f})(f_{k+j} - \bar{f}) \]

\[ i, j = 1, 2, \cdots, L, \quad (18) \]

where \( \bar{f} \) is defined in (2) and N is the total number of points in the record that includes \( N - L + 1 \) distinct windows with maximum overlap. The overlapping windows may be viewed as random phase samples (Vautard et al. 1992; Mahrt and Howell 1994). They are not random samples, however, since the windows overlap.

For a “second-order stationary” process, the lagged covariance for a fixed lag is independent of the within-window position, and the resolved EOFs are Fourier modes (Panofsky and Dutton 1984, section 12.5.1). For a finite number of windows (finite record), lagged covariances can only be estimated, so even if the process being sampled is second-order stationary, the eigenvectors may only resemble Fourier modes. Instead the EOF decomposition is statistically optimum in that the leading eigenvector maximizes the variance explained for that finite record. The second eigenvector maximizes the remaining variance, and so forth. A comparison between spectra based on eigenvector and Fourier basis sets is presented in Mahrt and Howell (1994).

The eigenvectors \( \phi^{(p)} \) of the lagged covariance matrix \( C \) associated with positive eigenvalues \( \lambda^{(p)} \) provide an orthogonal and complete basis set for describing the within-window variations (Loève 1963, section 34). The eigenvectors are defined such that

\[ \sum_{j=1}^{L} c_{ij} \phi^{(p)}_j = \lambda^{(p)} \phi^{(p)}_i, \quad (19) \]

where the eigenvalue \( \lambda^{(p)} \) corresponds to the average within-window variance described by the \( p \)th eigenvector. Values associated with the largest 16 eigenvalues are plotted in Fig. 8. The eigenvalues have been normalized by the average within-window variance. The rapid falloff in the SOI singular spectrum suggests a division between the coherent variations and more

![Fig. 8. Average within-window SOI variance associated with eigenvalues 1–16. The window width is 60 months and the first four eigenvectors describe 51% of the variance.](image-url)
random small-scale variations in the data and is a key to the success of SSA.

A datum located at the jth within-window position is reconstructed as

\[ f_{k+j} = \bar{f}_k + \sum_{p=1}^{L-1} d_k^{(p)} \phi_j^{(p)}, \]  

(20)

where the eigenvectors have been normalized to have a variance of unity. The expansion coefficient \( d_k^{(p)} \) is the inner product of the \( p \)th normalized eigenvector with the window of data starting at time \( k \delta \), which is written as

\[ d_k^{(p)} = \sum_{j=1}^{L} \phi_j^{(p)} f_{k+j}. \]  

(21)

A filter is formulated by truncating (20), so the summation is terminated at eigenvector \( P \) where \( P < L - 1 \). The truncated reconstruction of a given point in a given window is then written as

\[ \hat{f}_{k+j} = \bar{f}_k + \sum_{p=1}^{P} d_k^{(p)} \phi_j^{(p)}. \]  

(22)

According to the SOI singular spectrum (Fig. 8), \( P = 4 \) is selected. The reconstructions of a given point corresponding to the different windows that cover that point are in turn averaged to yield the final filtered output according to (15).

c. Filtering the SOI

The SOI time series includes small-scale (1–3 month) departures that have large amplitudes relative to the large-scale structure. These "spikes" are traditionally reduced with a 5-month running mean. The 5-month running mean will be used as the input for the VCM filter. This prefiltering is unnecessary for the SSA filter since small-scale spikes are primarily described by the higher-order eigenvectors that are removed by the truncated reconstruction (22). This is an advantage of the SSA filter.
The 5-month running mean contains 67.3% of the original total SOI variance (Fig. 9a); however, the running mean still includes small-scale fluctuations that have random phase with respect to the large-scale variations. By applying the VCM filter to the 5-month running mean, these fluctuations are removed effectively without smoothing the large-scale structure.

The VCM algorithm requires a window that includes $L = 2^M$ data points. Accordingly, a 64-month reconstruction window is selected. As stated in the turbulence example (section 4), the VCM filter is less sensitive to the window width and, instead, primarily depends on the specified variance. The specified fraction $\alpha$ of within-window variance to be reconstructed is selected just close enough to unity, here $\alpha = 0.9$, to reconstruct the large-scale coherent variations while truncating small amplitude, small-scale fluctuations. The VCM filter is then applied iteratively, selectively removing a little more localized small-scale variance upon each iteration (section 4b).

The initial VCM filtered output captures the large-scale structure at the cost of retaining some unwanted small-scale variance (figure not shown). Using the same $\alpha$, the filter is then applied to the output of the first filter application. After this second application, the filtered signal contains 51.7% of the total SOI variance (Fig. 9b). Coincidently, this is the same amount of variance associated with the SSA filtered signal (Fig. 9c), corresponding to EOFs 1–4.

To isolate the largest amplitude events, the VCM filter is iterated three more times corresponding to a total of five applications of the filter (Fig. 9d). The final filtered signal contains only 37% of the total SOI variance but still includes steep small-scale gradients associated with coherent change on a longer timescale. The filtered signal in Fig. 9d highly resolves the sudden occurrences of large amplitude, large-scale events, while random phase small-scale variance and relatively small amplitude large-scale variations are removed.

d. Comparison of SOI filtered signals

To further characterize the different filtering techniques, the filtered SOI signals are overlaid for a decade (1977–87) of the record containing the 1983 anomaly (Fig. 10). The 5-month running mean (dotted line, Fig. 10) contains considerably more small-scale variance than the other two signals. The VCM filtered signal (solid line, Fig. 10) closely follows the sharp changes associated with the appearance and subsequent termination of the 1983 anomaly. The SSA filtered signal (dashed line, Fig. 10) smooths these sharp changes.

Consequently, the structures contained in the SSA and VCM filtered signals are quite different even though they both contain the same amount of total variance. The SSA filtered SOI includes small amplitude large-scale oscillations. For example, a small amplitude cycle occurs in the SSA filtered signal following the 1983 event. This cycle is nearly 180° out of phase with the 5-month running mean (see arrows, Fig. 10). The 1983 anomaly appears localized in time, yet includes a broad range of timescales associated with the onset, duration, and decay of the event. Representation of this localized event with EOFs 1–4, which roughly correspond to timescales of 2–5 yr, leads to oscillations in the neighboring years with approximately 2-yr periods. These neighboring oscillations can be interpreted in terms of the Gibbs phenomenon (Oppenheim and Schafer 1975, section 5.5).

Because of the Gibbs phenomenon, truncating an eigenvector or Fourier representation of the data variability can lead to artificial oscillations. The Gibbs phenomenon is reduced by the VCM filter because the data are represented in terms of a local (wavelet) basis. On the other hand, does the VCM filter introduce artificially steep gradients? From numerous test applications to both real and artificial time series, the filter appears only to reconstruct steep gradients when the input data contain such steep gradients. This follows from the fact that the filtered output is a weighted average of the input data and the nonnegative weights are bounded by the triangle distribution (section 4d). However, a more evident difficulty is that the filter tends to include large amplitude small-scale variance with random phase, such as spikes. For this reason the SOI 5-month running mean was used for the input in the current application.

In this current example, the SSA filtered SOI corresponds to a wavy signal with modulated periods and amplitudes since the leading-order eigenvectors resemble sinoids. Such a signal is more predictable using an autoregressive model, as reported by Keppenne and Ghil (1992), which is one of the underlying motives for using an EOF approach (Barnett and Hasselmann 1979). In contrast, the VCM filter removes events with small amplitude relative to other nearby events. For example, the VCM filtered signals are “flat” during
the years following the 1983 anomaly. Between 1984 and 1986, interannual variations in the monthly SOI data appear minimal (Figs. 9 and 10, dotted lines). The VCM filter selects the largest amplitude events and reconstructs them in sharp detail. The SSA-based filter tends to include more events in smoothed form and is not as committed to the largest amplitude events. The two filters contain complementary information.

6. Conclusions

The variance-conserving multiresolution (VCM) filter introduced in this paper adapts to the data by decreasing the low-pass cutoff scale in regions of rapid change. The low-pass VCM filter is effective at eliminating noise while retaining sharp variations in the larger-scale signal. This filter can be a useful diagnostic tool when the coherency is buried in noise. The filter requires specification of the decomposition/reconstruction window width and the fraction of within-window variance to be reconstructed. The specified variance determines the amount of detail retained in the filtered signal. By only removing small amplitude, small-scale variations, the filter effectively retains any abrupt changes in the large-scale coherent signal.

When applied to turbulence data, the low-pass VCM filter preserves the sharpness of the wind gusts, yet removes finescale turbulent fluctuations that have random phase with respect to the large eddies. Conventional low-pass filters would smooth the sudden increase in wind speed at the gust microfronts. When applied to the SOI data, the VCM cutoff scale adapts to the variable range of timescales associated with El Niño/La Niña—Southern Oscillation events as represented by an SOI. The variable timescales may include the onset (monthly timescales), duration (yearly timescales), and subsequent disappearance (monthly timescales) of sea level pressure anomalies. The VCM filter extracts information that is complementary to the SSA adaptive data filter reviewed by Vautard et al. (1992). For the SOI example, the VCM filter better represents a sudden onset of an anomaly, whereas the SSA filter better represents the oscillatory nature of the anomalies. Application of both filters is recommended.

Acknowledgments. We appreciate the insightful comments of S. K. Esbensen, M. Ghil, and the reviewers. We gratefully acknowledge Michael Courtney and Søren Larsen for providing the turbulence data and George Taylor for providing the SOI data. This material is based upon work supported by Grant DAA H04-93-G-0019 from the Army Research Office and Grant ATM-8912736 from the Physical Meteorology Program of the National Science Foundation.

REFERENCES


