Brigham Young University
Department of Electrical and Computer Engineering

459 Clyde Building
Provo, Utah 84602

The Field-wise Wind Retrieval Objective Function

Andrew Fletcher

20 Sept. 2000

MERS Technical Report # MERS 00-01
ECEN Department Report # TR-L130-00.01

Microwave Earth Remote Sensing (MERS) Laboratory

The Field-wise Wind Retrieval Objective Function

Andrew Fletcher

September 20, 2000

Abstract

The Field-wise Objective Function is an important concept in field-wise wind retrieval. A discussion on the purpose of the field-wise objective function is presented. Several variations of the field-wise objective function are compared. The gradient vectors and Hessian matrices are derived for each variation.

1 An Overview of Field-wise Objective Functions

An objective function is an error metric, providing a scalar value quantifying the distance between an estimate and the observed measurements. Thus, the estimate that minimizes the objective function is considered "closest" to the observed measurements. The well-known "least-squares" problem is an implementation of a minimized objective function.

In the case of field-wise wind retrieval, the error metric is a function of the model parameters. A model parameter vector $\mathbf{x}$ parameterizes an M$\times$N wind field $\mathbf{w}$ through the linear model $\mathbf{w} = F\mathbf{x}$. An objective function $J(\mathbf{x})$ measures the difference between the wind field $\mathbf{w}$ and the observed measurements.

A direct error metric between the wind field $\mathbf{w}$ and the observed radar backscatter (σ^o) is meaningless, because $\mathbf{w}$ is not in the σ^o measurement space. To create a metric between $\mathbf{w}$ and σ^o, they must be transformed into the same space. As there is no model to transorm σ^o into the wind vector space, all of the objective functions employ a metric in the σ^o measurement space.

The Geophysical Model Function (GMF) enables $\mathbf{w}$ to be transformed into the σ^o measurement space. The GMF returns the σ^o value that would result from examining a wind vector under a given set of measurement conditions (i.e., instrument azimuth and incidence angles, ψ and θ). The transformation from the wind field space into the σ^o measurement space is called the forward projection.

To mathematically represent the measurements in a region, let Z be a three-dimensional array containing the observed values (σ^o) for each measurement. The first two dimensions index the measurement location (along-track and cross-track position) of a wind field region. The third dimension indexes the measurement number in each cell. Thus Z_{ijk} is the k^{th} measurement of the ij^{th} cell in the region. (Note that the number of measurements may vary at each swath location).

To represent the forward projection of one wind vector measurement, let $\mathcal{M}(U, \psi - \phi, \theta)$ be the GMF, where ϕ is the wind direction and U is the wind speed. As $\mathbf{x}$ parameterizes every wind vector in the region, the forward projection of the k^{th} measurement of the ij^{th} cell can be denoted $\mathcal{M}_{ijk}(\mathbf{x})$ where the measurement geometry is implied by k.

1.1 The Squared Error (SE) Objective Function

Perhaps the simplest and most common objective function is the squared error objective function. This error metric can be described as the Euclidean distance or the $\mathcal{L}_2$ norm, which gives rise to the "least squares" solution. Thus, the objective function can be written as

$$J_{SE}(\mathbf{x}) = \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} (Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))^2, \tag{1}$$

where K_{ij} is the number of measurements per cross-track cell, ij.

1.2 The Weighted Squared Error (WSE) Objective Function

While certainly the simplest option, the squared error objective function fails to make use of all available information, and, as a result, can be overly sensitive to noise. The measuring instrument introduces noise that has been well studied. The noise is represented by a zero-mean, Gaussian random variable ν, with variance (K_{PC}^2) given by

$$\varsigma^2 = \alpha(\sigma_T^o)^2 + \beta\sigma_T^o + \gamma. \tag{2}$$

The parameters α, β, and γ are functions of the instrument design and signal to noise ratio (SNR), and σ_T^o is the "true" σ^o measurement (i.e. the σ^o that would be observed in the absence of measurement noise). Thus, σ^o is a realization of the random variable equation

$$\sigma^o = \sigma_T^o + \nu. \tag{3}$$

Using the variance estimate from Eq. (2), instrument noise can be accounted for, by dividing each term in the squared error sum by the measurement variance. Thus,

$$J_{WSE}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2, \tag{4}$$

represents an objective function that can be classified as a "weighted squared error." It may be valuable to note that minimizing this objective function can be considered a maximum likelihood estimator, assuming that the variance of each measurement is constant with respect to $\mathbf{x}$. This assumption will be examined in greater detail in the following section. $J_{WSE}(\mathbf{x})$ is also a quantity known as a "chi-square" (χ^2).

1.3 Maximum Likelihood (ML) Estimation

In the preceding section, the weighted squared error objective function was casually mentioned to be a maximum likelihood (ML) estimator given a constant measurement variance. The ML estimator is explicitly derived in this section.

The ML estimator calculates the model parameters most likely to give rise to the observed measurements. For a given $\mathbf{x}$, the estimator evaluates the probability that the observed measurements Z would occur. The estimated quantity $\hat{\mathbf{x}}_{ML}$ is the $\mathbf{x}$ that maximizes this probability. Thus,

$$\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x}} p_Z(Z|\mathbf{x}). \tag{5}$$

If the measurements are assumed to be independently Gaussian, with variance ς^2 defined above, then

$$p_Z(Z|\mathbf{x}) = \prod_{i=1}^{N}\prod_{j=1}^{M}\prod_{k=1}^{K_{ij}} \frac{1}{\sqrt{2\pi\varsigma_{ijk}^2}} \exp\left\{-\frac{1}{2}\left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2\right\}. \tag{6}$$

Computing the maximum of $p_Z(Z|\mathbf{x})$ is equivalent to computing the minimum of the negative log-likelihood function $\mathcal{L}(\mathbf{x}) = -\ln p_Z(Z|\mathbf{x})$, which is

$$\mathcal{L}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left\{\frac{1}{2}\ln 2\pi + \frac{1}{2}\ln \varsigma_{ijk}^2 + \frac{1}{2}\left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2\right\}. \tag{7}$$

Note that the first two terms in the sum are constant with respect to $\mathbf{x}$, so they may be disregarded when calculating the arg min. The common scale factor of $\frac{1}{2}$ may also be ignored. Therefore,

$$\begin{aligned}\hat{\mathbf{x}}_{ML} &= \arg\min_{\mathbf{x}} p_Z(Z|\mathbf{x}) \\ &= \arg\min_{\mathbf{x}} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} \left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}} \right)^2 \right\} \\ &= \arg\min_{\mathbf{x}} \{ J_{SE}(\mathbf{x}) \}, \end{aligned} \tag{8}$$

the weighted squared error objective function.

Before declaring the weighted squared error a maximum likelihood estimator, the constant variance assumption needs to be examined in greater detail. Recall from Eq. (2) that ς^2 depends upon the value of σ_T^o. Also recall that in computing $p_Z(Z|\mathbf{x})$, we estimate the probability of the observed measurements under the assumption that the true wind field is represented by $\mathbf{x}$. Under this assumption, $\sigma_T^o = \mathcal{M}(\mathbf{x})$. Thus, ς^2 is a function of $\mathbf{x}$:

$$\varsigma_{ijk}^2(\mathbf{x}) = \alpha \mathcal{M}_{ijk}^2(\mathbf{x}) + \beta \mathcal{M}_{ijk}(\mathbf{x}) + \gamma. \tag{9}$$

This dependence on $\mathbf{x}$ changes the simplification of $\mathcal{L}(\mathbf{x})$; the $\frac{1}{2} \ln \varsigma^2$ term must be retained the minimization. Thus the objective function for maximum likelihood estimation is

$$J_{ML}(\mathbf{x}) = \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} \left\{ \left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})} \right)^2 + \ln \varsigma_{ijk}^2(\mathbf{x}) \right\}. \tag{10}$$

While both Eqs. (4) and (10) can be said to represent objective functions of maximum likelihood estimation, the constant variance assumption in Eq. (4) is inconsistent with the probability model $p_Z(Z|\mathbf{x})$. Therefore, for the duration of this paper "maximum likelihood" will refer exclusively to Eq. (10).

1.4 The Reduced Maximum Likelihood (RML) Objective Function

While a theoretically sound objective function, in practice Eq. (10) presents some difficulties. Examining the scale of the terms in the summation reveals one reason. The first term is the square of a zero-mean, unit-variance Gaussian random variable, thus a χ^2 random variable with one degree of freedom, which has mean value of 1. ς^2 is on the order of 10^{-5}, so $\ln\varsigma^2$ is on the order of -11.5, so summed over all of the measurements, the $\ln\varsigma^2$ dominates. The parameters α, β, and γ are only rough approximations, though, so the dominant term is not as accurate as the χ^2 term.

For this reason, the final objective function analyzed here is the reduced maximum likelihood objective function:

$$J_{RML}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})}\right)^2. \tag{11}$$

2 Objective Function Gradients

As mentioned before, to be useful as an estimation tool, the objective function minima must be obtained. Many minimization routines require the calculation of the objective function gradient. Below, the gradient is analytically derived for the four cited objective functions.

2.1 SE Objective Function Gradient

Evaluation of the gradient requires a straightforward application of the chain rule, differentiating with respect to each model parameter. With respect to the p^{th} model parameter, the partial derivative of Eq. (1) is:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{SE}(\mathbf{x}) = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial}{\partial \mathbf{x}_p}\mathcal{M}_{ijk}(\mathbf{x}) \tag{12}$$

where

$$\frac{\partial}{\partial \mathbf{x}_p}\mathcal{M}_{ijk}(\mathbf{x}) = \frac{\partial \mathcal{M}(u_{ij}, v_{ij})}{\partial u_{ij}}\frac{\partial u_{ij}}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}(u_{ij}, v_{ij})}{\partial v_{ij}}\frac{\partial v_{ij}}{\partial \mathbf{x}_p}. \tag{13}$$

The terms u_{ij} and v_{ij} represent the rectangular components of the wind field at the ij_{th} wvc. Note that these may be represented in terms of the F matrix representing any linear wind field model:

$$u_{ij} = F_l^T \mathbf{x}, \tag{14}$$

$$v_{ij} = F_{l+MN}^T \mathbf{x}, \tag{15}$$

$$l = N(i-1) + j, \tag{16}$$

where l is the index into the column scanned representation of the wind region and F_l^T is the l^{th} row of the wind field model transform matrix. Therefore,

$$\frac{\partial u_{ij}}{\partial \mathbf{X}_p} = F_{l,p} \tag{17}$$

$$\frac{\partial v_{ij}}{\partial \mathbf{X}_p} = F_{l+MN,p}. \tag{18}$$

The Geophysical Model Function is an empirically derived table of values with no closed form solution. The table has three dimensions: wind speed U, relative azimuth χ (instrument azimuth ψ - wind direction ϕ), and incidence angle θ. In order to evaluate the function, an interpolation routine must be used. In the MERS lab, a bspline function is used, interpolating in all three directions. Through this function, partial derivatives can be easily obtained with respect to wind speed and relative azimuth, i.e. $\frac{\partial \mathcal{M}}{\partial s}$ and $\frac{\partial \mathcal{M}}{\partial \chi}$. These are related to the rectangular components by

$$s = \sqrt{u^2 + v^2}, \tag{19}$$

$$\chi = \psi - \phi, \tag{20}$$

$$\phi = \tan^{-1}(\frac{v}{u}). \tag{21}$$

Thus,

$$\frac{\partial \mathcal{M}}{\partial v} = \frac{\partial \mathcal{M}}{\partial s}\frac{\partial s}{\partial v} + \frac{\partial \mathcal{M}}{\partial \phi}\frac{\partial \phi}{\partial v} \tag{22}$$

$$\frac{\partial \mathcal{M}}{\partial u} = \frac{\partial \mathcal{M}}{\partial s}\frac{\partial s}{\partial u} + \frac{\partial \mathcal{M}}{\partial \phi}\frac{\partial \phi}{\partial u} \tag{23}$$

$$\frac{\partial s}{\partial u} = \frac{u}{\sqrt{u^2 + v^2}} \tag{24}$$

$$\frac{\partial s}{\partial v} = \frac{v}{\sqrt{u^2 + v^2}} \tag{25}$$

$$\frac{\partial \mathcal{M}}{\partial \phi} = \frac{\partial \mathcal{M}}{\partial \chi}\frac{\partial \chi}{\partial \phi} = -\frac{\partial \mathcal{M}}{\partial \chi}. \tag{26}$$

The partials of ϕ with respect to u and v require more careful attention. The inverse tangent with only one argument, defined on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, has a well known derivative:

$$\frac{d}{dx}\tan^{-1}(x) = \frac{1}{1+x^2}. \tag{27}$$

For purposes in wind retrieval, the four quadrant inverse tangent (defined on the interval $[-\pi, \pi]$, and denoted $\tan_4^{-1}$) is necessary. This can be defined in the following way:

$$\tan_4^{-1}(\frac{v}{u}) = \begin{cases} \tan^{-1}(\frac{v}{u}) & \text{Quadrants I and IV,} \\ \tan^{-1}(\frac{v}{u}) + \pi & \text{Quadrant II,} \\ \tan^{-1}(\frac{v}{u}) - \pi & \text{Quadrant III} \end{cases} \tag{28}$$

Thus the partial derivatives will be the same in all quadrants, i.e.

$$\begin{aligned} \frac{\partial \phi}{\partial u} = \frac{\partial}{\partial u}\tan_4^{-1}(\frac{v}{u}) &= \frac{1}{1+\frac{v^2}{u^2}}\frac{-v}{u^2} \\ &= \frac{-v}{u^2+v^2}. \end{aligned} \tag{29}$$

$$\begin{aligned} \frac{\partial \phi}{\partial v} = \frac{\partial}{\partial v}\tan_4^{-1}(\frac{v}{u}) &= \frac{1}{1+\frac{v^2}{u^2}}\frac{1}{u} \\ &= \frac{u}{u^2+v^2}. \end{aligned} \tag{30}$$

Therefore,

$$\frac{\partial \mathcal{M}}{\partial v} = \frac{\partial \mathcal{M}}{\partial s}\frac{v}{\sqrt{u^2+v^2}} + \frac{\partial \mathcal{M}}{\partial \phi}\frac{u}{u^2+v^2}, \tag{31}$$

$$\frac{\partial \mathcal{M}}{\partial u} = \frac{\partial \mathcal{M}}{\partial s}\frac{u}{\sqrt{u^2+v^2}} - \frac{\partial \mathcal{M}}{\partial \phi}\frac{v}{u^2+v^2}. \tag{32}$$

2.2 WSE Objective Function Gradient

The WSE objective function differs from the SE objective function by only the ς term which is constant with respect to $\mathbf{x}$. Thus, the gradient differs from Eq. (12) by the same term:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{WSE}(\mathbf{x}) = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left[\frac{(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))}{\varsigma_{ijk}^2}\right]\frac{\partial}{\partial \mathbf{x}_p}\mathcal{M}_{ijk}(\mathbf{x}). \tag{33}$$

2.3 RML Objective Function Gradient

The RML objective function differs from the WSE objective function only in that ς^2 depends upon $\mathbf{x}$. Computation of the gradient requires use of the derivative quotient rule:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{RML}(\mathbf{x}) = -\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left\{ \frac{2(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial}{\partial \mathbf{x}_p}\mathcal{M}_{ijk}(\mathbf{x}_p)}{\varsigma_{ijk}^2} + \frac{(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))^2 \frac{\partial \varsigma_{ijk}^2}{\partial \mathbf{x}_p}}{(\varsigma_{ijk}^2)^2} \right\}, \quad (34)$$

where

$$\frac{\partial \varsigma_{ijk}^2}{\partial \mathbf{x}_p} = 2\alpha \mathcal{M}_{ijk}(\mathbf{x}) \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p} + \beta \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}. \quad (35)$$

2.4 ML Objective Function Gradient

Differentiaing Eq. (10) requires only the addition of one term to Eq. (34):

$$\frac{\partial}{\partial \mathbf{x}_p} J_{ML}(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}_p} J_{RML}(\mathbf{x}) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \frac{1}{\varsigma_{ijk}^2(\mathbf{x})} \frac{\partial \varsigma_{ijk}^2(\mathbf{x})}{\partial \mathbf{x}_p}. \quad (36)$$

3 Objective Function Hessian Matrices

Several minimization algorithms for the objective function require a realization of the Hessian matrix, or the matrix of double partials. Although the derivation is involved, like the gradient, it is a straightforward implementation of the chain rule.

3.1 SE Objective Function Hessian

To completely specify the derivation, it is sufficient to derive expressions for the following:

$$(37)$$

$$\frac{\partial^2 J_{SE}}{\partial \mathbf{x}_p^2} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left[(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p^2} - \left(\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}\right)^2 \right],$$

and

$$\frac{\partial^2 J_{SE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = -2 \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} \left[(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})) \frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p \partial \mathbf{x}_q} - \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p} \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q} \right]. \tag{38}$$

where $p \neq q$. By the chain rule,

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_p^2} &= \frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial u} \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial v} \frac{\partial v}{\partial \mathbf{x}_p} \right) \\
&= \frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial u} \right) \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial u} \frac{\partial^2 u}{\partial \mathbf{x}_p^2} + \frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial v} \right) \frac{\partial v}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial v} \frac{\partial^2 v}{\partial \mathbf{x}_p^2} \\
&= \frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial u} \right) \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial v} \right) \frac{\partial v}{\partial \mathbf{x}_p}, \qquad (39) \\
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_q \partial \mathbf{x}_p} &= \frac{\partial}{\partial \mathbf{x}_q} \left(\frac{\partial \mathcal{M}}{\partial u} \right) \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial}{\partial \mathbf{x}_q} \left(\frac{\partial \mathcal{M}}{\partial v} \right) \frac{\partial v}{\partial \mathbf{x}_p}. \qquad (40)
\end{aligned}$$

Note that the double partials, $\frac{\partial^2 v}{\partial \mathbf{x}_p^2}$ and $\frac{\partial^2 u}{\partial \mathbf{x}_p^2}$ are both 0. The mixed partials of the model function are further developed as

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial u} \right) &= \frac{\partial^2 \mathcal{M}}{\partial u^2} \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial u \partial v} \frac{\partial v}{\partial \mathbf{x}_p}, \qquad (41) \\
\frac{\partial}{\partial \mathbf{x}_p} \left(\frac{\partial \mathcal{M}}{\partial v} \right) &= \frac{\partial^2 \mathcal{M}}{\partial u \partial v} \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2} \frac{\partial v}{\partial \mathbf{x}_p}, \qquad (42)
\end{aligned}$$

and thus, the above simplifies to

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_p^2} &= \frac{\partial^2 \mathcal{M}}{\partial u^2} \left(\frac{\partial u}{\partial \mathbf{x}_p} \right)^2 + 2 \frac{\partial^2 \mathcal{M}}{\partial u \partial v} \frac{\partial v}{\partial \mathbf{x}_p} \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2} \left(\frac{\partial v}{\partial \mathbf{x}_p} \right)^2, \qquad (43) \\
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_q \partial \mathbf{x}_p} &= \frac{\partial^2 \mathcal{M}}{\partial u^2} \frac{\partial u}{\partial \mathbf{x}_q} \frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial u \partial v} \frac{\partial v}{\partial \mathbf{x}_q} \frac{\partial u}{\partial \mathbf{x}_p} \\
&\quad + \frac{\partial^2 \mathcal{M}}{\partial u \partial v} \frac{\partial u}{\partial \mathbf{x}_q} \frac{\partial v}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2} \frac{\partial v}{\partial \mathbf{x}_q} \frac{\partial v}{\partial \mathbf{x}_p}. \qquad (44)
\end{aligned}$$

When using the bspline version of the Geophysical Model Function, the double partials $\frac{\partial^2 \mathcal{M}}{\partial u^2}, \frac{\partial^2 \mathcal{M}}{\partial v^2}$, and $\frac{\partial^2 \mathcal{M}}{\partial u \partial v}$ are not directly available, as the model function is splined with respect to s and χ. Thus, expressions for these partials must also be derived to implement the Hessian matrix.

$$
\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial u^2} &= \frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial u}\right) \\
&= \frac{\partial}{\partial u}\left(\frac{u}{\sqrt{u^2+v^2}}\frac{\partial \mathcal{M}}{\partial s} - \frac{v}{u^2+v^2}\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial s}\right) + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad -\frac{v}{(u^2+v^2)}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) + \frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} && (45) \\
\frac{\partial^2 \mathcal{M}}{\partial v^2} &= \frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial v}\right) \\
&= \frac{\partial}{\partial v}\left(\frac{v}{\sqrt{u^2+v^2}}\frac{\partial \mathcal{M}}{\partial s} + \frac{u}{u^2+v^2}\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial s}\right) + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad -\frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} + \frac{u}{u^2+v^2}\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right), && (46)
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial s}\right) &= \frac{\partial^2 \mathcal{M}}{\partial s^2}\frac{\partial s}{\partial u} + \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial \phi}{\partial u} \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial s^2} - \frac{v}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}, && (47) \\
\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) &= \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial s}{\partial u} + \frac{\partial^2 \mathcal{M}}{\partial \phi^2}\frac{\partial \phi}{\partial u} \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} - \frac{v}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}, && (48) \\
\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial s}\right) &= \frac{\partial^2 \mathcal{M}}{\partial s^2}\frac{\partial s}{\partial v} + \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial \phi}{\partial v} \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial s^2} + \frac{u}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}, && (49) \\
\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) &= \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial s}{\partial v} + \frac{\partial^2 \mathcal{M}}{\partial \phi^2}\frac{\partial \phi}{\partial v} \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} + \frac{u}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}. && (50)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial u^2} &= \frac{u^2}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial s^2} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} + \frac{v^2}{(u^2+v^2)^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2} + \frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} \\
&= \frac{u^2}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial s^2} + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} - \frac{2uv}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} \\
&\quad + \frac{v^2}{(u^2+v^2)^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2} + \frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi}, \qquad (51) \\
\frac{\partial^2 \mathcal{M}}{\partial v^2} &= \frac{v^2}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial s^2} + \frac{2uv}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad - \frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} + \frac{u^2}{(u^2+v^2)^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}. \qquad (52)
\end{aligned}
$$

The mixed partial is found to be

$$
\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial u \partial v} &= \frac{\partial}{\partial u}\left(\frac{v}{\sqrt{u^2+v^2}}\frac{\partial \mathcal{M}}{\partial s} + \frac{u}{u^2+v^2}\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial s}\right) - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad + \frac{u}{u^2+v^2}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) + \frac{v^2-u^2}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} \\
&= \frac{uv}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial s^2} - \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad - \frac{2u^2}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} - \frac{uv}{(u^2+v^2)^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2} \\
&= \frac{uv}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial s^2} + \frac{u^2-v^2}{(u^2+v^2)^{3/2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad + \frac{v^2-u^2}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} - \frac{uv}{(u^2+v^2)^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}. \qquad (53)
\end{aligned}
$$

Obtaining $\frac{\partial^2 \mathcal{M}}{\partial \phi^2}$ and $\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}$ from the model function requires only a simple application of the chain rule. Recall that

$$
\begin{aligned}
\chi &= \psi - \phi, \qquad (54) \\
\frac{\partial \mathcal{M}}{\partial \phi} &= -\frac{\partial \mathcal{M}}{\partial \chi}. \qquad (55)
\end{aligned}
$$

Therefore,

$$\begin{aligned}\frac{\partial^2 \mathcal{M}}{\partial \phi^2} &= \frac{\partial}{\partial \phi}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) \\ &= \frac{\partial}{\partial \phi}\left(-\frac{\partial \mathcal{M}}{\partial \chi}\right) \\ &= -\frac{\partial^2 \mathcal{M}}{\partial \chi^2}\frac{\partial \chi}{\partial \phi} \\ &= \frac{\partial^2 \mathcal{M}}{\partial \chi^2}, \qquad (56) \\ \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} &= \frac{\partial}{\partial \phi}\left(\frac{\partial \mathcal{M}}{\partial s}\right) \\ &= \frac{\partial^2 \mathcal{M}}{\partial \chi \partial s}\frac{\partial \chi}{\partial \phi} \\ &= -\frac{\partial^2 \mathcal{M}}{\partial \chi \partial s}. \qquad (57)\end{aligned}$$

3.2 The WSE Hessian Matrix

As with the gradient of J_{WSE}, the hessian requires only the addition of the constant ς^2 term to the J_{SE} hessian matrix:

$$\frac{\partial^2 J_{WSE}}{\partial \mathbf{x}_p^2} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left[\frac{(Z_{ijk}-\mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p^2}}{\varsigma_{ijk}^2} - \left(\frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}}{\varsigma_{ijk}^2}\right)^2\right], \qquad (58)$$

$$\frac{\partial^2 J_{WSE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left[\frac{(Z_{ijk}-\mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p \partial \mathbf{x}_q}}{\varsigma_{ijk}^2} - \frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q}}{\varsigma_{ijk}^2}\right]. \qquad (59)$$

4 Conclusion

While four objective functions have been derived, in general, only the Weighted Squared Error objective function and the Reduced Maximum Likelihood Objective function are used in field-wise wind retrieval. The Squared Error objective function can be overly sensitive to noise, while the Maximum Likelihood objective function is dominated by less accurate estimates.