P1.23 THE RESPONSE FILTER

Mark Askelson*
University of North Dakota, Grand Forks, North Dakota

Haidao Lin
Florida State University, Tallahassee, Florida

Matt Solum
National Weather Service Forecast Office, Billings, Montana

Charles Chambers
University of North Dakota, Grand Forks, North Dakota

1. INTRODUCTION

Accurate analyses and short-term forecasts of weather conditions can greatly assist the Army in its battlespace operations. Such information can be used to predict impact areas of a chemical or biological attack, to diagnose atmospheric effects on electromagnetic surveillance, and in general to manage personnel and materiel in a battlefield.

Because of the complexity of atmospheric processes and because of a lack of observations, producing highly accurate analyses and short term forecasts of atmospheric conditions is very challenging. The latter problem of a lack of observations is a particularly difficult one, as regions in which battlefield operations are ongoing are less likely to have routine weather observations available.

To address this problem and, thus, to enhance the Army's ability to diagnose and predict atmospheric conditions, research has been directed towards developing an improved analysis technique. This technique, termed the response filter, has shown significant improvements relative to simpler techniques that are oftentimes applied to this problem.

2. BACKGROUND

In atmospheric sciences, a multitude of techniques, including function fitting (linear interpolation, splines, etc.), successive correction methods [e.g., Cressman (1959); Barnes (1964)], statistical objective analysis (also called optimal interpolation), and variational analysis, are used. As discussed by Askelson et al. (2005), function fitting, successive correction methods, and statistical objective analysis schemes can be recast in terms of one-pass distance dependent weighted averaging (OPDDWA). It can also be argued that if a response function can be defined for variational schemes, then they too can be recast in terms of OPDDWA (Doswell and Caracena 1988). OPDDWA, therefore, is a fundamental objective analysis process.

* *Corresponding author address*: Mark A. Askelson, University of North Dakota, Dept. of Atmospheric Sciences, P.O. Box 9006, Grand Forks, ND 58202-9006; e-mail: askelson@aero.und.edu.

For one-dimensional data, OPDDWA can be expressed as

$$f_A(x) = \sum_{i=1}^{N_o} f(x_{oi}) w_N(x_{oi} - x) , \qquad (1)$$

where $f_A(x)$ denotes the analysis field, N_o is the number of observations used in the analysis, $f(x_{oi})$ is the observation value at location i, and $w_N(x_{oi} - x)$ is the weight applied to observation $f(x_{oi})$, with the subscript N indicating that weights are typically normalized such that $\sum_{i=1}^{N_o} w_N(x_{oi} - x) = 1$ so that the average (zeroth harmonic) of the input field remains unchanged.

The response function is commonly used to interpret the effects an analysis scheme has on the input. As shown by Askelson and Straka (2005), for real data the response function can be interpreted in terms of changes to the amplitudes and phases of the sinusoids that comprise the input field. Because the response function provides this information at multiple frequencies, it can enhance understanding of the properties of an analysis scheme relative to single-valued measures like root mean square differences and mean absolute differences.

To illustrate the use of the response function in the interpretation of an analysis scheme and to set the stage for later illustration, consider Fig. 1. As is apparent in Fig. 1a, the quality of the Barnes analysis (dashed line), which is applied using (1) with

$$w_N(x_{oi} - x, \kappa_d) = \frac{\exp\left[-(x_{oi} - x)^2 / \kappa_d\right]}{\sum_{i=1}^{N_o} \exp\left[-(x_{oi} - x)^2 / \kappa_d\right]} , \qquad (2)$$

where κ_d is the smoothing parameter, is affected by the distribution of the observations. This is made clear when one compares the Barnes analysis [dashed line; $f_{Barnes}(x)$] to the analysis one would obtain using the Barnes scheme if the domain of the observations were infinite and continuous [solid line; $f_{theor}(x)$]. The situation where the domain of the observations is infinite and

continuous is optimal in the sense that all possible information regarding the observation field is available to the analysis scheme. This situation is never realized, of course, and thus analysis schemes must generally work with a limited amount of information regarding the input field.

The reasons why $f_{Barnes}(x)$ is not equal to $f_{theor}(x)$ are illustrated in Figs. 1b,c. Figure 1b illustrates the amplitude modulations for $f_{theor}(x)$ (dotted line), $f_{Barnes}(x)$ (dashed line), and for the response filter analysis (solid line), which will be discussed in section 4. As shown by Askelson and Straka (2005), the amplitude modulation is the ratio of a wave's output amplitude to its input amplitude. If the domain of the input field is infinite and continuous, the amplitude modulation of the Barnes scheme is constant, as shown by the dotted line in Fig. 1b. When the domain of the input field is not infinite and continuous, however, the amplitude modulation of the Barnes scheme depends on analysis location, as shown by the dotted line in Fig. 1b. Because the amplitude modulation of the Barnes scheme varies with location when the domain of the input field is discrete, the Barnes scheme cannot replicate $f_{theor}(x)$ in this situation. Moreover, this variation in amplitude modulation results in a distorted post-analysis representation of the input wave, which can readily be seen by comparing the Barnes analysis on the left side of the domain to the Barnes analysis on the right side of the domain and noting that the extremes of the analysis field are different on the two halves of the analysis domain.

Distortion of the analysis field also results from phase modulations, which are illustrated in Fig. 1c. Phase modulation is the amount by which the phase of an input wave is altered by an analysis scheme. For an input field having an infinite and continuous domain, the Barnes scheme produces no phase modulation. When the domain of the input field is discrete, however, the Barnes scheme produces nonzero phase modulations, as shown by the dashed line in Fig. 1c. These nonzero phase shifts are apparent in Fig. 1a in the non-alignment of the extremes of the Barnes analysis field with the extremes of the input field. Because phase modulation is not generally constant across an analysis domain, input waves are distorted.

Consequently, observation distribution can significantly impact the quality of the output of an analysis scheme. It does so by producing amplitude and phase modulations that vary with analysis location.

3. FORMULATION OF THE RESPONSE FILTER

The purpose of the response filter is to explicitly incorporate information regarding the distribution of observations so as to produce analyses as close to those desired [e.g., $f_{theor}(x)$ in Fig. 1] as possible. In doing so, variations in amplitude and phase modulations across the analysis domain (e.g., dashed lines in Figs. 1b,c) are mitigated.

The response filter is based upon the response function. As shown by Askelson et al. (2005), the response function for an OPDDWA analysis of one-dimensional, discrete data is given by

$$W_N^*(v,x) = \sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)] - j\left\{-\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]\right\}, \quad (3)$$

where $W_N^*(v,x)$ is the complex conjugate of the Fourier transform of the normalized effective weight function and v is frequency (= 1/wavelength). As shown by Askelson and Straka (2005) and Askelson et al. (2005), the corresponding amplitude and phase modulations are given by

$$\left|W_N^*(v,x)\right| = \sqrt{\left\{\sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)]\right\}^2 + \left\{\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]\right\}^2} \quad (4)$$

and

$$\varphi_{W_N^*(v,x)} = \tan^{-1}\left\{\frac{\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]}{\sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)]}\right\} + 2\pi n, \quad (5)$$

where n is an integer. The $2\pi n$ term is required in (5) because $\tan^{-1}(\ )$ is 2π periodic and thus, because n is not generally known, $\varphi_{W_N^*(v,x)}$ is $2\pi n$ ambiguous.*

The approach with the response filter is to dictate desired amplitude and phase modulations as a function of frequency and to minimize the differences between these amplitude and phase modulations and those produced by an analysis scheme. This is accomplished using the response filter functional J_{RF} given by

$$J_{RF}\left[w_{N_1},\ldots,w_{N_{N_o}}\right] = \sum_{k=1}^{N_k} I_{M_k}\left[\left|W_N^*(v_k,x)\right| - M_k\right]^2 + \gamma\sum_{k=1}^{N_k} I_{P_k}\left[\varphi_{W_N^*(v_k,x)} - P_k\right]^2, \quad (6)$$

where $w_{N_i} = w_N(x_{oi}-x)$, N_k is the number of frequencies, I_{M_k} is a user-controlled parameter that determines the relative importance placed on matching M_k, M_k is the desired amplitude modulation at frequency v_k, γ is a user-controlled parameter that determines how important the second term in J_{RF} is relative to the first term, I_{P_k} is a user-controlled parameter that determines the relative importance placed on matching P_k, and P_k is the desired phase modulation at frequency v_k. The first

* For the interested reader, techniques do exist for resolving this phase ambiguity. See, for instance, Steiglitz (1982).

term on the rhs of (6) is the summation, over frequency, of I_{M_k}-weighted squared differences between desired amplitude modulations M_k and actual amplitude modulations $\left|W_N^*(v_k,x)\right|$, which, from (4), depend upon both the weights $\left[w_{N_1},\ldots,w_{N_{N_o}}\right]$ and the observation distribution relative to the analysis location x. The second term on the rhs of (6) is the summation, over frequency, of I_{P_k}-weighted squared differences between desired phase modulations P_k and actual phase modulations $\varphi_{W_N^*(v_k,x)}$, which, from (5), again depend upon both the weights $\left[w_{N_1},\ldots,w_{N_{N_o}}\right]$ and the observation distribution relative to the analysis location x.

Under appropriate conditions, simplifications can be applied to $\left|W_N^*(v_k,x)\right|$ and $\varphi_{W_N^*(v_k,x)}$ in (6). These conditions tend be satisfied when one is producing an analysis and not trying to compute something like a derivative. In that case, one would not want to move waves and thus would set $P_k=0$. When phase shifts are nearly zero, the first term under the radical in (4) tends to dominate the second (especially for waves that are retained at a non-negligible amplitude) and

$$\left|W_N^*(v,x)\right|_{\text{sp}} \approx \sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)], \qquad (7)$$

where the subscript sp indicates small phase shift. Moreover, since $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)]$ is typically much larger than $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]$ when phase shifts are small, the phase modulation $\varphi_{W_N^*(v,x)}$ in (5) is directly related to the magnitude of $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]$.* With these simplifications and a multiplication by ½ to simplify subsequent analysis, one can form, from (6), what is termed the linear response filter functional J_{RF_lin} given by

$$\begin{aligned} &J_{RF\_lin}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]= \\ &\frac{1}{2}\sum_{k=1}^{N_k} I_{M_k}\left\{\sum_{i=1}^{N_o} w_{N_i}\cos[2\pi v_k(x_{oi}-x)]-M_k\right\}^2 \\ &+\frac{1}{2}\gamma\sum_{k=1}^{N_k} I_{P_k}\left\{\sum_{i=1}^{N_o} w_{N_i}\sin[2\pi v_k(x_{oi}-x)]\right\}^2. \end{aligned} \qquad (8)$$

This form of the response filter functional is very useful in that it defines a linear least squares minimization problem for $w_N(x_{oi}-x)$. For this reason, the adjective 'linear' is ascribed to this functional.

The linearized response filter (hereinafter lrf) is obtained from (8) by obtaining $\partial J_{RF\_lin}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]/\partial w_N(x_{oj}-x)$ and setting the result to zero. Denoting $x_{oi}-x$ as x'_{oi}, this operation produces

$$\begin{aligned} &\frac{\partial J_{RF\_lin}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]}{\partial w_{N_j}} \\ &=\sum_{k=1}^{N_k}\left\{I_{M_k}\left[\sum_{i=1}^{N_o} w_{N_i}\cos(2\pi v_k x'_{oi})-M_k\right]\cos(2\pi v_k x'_{oj})\right\} \\ &\quad+\gamma\sum_{k=1}^{N_k}\left[I_{P_k}\sum_{i=1}^{N_o} w_{N_i}\sin(2\pi v_k x'_{oi})\sin(2\pi v_k x'_{oj})\right] \\ &=0. \end{aligned} \qquad (9)$$

Algebraic manipulation of (9) results in

$$\begin{aligned} &\sum_{k=1}^{N_k}\sum_{i=1}^{N_o} I_{M_k} w_{N_i}\cos(2\pi v_k x'_{oi})\cos(2\pi v_k x'_{oj}) \\ &+\gamma\sum_{k=1}^{N_k}\sum_{i=1}^{N_o} I_{P_k} w_{N_i}\sin(2\pi v_k x'_{oi})\sin(2\pi v_k x'_{oj}) \\ &=\sum_{k=1}^{N_k} I_{M_k} M_k\cos(2\pi v_k x'_{oj}), \end{aligned} \qquad (10)$$

which can be further manipulated into

$$\begin{aligned} &\sum_{i=1}^{N_o} w_{N_i}\sum_{k=1}^{N_k}\left[\begin{aligned}&I_{M_k}\cos(2\pi v_k x'_{oi})\cos(2\pi v_k x'_{oj}) \\ &+\gamma I_{P_k}\sin(2\pi v_k x'_{oi})\sin(2\pi v_k x'_{oj})\end{aligned}\right] \\ &=\sum_{k=1}^{N_k} I_{M_k} M_k\cos(2\pi v_k x'_{oj}). \end{aligned} \qquad (11)$$

This is the general form of the lrf.

If in (11) $I_{M_k}=I_{P_k}$ and $\gamma=1$, the lrf simplifies to

* A power series expansion for $\tan^{-1}x$ is $\tan^{-1}x = x-\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^7}{7}+\cdots$ (Salas et al. 1986, p. 703). If the magnitude of x is small, then the cubic and higher order terms are small relative to the first term in the expansion and thus the magnitude of the phase modulation is directly related to $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]$. If, in addition $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\cos[2\pi v(x_{oi}-x)]$ is positive, which is typical for low-pass filters (especially for lower frequencies), then $\sum_{i=1}^{N_o} w_N(x_{oi}-x)\sin[2\pi v(x_{oi}-x)]$ also indicates the correct direction (sign) of the phase shift.

$$\sum_{i=1}^{N_o} w_{N_i} \sum_{k=1}^{N_k} I_{M_k} \cos\left[2\pi\nu_k\left(x'_{oi} - x'_{oj}\right)\right] = \sum_{k=1}^{N_k} I_{M_k} M_k \cos\left(2\pi\nu_k x'_{oj}\right), \quad (12)$$

where the trigonometric identity $\cos(a-b) = \cos a \cos b + \sin a \sin b$ has been applied.

Equations (11) and (12) contain the free parameter *j*. Since there are N_o weights (*j*s), (11) and (12) encapsulate N_o equations for each analysis location *x*. These two sets of equations can be expressed in the matrix form $\underline{\underline{\mathbf{A}}}\underline{\mathbf{x}} = \underline{\mathbf{B}}$ as

$$\begin{bmatrix} A_{11} & \cdots & A_{1N_o} \\ A_{21} & \cdots & A_{2N_o} \\ \vdots & \vdots & \vdots \\ A_{N_o1} & \cdots & A_{N_oN_o} \end{bmatrix} \begin{bmatrix} w_{N_1} \\ w_{N_2} \\ \vdots \\ w_{N_{N_o}} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{N_k} I_{M_k} M_k \cos(2\pi\nu_k x'_{o1}) \\ \sum_{k=1}^{N_k} I_{M_k} M_k \cos(2\pi\nu_k x'_{o2}) \\ \vdots \\ \sum_{k=1}^{N_k} I_{M_k} M_k \cos\left(2\pi\nu_k x'_{oN_o}\right) \end{bmatrix}, \quad (13)$$

where

$$A_{ji} = \sum_{k=1}^{N_k} \left[I_{M_k} \cos(2\pi\nu_k x'_{oi}) \cos\left(2\pi\nu_k x'_{oj}\right) + \gamma I_{P_k} \sin(2\pi\nu_k x'_{oi}) \sin\left(2\pi\nu_k x'_{oj}\right) \right] \quad (14)$$

for (11) and

$$A_{ji} = \sum_{k=1}^{N_k} I_{M_k} \cos\left[2\pi\nu_k\left(x'_{oi} - x'_{oj}\right)\right] \quad (15)$$

for (12), with *j* giving the row index and *i* giving the column index.

It is noted that both the number of rows and the number of columns in $\underline{\underline{\mathbf{A}}}$ are determined by the number of weights N_o. Thus, $\underline{\underline{\mathbf{A}}}$ is always a square matrix. Furthermore, $\underline{\underline{\mathbf{A}}}$ is symmetric regardless of whether (14) or (15) is used because swapping the *j* and *i* indices does not change either (14) or (15). Finally, if two (or more observations) are collocated, then $\underline{\underline{\mathbf{A}}}$ is singular and weights cannot be determined using (13). In this situation $\underline{\underline{\mathbf{A}}}$ is singular for either (14) or (15) because two or more rows of $\underline{\underline{\mathbf{A}}}$ are equivalent. This equivalence arises because these rows have equivalent x'_{oj} values.

When (15) is used, $\underline{\underline{\mathbf{A}}}$ has a couple other useful properties. First, in this situation the diagonal elements of $\underline{\underline{\mathbf{A}}}$ are equal to $\sum_{k=1}^{N_k} I_{M_k}$. Second, if (15) is used and if the set of observations used to obtain analysis values remains the same, the frequencies used in $J_{RF\_lin}(x)$ remain the same, and the I_{M_k} remain the same, then $\underline{\underline{\mathbf{A}}}$ only needs to be computed once to produce an analysis.

The symmetry of $\underline{\underline{\mathbf{A}}}$ and the form of (13) means that (13) results from the attempt to find an extremum, with respect to the weights w_{N_i}, of

$$\frac{1}{2}\underline{\mathbf{w_N}}^T \underline{\underline{\mathbf{A}}}\,\underline{\mathbf{w_N}} - \underline{\mathbf{B}}^T \underline{\mathbf{w_N}} = 0, \quad (16)$$

where $\underline{\mathbf{w_N}}^T = \begin{bmatrix} w_{N_1} & w_{N_2} & \cdots & w_{N_{N_o}} \end{bmatrix}$, superscript *T* indicates matrix transpose, and $\underline{\mathbf{B}}^T = \left[\sum_{k=1}^{N_k} I_{M_k} M_k \cos(2\pi\nu_k x'_{o1}) \quad \cdots \quad \sum_{k=1}^{N_k} I_{M_k} M_k \cos\left(2\pi\nu_k x'_{oN_o}\right) \right]$. Equation (16) is simply an alternate expression of the linear response filter functional J_{RF_lin} (8). If $\frac{1}{2}\underline{\mathbf{w_N}}^T \underline{\underline{\mathbf{A}}}\,\underline{\mathbf{w_N}} \geq 0$ for all possible $\underline{\mathbf{w_N}}$, then $\underline{\underline{\mathbf{A}}}$ would be a positive semidefinite matrix. At this point, however, it has not been shown that $\frac{1}{2}\underline{\mathbf{w_N}}^T \underline{\underline{\mathbf{A}}}\,\underline{\mathbf{w_N}} \geq 0$ for all possible $\underline{\mathbf{w_N}}$. For positive I_{M_k}, I_{P_k}, and γ, the lhs of (16) is greater than or equal to zero for all possible $\underline{\mathbf{w_N}}$, however, because it is equivalent to J_{RF_lin}.

With the lrf, equations (13) and (14) or (13) and (15) are solved at each location where an analysis is desired, with the resulting w_N subsequently used in (1) to obtain the analysis value.

Two and three-dimensional versions of the lrf have also been developed (Lin 2004). The mathematical development of these versions of the lrf is essentially the same as the one-dimensional version, with the primary differences being housekeeping associated with the extra dimensions. In the interest of brevity, the details are omitted.

4. RESULTS

An example of the capabilities of the lrf is shown in Fig. 1. In this example, the lrf was constrained to produce the amplitude and phase modulations that the Barnes scheme would have produced if the observation domain were infinite and continuous. In Fig. 1a, it is apparent that the analysis produced by the lrf (plus symbols) is much closer to the analysis that would result if the observation domain were infinite and continuous (solid line). The reasons for this are shown in Figs. 1b,c. As is apparent in these figures, the lrf produces amplitude and phase modulations (solid lines) that are much closer to those that would result if the observation domain were infinite and continuous. In doing so, it greatly reduces the spatial variations in amplitude and phase modulations that are present in the Barnes analysis. Consequently, the lrf produces a more coherent analysis. The reason the lrf is superior to the

Barnes analysis is the lrf, in minimizing $J_{RF\_lin}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]$, finds the weights that produce, as closely as possible, the desired amplitude and phase modulations. In doing so, the lrf explicitly takes into account the distribution of the observations whereas the Barnes scheme does not. Thus, the lrf is an *a posteriori* scheme whereas the Barnes scheme is an *a priori* scheme.

Results obtained with the two-dimensional lrf are illustrated in Figs. 2 and 3. The lrf (Fig. 2b) produces a more accurate representation of the theoretical result (Fig. 2c), which is the analysis that would result if the input field were known everywhere, than does the one-pass Barnes analysis (Fig. 2a). The reasons for this superior performance of the lrf are again the amplitude and phase modulations produced by the lrf. As shown in Fig. 3, the lrf produces, across the domain, amplitude modulations (Fig. 3b) that are much closer than those of the one-pass Barnes scheme (Fig. 3a) to the 0.24 theoretical amplitude modulation. In addition, the phase modulations of the lrf (Fig. 3d) are generally closer to zero than are those of the one-pass Barnes scheme (Fig. 3c). This superior amplitude and phase modulation performance results in a more coherent analysis. Interestingly, the lrf, while producing smaller variations in amplitude and phase modulations, also produces finer-scale variations in amplitude and phase modulations. These result in the somewhat rippled appearance of Fig. 2b. In contrast, the larger-scale variations of its amplitude and phase modulations results in the one-pass Barnes scheme producing an analysis that has larger-scale departures from the theoretical result (Fig. 2a).

Tests of the three-dimensional lrf have also been performed. Because of the difficulties associated with illustrating those tests in a manner similar to Figs. 1-3, a single-valued measure of performance is used instead. This single-valued measure is the root mean square (rms) error, computed here as the difference between values obtained using an analysis scheme and those that would be obtained if the observational domain were infinite and continuous. It is noted that this measure is labeled with the adjective 'error' since results obtained for an infinite and continuous observational domain are those that are desired. Results for a one-pass Barnes scheme and the three-dimensional lrf are shown in Fig. 4 as a function of scatter number, which is used as in Doswell and Lasher-Trapp (1997) and dictates the irregularity of the observation distribution. As Fig. 4 illustrates, the three-dimensional lrf produces analysis fields that are closer to the desired analysis fields than the one-pass Barnes analysis fields. Moreover, as scatter number increases and the observation distribution becomes more irregular, the performance of the lrf, according to this measure, remains relatively constant while the performance of the one-pass Barnes scheme deteriorates rapidly.

Because weight functions like that of Barnes (1964) are oftentimes employed in successive correction schemes that use multiple passes, the lrf has also been tested against those schemes. In these tests, the Barnes weight function is applied to one-dimensional data, no background field is used, and the smoothing parameter used in the Barnes weight function is held constant with each pass. Then, following Caracena (1987), the equivalent one-pass weights are computed. The equivalent one-pass weights are the weights that, when used in (1), produce the same analysis field in one analysis pass as that produced by the successive corrections scheme in multiple analysis passes. These equivalent one-pass weights are used, along with (4) and (5), to diagnose the amplitude and phase modulations of the analysis fields produced by the successive corrections scheme. Finally, a relation derived by Pedder (1993) is used to compute the amplitude modulations the successive corrections scheme would produce if the observational domain were infinite and continuous. These amplitude modulations are then used to drive the lrf. An example of results obtained is provided in Fig. 5. As in the case of the one-pass analysis tests (Fig. 1), the lrf produces amplitude and phase modulations that are much closer to those desired than does the successive corrections method. Furthermore, because the lrf reduces the spatial variability of the amplitude and phase modulations, it produces more coherent analyses.

In addition to the lrf, a one-dimensional nonlinear response filter (hereinafter nrf) that utilizes J_{RF} (6) has been developed. This has been accomplished using minpack, a minimization software package for multi-dimensional nonlinear functions developed at Argonne National Laboratory that utilizes a modified Levenberg-Marquardt algorithm (Marquardt 1963). An example of test results is shown in Fig. 6. As Fig. 6a indicates, the lrf and nrf weights are nearly equal. The small differences in the weights for these two schemes, however, do result in differences in amplitude (Fig. 6b) and phase (Fig. 6c) modulations. As illustrated in Fig. 6b, the lrf and nrf schemes produce amplitude modulations that are much closer to those desired than does the one-pass Barnes scheme. The lrf and nrf schemes have the greatest difficulty at higher frequencies, which results in part because the I_{M_k} and I_{P_k} values are weighted towards lower frequencies in this test and in part because these schemes are expected to struggle at higher frequencies because of the irregularity of the data spacing. With regard to phase modulation, the lrf and nrf schemes also vastly outperform the one-pass Barnes scheme, except at the highest frequencies (Fig. 6c). The performance at the very high frequencies is not considered to be highly detrimental because, as requested, the lrf and nrf schemes significantly reduce the amplitudes of the waves at these frequencies and thus the phase-shifted results in the analyses at these frequencies show up with very little amplitude. In fact, the more significant phase shifts for the lrf and nrf schemes likely result because these schemes are more successful at reducing the amplitudes of waves at these frequencies and in doing so increase the likelihood of the numerator in (5) becoming larger relative to the denominator and thus of phase modulations becoming larger.

Finally, the increased performance provided by the response filter comes with a cost. Timing tests for one-dimensional analyses having 801 analysis locations indicate that the lrf, when used with (14), takes about ten times longer than a three pass successive corrections scheme. If, however, the lrf is used with (15) and the previously stated conditions for which $\underline{\underline{\mathbf{A}}}$ only has to computed once hold, then the lrf takes only about four times longer than a three pass successive corrections scheme.

5. DISCUSSION

The response filter is not necessarily intended to replace other schemes that adapt to data distributions, like statistical objective analysis, but is meant to be a high-quality alternative that can be especially useful when implementation of other schemes is difficult. For instance, implementation of the statistical objective analysis scheme can be complicated by lack of knowledge regarding background and observation error covariances and by data that do not satisfy the assumptions commonly used in this scheme (Daley 1991, §4.2, §4.9).

It is noted that for both $J_{RF}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]$ and $J_{RF\_lin}\left[w_{N_1},\ldots,w_{N_{N_o}}\right]$, these functionals are expressed as depending only upon the weights and not upon the analysis location *x*, the observation locations $\left\{x_{o_1},\ldots,x_{o_{N_o}}\right\}$, or elements like I_{M_k} and I_{P_k}. The reason for this is that once the analysis conditions (e.g., *x*, I_{M_k}, M_k, etc.) are set, the remaining problem is to find the weights. One might argue that these functionals should be expressed as depending upon *x* and $\left\{x_{o_1},\ldots,x_{o_{N_o}}\right\}$. This, in fact, is an excellent argument since how well a set of weights is able to replicate the requested amplitude and phase modulations depends to a large extent upon the distribution of the observations around an analysis location. However, because during an analysis observations generally cannot be moved and because moving an analysis location defeats the purpose of obtaining an analysis value at that location, these functionals are expressed as being dependent solely on the weights.

As stated earlier, the mathematical development of the multidimensional lrf is very similar to that for the one-dimensional lrf, with only additional housekeeping associated with the extra dimensions. The implementation of the multidimensional lrf, however, is somewhat more complicated owing to the need to select frequencies for constraining the lrf from a multidimensional frequency space. Numerous strategies can be employed, including what Lin (2004) calls the 'circle' and 'box' methods for two-dimensional analyses and the 'sphere' and 'cuboid' methods for three-dimensional analyses. The interested reader is referred to Lin (2004) for more details.

It is important to note that the performance of the lrf depends upon the I_{M_k} and I_{P_k}. The results presented herein were obtained with I_{M_k} and I_{P_k} values that more heavily weight the lower frequencies in (8). If this approach is not followed, then the lrf can produce unrealistic analysis values near data boundaries and large data gaps. It is thought that these failures result because the lrf is attempting to handle scales for which little information is available from the irregularly-spaced observations. This issue remains as a topic for future investigation.

Tests have shown that the nrf takes longer to run than the lrf and occasionally fails to converge to a solution. Furthermore, with the nrf there are the additional uncertainties (relative to the lrf) regarding the dependence of a solution upon the initial guess and regarding whether a local, rather than the global, minimum is obtained. Because of these factors and because the lrf and nrf produce very similar results (e.g., Fig. 6), development has, and is continuing to be, focused on the lrf.

Numerous issues remain regarding the lrf. These include the dependence of its efficacy on observation distribution irregularity, potential trade-offs between performance at one frequency versus another, anisotropic application to multidimensional data, and performance with more complicated (e.g., real) data sets. Investigations of these issues, via the integration of the lrf into analysis subsystems of the Local Analysis and Prediction System (LAPS; http://laps.fsl.noaa.gov/) for research supported by the Army High Performance Computing Research Center AHPCRC), are continuing.

6. CONCLUSIONS

The following summarize the results of this work:

1) A filter based upon providing desired amplitude and phase modulations can be designed even for irregularly-spaced data.
2) The response filter attempts to provide, as closely as possible, the desired amplitude and phase modulations given the observation distribution.
3) The response filter (both lrf and nrf) can provide vastly superior analyses relative to single- and multiple-pass successive corrections schemes.
4) The lrf is extensible to multiple dimensions.
5) The response filter (both lrf and nrf) can struggle in situations (e.g., near data boundaries or for small-wavelength waves) where the observations do not provide sufficient information regarding waves of interest. These problems are not specific to the response filter, but instead owe to a lack of data, for which no scheme can compensate.

ACKNOWLEDGEMENTS

This work is supported by the Army High Performance Computing Research Center (AHPCRC) under the auspices of the Department of the Army, Army Research Laboratory (ARL) under Cooperative

Agreement number **DAAD19-01-2-0014**. The content does not necessarily reflect the position or the policy of the government and no official endorsement should be inferred.

REFERENCES

Askelson, M. A., and J. M. Straka, 2005: Response functions for arbitrary weight functions and data distributions. Part I: Framework for interpreting the response function. *Mon. Wea. Rev.*, in press.

———, P. M. Pauley, and J. M. Straka, 2005: Response functions for arbitrary weight functions and data distributions. Part II: Response function derivation and verification. *Mon. Wea. Rev.*, in press.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396-409.

Caracena, F., 1987: Analytic approximation of discrete field samples with weighted sums and the gridless computation of field derivatives. *J. Atmos. Sci.*, **44**, 3753-3768.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367-374.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge Atmospheric and Space Science Series, Vol. 2, Cambridge University Press, 457 pp.

Doswell, C. A., III, and F. Caracena, 1988: Derivative estimation from marginally sampled vector point functions. *J. Atmos. Sci.*, **45**, 242-253.

———, and S. Lasher-Trapp, 1997: On measuring the degree of irregularity in an observing network**.** *J. Atmos. Oceanic Technol.*, **14**, 120–132.

Lin, H., 2004: The response filter–An approach for adapting to irregular data distributions. M.S. thesis, Dept. of Atmospheric Sciences, University of North Dakota, 81 pp.

Marquardt, D. W., 1963: *J. Soc. Ind. Applied Math.*, **11**, 431-441.

Pedder, M. A., 1993: Interpolation and filtering of spatial observations using successive corrections and Gaussian filters. *Mon. Wea. Rev.*, **121**, 2889–2902.

Salas, S. L., E. Hille, and J. T. Anderson, 1986: *Calculus: One and Several Variables*. 5th Ed., John Wiley and Sons, 1079 pp.

Solum, M., 2005: A comparison between two analysis techniques: The response filter and the multipass Barnes scheme. Senior Project Final Report, Dept. of Atmospheric Sciences, University of North Dakota, 36 pp.

Steiglitz, K., 1982: Phase unwrapping by factorization. *IEEE Trans. Acoustics, Speech, Signal. Proc.*, **ASSP-30**, 984-991.

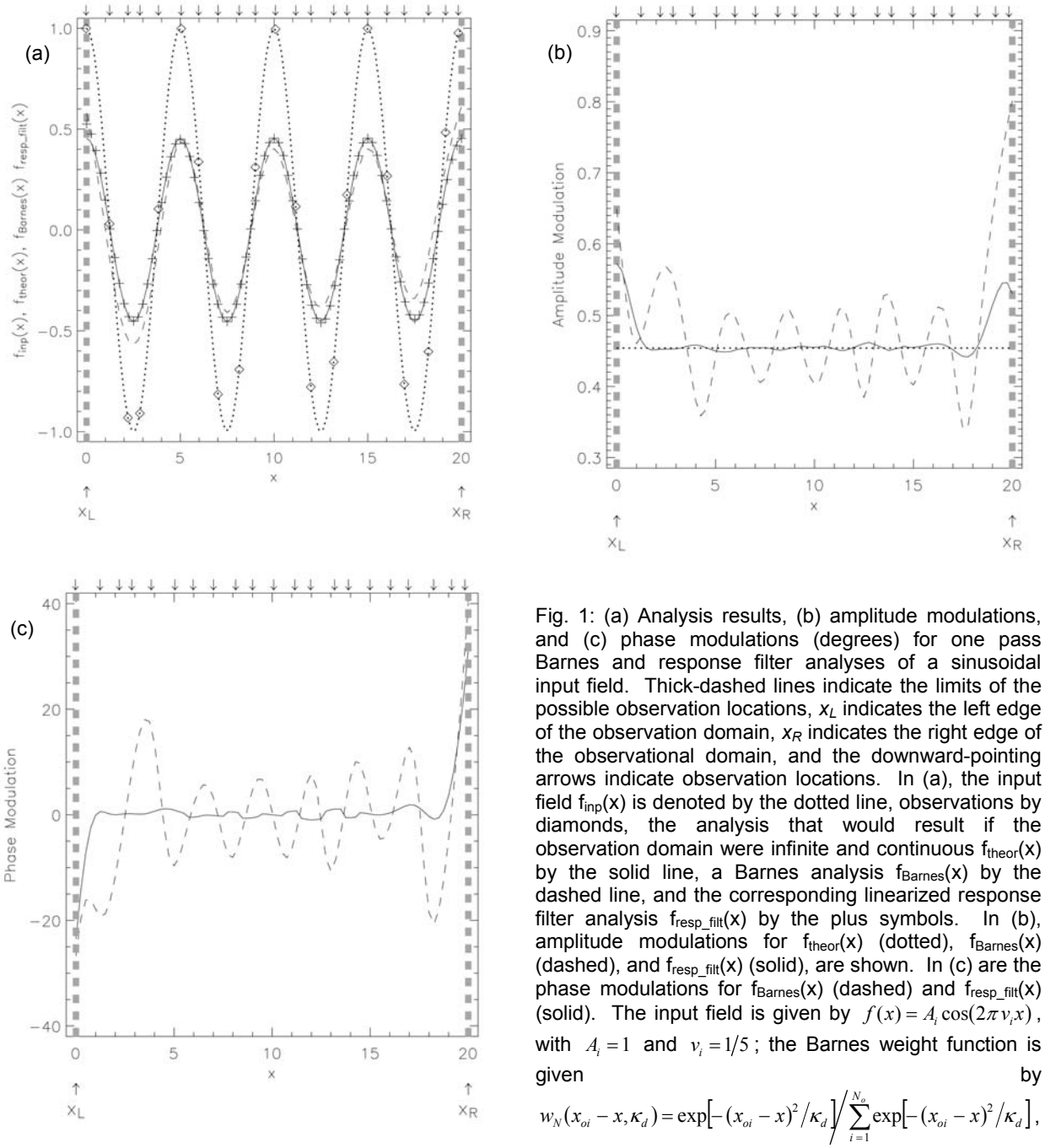


Fig. 1: (a) Analysis results, (b) amplitude modulations, and (c) phase modulations (degrees) for one pass Barnes and response filter analyses of a sinusoidal input field. Thick-dashed lines indicate the limits of the possible observation locations, x_L indicates the left edge of the observation domain, x_R indicates the right edge of the observational domain, and the downward-pointing arrows indicate observation locations. In (a), the input field $f_{inp}(x)$ is denoted by the dotted line, observations by diamonds, the analysis that would result if the observation domain were infinite and continuous $f_{theor}(x)$ by the solid line, a Barnes analysis $f_{Barnes}(x)$ by the dashed line, and the corresponding linearized response filter analysis $f_{resp\_filt}(x)$ by the plus symbols. In (b), amplitude modulations for $f_{theor}(x)$ (dotted), $f_{Barnes}(x)$ (dashed), and $f_{resp\_filt}(x)$ (solid), are shown. In (c) are the phase modulations for $f_{Barnes}(x)$ (dashed) and $f_{resp\_filt}(x)$ (solid). The input field is given by $f(x) = A_i \cos(2\pi v_i x)$, with $A_i = 1$ and $v_i = 1/5$; the Barnes weight function is given by $w_N(x_{oi} - x, \kappa_d) = \exp\left[-(x_{oi} - x)^2/\kappa_d\right] \Big/ \sum_{i=1}^{N_o} \exp\left[-(x_{oi} - x)^2/\kappa_d\right]$, with $\kappa_d = 2$.

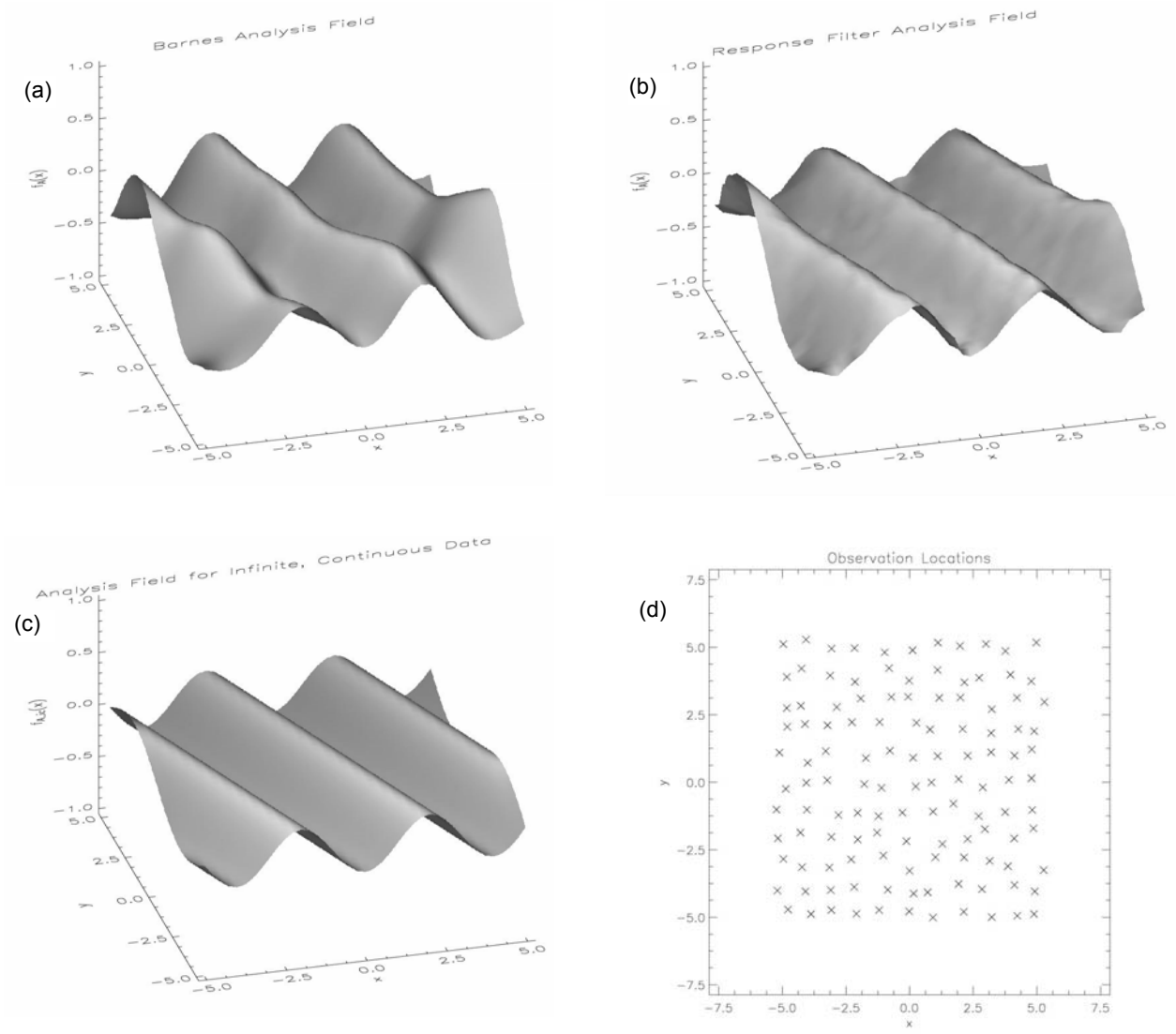


Fig. 2: One-pass Barnes analysis (a), two-dimensional lrf analysis (b), and theoretical analysis (realized for observations on an infinite, continuous domain) (c) for a two-dimensional sinusoidal field sampled at the locations indicated in (d). The input field is given by $f(x,y) = A_i \cos[2\pi(u_i x + v_i y)]$, with $A_i = 1$, $u_i = 1/4$, and $v_i = 1/10$; the Barnes weight function is given by $w_N(x_{oi} - x, y_{oi} - y, \kappa_x, \kappa_y) = \exp\left[-(x_{oi} - x)^2/\kappa_x - (y_{oi} - y)^2/\kappa_y\right] \Big/ \sum_{i=1}^{N_o} \exp\left[-(x_{oi} - x)^2/\kappa_x - (y_{oi} - y)^2/\kappa_y\right]$, with $\kappa_x = \kappa_y = 2$.

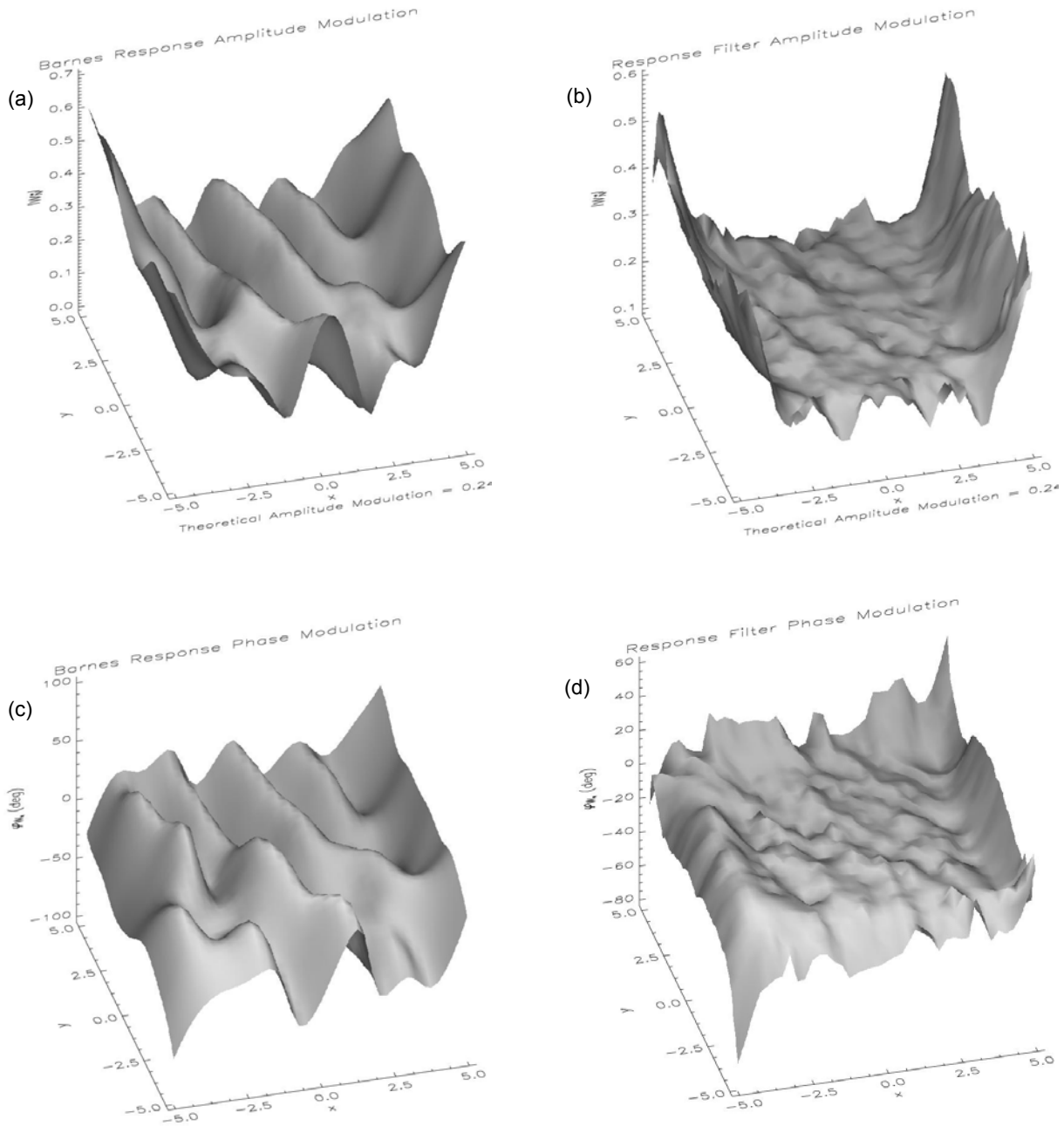


Fig. 3: Amplitude modulation at the frequency of the input wave as a function of location for (a) one-pass Barnes analysis and (b) two-dimensional lrf analysis; phase modulation at the frequency of the input wave as a function of location for (c) one-pass Barnes analysis and (d) two-dimensional lrf analysis. Values correspond to the analyses depicted in Fig. 2.

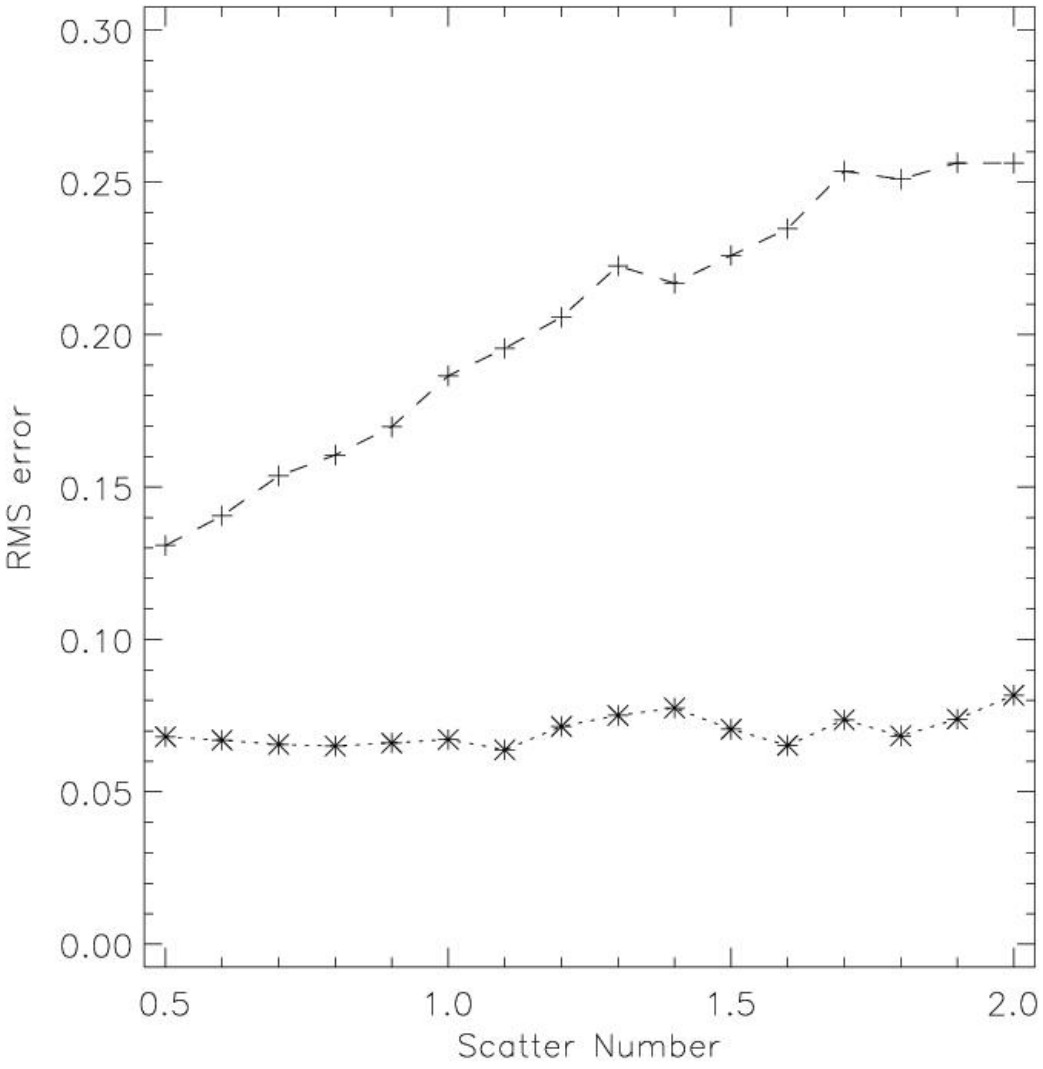


Fig. 4: RMS errors for one-pass Barnes and three-dimensional lrf analyses as a function of scatter number, which dictates the degree of irregularity of the observation distribution. The input field is given by $f(x,y,z) = A_i \cos[2\pi(u_i x + v_i y + w_i z)]$, with $A_i = 1$, $u_i = 1/8$, $v_i = 1/10$, and $w_i = 1/15$; the Barnes weight function is given by $w_N(x_{oi} - x, y_{oi} - y, z_{oi} - z, \kappa_x, \kappa_y, \kappa_z) = \exp\left[-(x_{oi} - x)^2/\kappa_x - (y_{oi} - y)^2/\kappa_y - (z_{oi} - z)^2/\kappa_z\right] \Big/ \sum_{i=1}^{N_o} \exp\left[-(x_{oi} - x)^2/\kappa_x - (y_{oi} - y)^2/\kappa_y - (z_{oi} - z)^2/\kappa_z\right]$, with $\kappa_x = \kappa_y = \kappa_z = 0.93$. From Lin (2004).

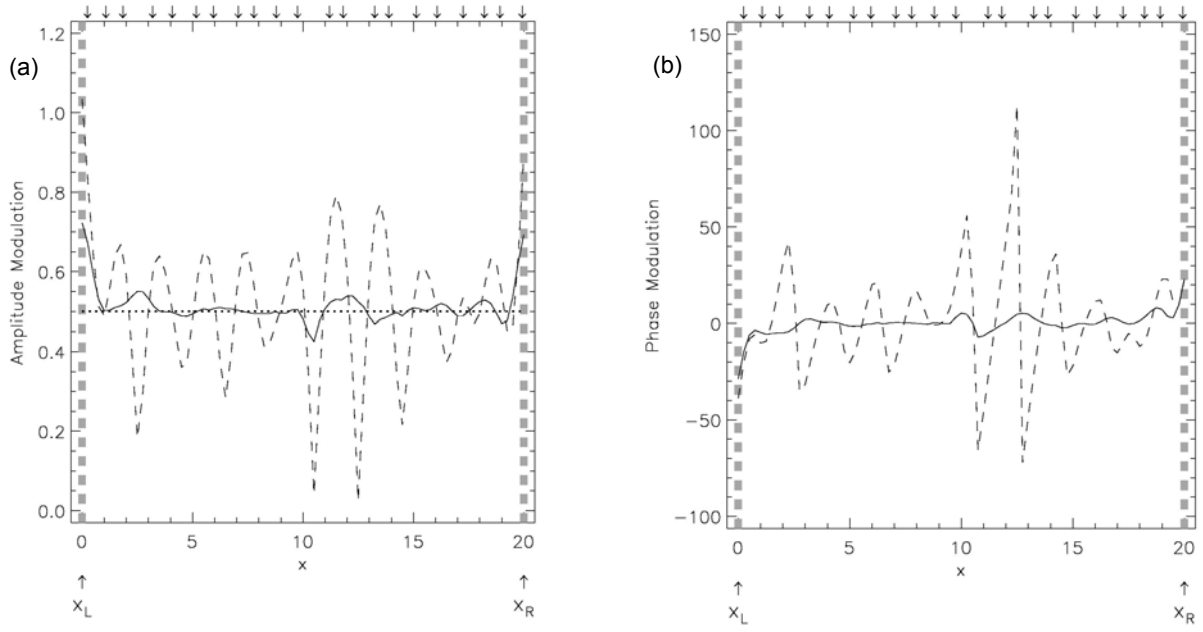
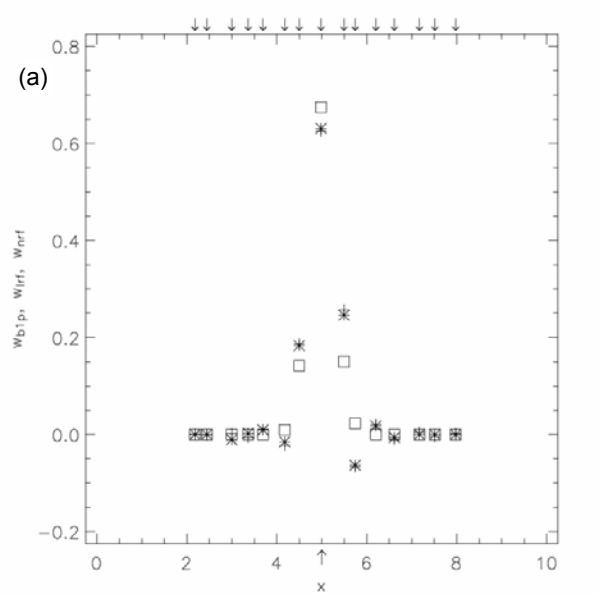
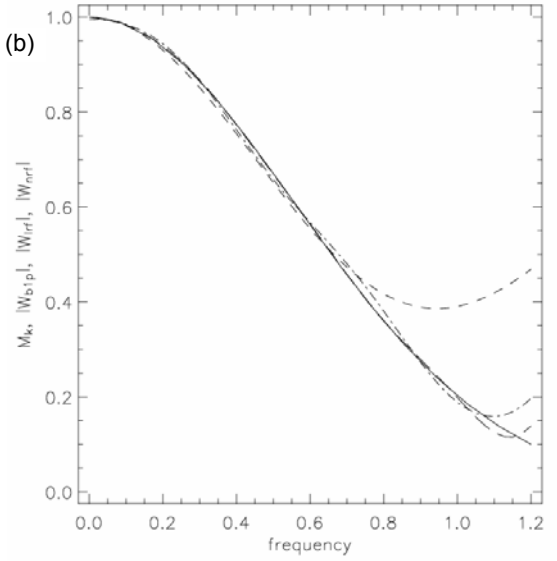


Fig. 5: Amplitude (a) and phase (b) modulations for a three-pass successive corrections method analysis produced using the Barnes weight function (dashed lines) and for the corresponding lrf analysis (solid line). The amplitude modulation for the theoretical analysis field, which would be produced if the observational domain were infinite and continuous, is indicated by the dotted line in (a). The phase modulation for the theoretical analysis field is zero. Other symbols are as in Fig. 1. The input field is given by $f(x) = A_i \cos(2\pi v_i x)$, with $A_i = 1$ and $v_i = 1/3$; the Barnes weight function is given by $w_N(x_{oi} - x, \kappa_d) = \exp\left[-(x_{oi} - x)^2/\kappa_d\right] \Big/ \sum_{i=1}^{N_o} \exp\left[-(x_{oi} - x)^2/\kappa_d\right]$, with $\kappa_d = 1.44$. From Solum (2005).





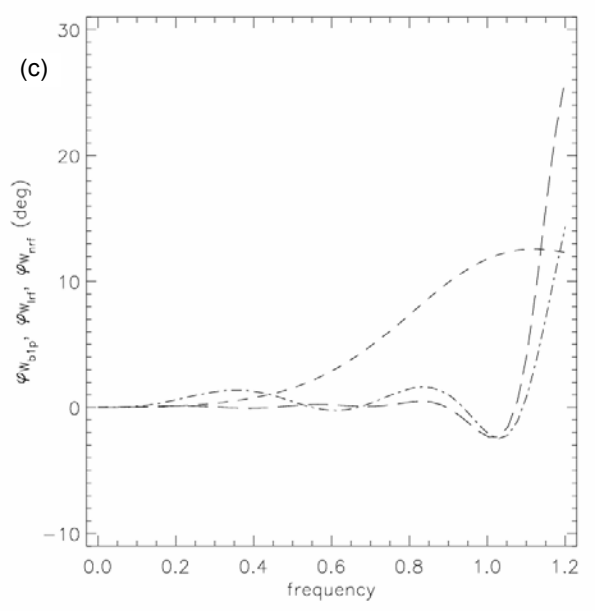


Fig. 6: Results for one-pass Barnes, lrf, and nrf schemes for analyses at x=5.0 [denoted by the upward-pointing arrow in (a)] and for a nominal (i.e., before observation locations are scattered) observational data spacing of 1.0 (nominal nyquist frequency = 1.2). (a) Weights for the one-pass Barnes scheme (boxes), the lrf (plus signs), and the nrf (asterisks) at the locations of observations used in the analyses (indicated by downward pointing arrows at the top of the plot). (b) Requested amplitude modulations M_k (solid line) and amplitude modulations of the one-pass Barnes scheme (dashed line), the lrf (long dashes), and the nrf (dash-dotted line) as a function of frequency. (c) Phase modulations of the one-pass Barnes scheme (dashed line), the lrf (long dashes), and the nrf (dash-dotted line) as a function of frequency. The requested phase modulation is zero at all frequencies.