DEPARTMENT OF STATISTICS

University of Wisconsin
1210 W. Dayton St.
Madison, WI 53706

(NASA-CR-169221) VECTOR SPLINES ON THE SPHERE WITH APPLICATION TO THE ESTIMATION OF VORTICITY AND DIVERGENCE FROM DISCRETE, NOISY DATA Scientific Interim Report (Wisconsin Univ.) 25 p HC A02/MF A01 — N82-33088 — Unclas G3/64 27799

TECHNICAL REPORT NO. 674

May 1982

VECTOR SPLINES ON THE SPHERE, WITH APPLICATION TO THE ESTIMATION OF VORTICITY AND DIVERGENCE FROM DISCRETE, NOISY DATA

by

Grace Wahba

This research was supported by the Office of Naval Research under Contract N00014-77-G-0675 and by the National Aeronautics and Space Administration under Contract NAG5-128.

VECTOR SPLINES ON THE SPHERE, WITH
APPLICATION TO THE ESTIMATION OF VORTICITY
AND DIVERGENCE FROM DISCRETE, NOISY DATA

Grace Wahba

Statistics Department
University of Wisconsin-Madison
USA

Vector smoothing splines on the sphere are defined. Theoretical properties are briefly alluded to. An approach to choosing the appropriate Hilbert space norms to use in a specific meteorological application is described and justified via a duality theorem. Numerical procedures for computing the splines as well as the cross validation estimate of two smoothing parameters are given. A Monte Carlo study is described which suggests the accuracy with which upper air vorticity and divergence can be estimated using measured wind vectors from the North American radiosonde network.

To appear in "Multivariate Approximation Theory", Vol. 2, proceedings of a conference at Oberwolfach February 7-13, 1982. W. Schempp and K. Zeller, eds. Birkhauser Verlag, Basel-Boston-Stuttgart.

1. Introduction

A theory of spline functions on the sphere is rapidly being developed, see WAHBA (1981a), FREEDEN (1981a,b), SHURE, PARKER AND BACKUS (1981). Dr. FREEDEN will be reporting on some of his results elsewhere in this volume. Much of the rich theory surrounding univariate splines and thin plate splines clearly is extendable to the theory of splines on the sphere, via the use of reproducing kernels, n-widths, etc. In particular convergence rates for smoothing splines on the sphere can be obtained from the known rate of decay of the eigenvalues of the relevant reproducing kernels, see e.g. MICCHELLI and WAHBA (1981), WAHBA (1977), UTRERAS (1981).

In this paper we propose a notion of vector splines on the sphere. It is clear that interesting approximation theoretical properties of these splines can be obtained. However, in this paper our focus will be on the solution of certain practical problems which must be solved so that these splines may be usefully applied to the analysis of meteorological data from the upper air radiosonde network.

For the purpose of numerical weather prediction the global radiosonde (weather balloon) network takes measurements every 12 hours of the horizontal wind velocity vectors and other variables, at 9 standardized vertical levels. From this data it is desired to estimate the horizontal wind field and its vorticity and divergence (and other variables) at a regular grid of points, for each level. These estimates on a grid are then merged with estimates of the same variables on the same grid, which have been obtained from a forecast, to provide an estimate of the present state of the atmosphere. This state estimate is then used as the initial conditions to a numerical integration scheme which integrates a set of differential equations describing the dynamics of the atmosphere, to provide a new forecast. Numerical weather forecasts can be quite sensitive to errors in the vorticity and divergence in the initial wind fields. Unfortunately, horizontal wind vectors at, for example the 500 millibar height, of the order of a few tens of meters per second, are measured with an error standard deviation in each component of the order of 2-4 meters per second. Thus, it is not a trivial matter to obtain accurate information concerning the vorticity and divergence from this data, even in areas such as the continental U.S. where the radiosonde network is relatively dense. We believe that the appropriate derivatives of the vector smoothing splines we propose have the potential for doing this relatively well.

Speaking intuitively, the vector smoothing splines we propose will behave like low pass filters. In the splines we propose there will be two regularization or smoothing parameters to be chosen and two (sets of) "shape" parameters. The first smoothing parameter to be chosen, may be thought of as governing the overall half power point of the low pass filter. The second parameter governs the relative distribution of power between vorticity and divergence in the estimate. The choice of the two sets of shape parameters correspond to the choice of Hilbert space norms, but in an important practical sense they govern the rates of decay of the energy spectrum of the solution, one "shape" for vorticity, and one "shape" for divergence. It is well known from the theory and practice of ill posed problems that the appropriate choice of certain of these parameters can affect the practical usefulness of the result.

In this paper we propose the use of generalized cross-validation (GCV) for choosing the two smoothing parameters. GCV can also be used to choose a small number of "shape" parameters (see CRAVEN and WAHBA (1979), WAHBA and WENDELBERGER (1980)). However, in this paper we show how historical meteorological data can be used to choose the "shape" parameters, or Hilbert space norm. We discuss some numerical methods, and we describe the results of some numerical experiments on synthetic data which mimics actual 500 millibar horizontal wind fields over the U.S. In our experiments we have observed that the accuracy in estimating both vorticity and divergence can be quite sensitive to the relative distribution of power allocated between them, (choice of second smoothing parameter) but that GCV can be quite effective in estimating the correct relative power distribution.

For the meteorological experts in the audience we remark that estimating the present state of the atmosphere from current data is not exactly the same problem as estimating the state of the atmosphere from a combination of present data and a forecast of the present. This is so because a data only estimate needs to take account of properties of the atmosphere and measurement system while a data plus forecast estimate needs to take into account the relative error of the data and the forecast. In this paper we are studying the data only problem. However, we believe that this class of techniques can be extended to the data plus forecast problem and hope to do that in a subsequent paper.

In Section 2 we define the vector smoothing splines. In Section 3 we discuss the choice of Hilbert space norms. In Section 4 we describe numerical methods and the cross validation estimate of the smoothing parameter and in Section 5 we describe a Monte Carlo test of the method.

2. Helmholz Theorem and The Definition of Vector Smoothing Splines

We let P be a point on the sphere S, P = (λ,ϕ), where λ = longitude $(0\le\lambda\le 2\pi)$ and ϕ = latitude $(-\frac{\pi}{2}\le\phi\le\frac{\pi}{2})$. $\underset{\sim}{V} = (U,V)$ is a (sufficiently regular) horizontal vector field on the sphere, where U(P) is the eastward component and V(P) is the northward component at P.

The vorticity ζ and the divergence D of $\underset{\sim}{V}$ are given by

$$\zeta = \frac{1}{a\cos\phi}\left[-\frac{\partial}{\partial\phi}(U\cos\phi) + \frac{\partial V}{\partial\lambda}\right] \tag{2.1}$$

$$D = \frac{1}{a\cos\phi}\left[-\frac{\partial U}{\partial\lambda} + \frac{\partial}{\partial\phi}(V\cos\phi)\right], \tag{2.2}$$

where a is the radius of the sphere. Then there exists (by Helmoltz Theorem) two functions $\Psi(P)$ and $\Phi(P)$, $P\epsilon S$, called the stream function and the velocity potential respectively, with the following properties:

$$\left.\begin{aligned} U &= \frac{1}{a}\left(-\frac{\partial\Psi}{\partial\phi} + \frac{1}{\cos\phi}\frac{\partial\Phi}{\partial\lambda}\right) \\ V &= \frac{1}{a}\left(\frac{1}{\cos\phi}\frac{\partial\Psi}{\partial\lambda} + \frac{\partial\Phi}{\partial\phi}\right) \end{aligned}\right. \tag{2.3a}$$

$$\left.\begin{aligned} \zeta &= \Delta\Psi \\ D &= \Delta\Phi \end{aligned}\right. \tag{2.3b}$$

where Δ is the (horizontal) Laplacian on the sphere

$$\Delta f = \frac{1}{a^2}\left[\frac{1}{\cos^2\phi} f_{\lambda\lambda} + \frac{1}{\cos\phi}(\cos\phi f_\phi)_\phi\right].$$

Ψ and Φ are uniquely determined up to a constant (which we will take to be determined by $\int_S \Psi(P)dP = \int_S \Phi(P)dP = 0$. We are interested in defining Hilbert spaces of vector fields whose divergence and vorticity exists pointwise. We will do this as follows. Let $f(P)$ be a square integrable function on the sphere which integrates to 0. Then f has an expansion in the normalized spherical harmonics Y_ℓ^s

$$f(P) = \sum_{\ell=1}^{\infty} \sum_{s=-\ell}^{\ell} f_{\ell s} Y_\ell^s(P)$$

where

$$Y_\ell^s(\lambda,\phi) = \begin{cases} \theta_{\ell s} \cos s\lambda P_\ell^s(\sin\phi) & 0 \le s \le \ell \\ \theta_{\ell s} \sin s\lambda P_\ell^{|s|}(\sin\phi) & -\ell \le s < 0, \end{cases} \qquad \ell = 1,2,\ldots$$

$$\theta_{\ell s} = \sqrt{2}\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-|s|)!}{(\ell+|s|)!}} \qquad s \ne 0$$

$$= \sqrt{\frac{2\ell+1}{4\pi}} \qquad s = 0$$

and the Fourier Bessel coefficients $f_{\ell s}$ are given by

$$f_{\ell s} = \int f(P) Y_\ell^s(P)dP$$

with

$$\int f^2(P)dP = \sum_{\ell,s} f_{\ell s}^2.$$

Now Y_ℓ^s are the eigenfunctions of the Laplacian

$$\Delta Y_\ell^s = -\ell(\ell+1)Y_\ell^s .$$

Thus

$$\Delta f = -\sum_{\ell,s}\sum \ell(\ell+1) f_{\ell s} Y_\ell^s.$$

Let $\lambda_{\ell s}$, $\ell = 1,2,\ldots$, $s = -\ell,\ldots,\ell$ be a set of nonnegative numbers with

$\bar{\lambda}_\ell = \max_{s=-\ell,\ldots\ell} \lambda_{\ell s}$ and

$$\sum_{\ell} \ell^2(\ell+1)^2(2\ell+1)\bar{\lambda}_\ell < \infty. \tag{2.5}$$

Using the addition formula for spherical harmonics

$$\sum_{s=-\ell}^{\ell} Y_\ell^{s}(P) Y_\ell^{s}(Q) = \frac{2\ell+1}{4\pi} P_\ell(\cos\gamma(P,Q))$$

where γ is the angle between P and Q, the Cauchy-Schwartz inequality and the fact that $P_\ell(1) = 1$ gives

$$|\Delta f(P)| = \Big|\sum_{\ell,s} \ell(\ell+1)\sqrt{\lambda_{\ell s}}\, Y_\ell^{s}(P)\frac{f_{\ell s}}{\sqrt{\lambda_{\ell s}}}\Big|$$

$$\leq \Big(\sum_{\ell} \ell^2(\ell+1)^2 \sum_{s=-\ell}^{\ell} \lambda_{\ell s}(Y_\ell^{s}(P))^2\Big)^{1/2}\Big(\sum_{\ell,s}\frac{f_{\ell s}^{2}}{\lambda_{\ell s}}\Big)^{1/2}$$

$$\leq \Big(\frac{1}{4\pi}\sum_{\ell} \ell^2(\ell+1)^2(2\ell+1)\bar{\lambda}_\ell\Big)^{1/2}\Big(\sum_{\ell,s}\frac{f_{\ell s}^{2}}{\lambda_{\ell s}}\Big)^{1/2}$$

Thus $\{\lambda_{\ell s}\}$ satisfying (2.5) and $\sum_{\ell,s}\frac{f_{\ell s}^{2}}{\lambda_{\ell s}} < \infty$ imply that $\Delta f(P)$ is well defined and finite for all P.

Let H be the collection of all pairs (Ψ,Φ) on the sphere which integrate to zero, are square integrable and

$$J^{(1)}(\Psi) = \sum_{\ell=1}^{\infty}\sum_{s=-\ell}^{\ell}\frac{\Psi_{\ell s}^{2}}{\lambda_{\ell s}(1)} < \infty \qquad \Psi_{\ell s} = \int \Psi(P) Y_\ell^{s}(P)\,dP$$

$$J^{(2)}(\Phi) = \sum_{\ell=1}^{\infty}\sum_{s=-\ell}^{\ell}\frac{\Phi_{\ell s}^{2}}{\lambda_{\ell s}(2)} < \infty \qquad \Phi_{\ell s} = \int \Phi(P) Y_\ell^{s}(P)\,dP$$

where $\{\lambda_{\ell s}(1)\}$ and $\{\lambda_{\ell s}(2)\}$ are sequences satisfying

$$\sum_{\ell=1}^{\infty} \ell^2(\ell+1)^2(2\ell+1)\max_{s}\lambda_{\ell s}(i) < \infty,\ i = 1,2.$$

H is clearly a Hilbert space with square norm

$$||(\Psi,\Phi)||^2 = J^{(1)}(\Psi) + \frac{1}{\delta}J^{(2)}(\Phi)$$

for any fixed $\delta > 0$ and both members of each pair possess Laplacians everywhere. It is easy to show that if $\lambda_{\ell s} = [\ell(\ell+1)]^{-m}$, then

$$\begin{aligned} J(f) &= \int(\Delta^{m/2}f)^2dP && \text{m even} \\ &= \int\{\frac{(\Delta^{(m-1)/2}f)_\lambda^2}{\sin^2\phi} + (\Delta^{(m-1)/2}f)_\phi^2\}dP && \text{m odd.} \end{aligned} \tag{2.6}$$

If $\lambda_{\ell s} = [\ell(\ell+1)]^{-m}$, them $m > 3$ guarantees the pointwise existence of the Laplacian.

The observations are assumed to be of the form

$$U_i = U(P_i) + \epsilon_i^U,\ V_i = V(P_i) + \epsilon_i^V,\ i = 1,2,\ldots,n \tag{2.7}$$

where $(U(P_i),V(P_i))$ is the true (wind) vector at P_i and $\epsilon_i^U,\epsilon_i^V$ are measurement errors. We propose estimating the stream function and velocity potential (Ψ,Φ) associated with U and V by finding $(\Psi,\Phi)\epsilon H$ to minimize

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}(-\frac{1}{a}\frac{\partial\Psi}{\partial\phi}(P_i) + \frac{1}{a\cos\phi_i}\ \frac{\partial\Phi}{\partial\lambda}(P_i)-U_i)^2 \\ &+ \frac{1}{n}\sum_{i=1}^{n}(\frac{1}{a\cos\phi_i}\ \frac{\partial\Psi}{\partial\lambda}(P_i) + \frac{1}{a}\frac{\partial\Phi}{\partial\phi}(P_i)-V_i)^2 \\ &+ \lambda[J_1(\Psi) + \frac{1}{\delta}J_2(\Phi)] \end{aligned} \tag{2.8}$$

Note that in the residual sum of squares above, $U(P_i)$ and $V(P_i)$ are expressed in terms of Ψ and Φ via (2.3a). A unique minimizer $(\Psi_{\lambda,\delta},\Phi_{\lambda,\delta})$ exists for each $\lambda > 0$, $\delta > 0$ and the resulting wind field $(U_{\lambda,\delta},V_{\lambda,\delta})$ constructed from $(\Psi_{\lambda,\delta},\Phi_{\lambda,\delta})$ may be termed a vector spline field. Its vorticity and divergence will be given by $\zeta_{\lambda,\delta} = \Delta\Psi_{\lambda,\delta}$, $D_{\lambda,\delta} = \Delta\Psi_{\lambda,\delta}$. (Obviously, interpolating splines can be defined as minimizers of $J_1(\Psi) + \frac{1}{\delta}J_2(\Psi)$ subject to the interpolating conditions, we will not discuss these further.) Using WAHBA (1981a) or FREEDEN (1981a) it is straightforward to write an explicit (infinite series) expression for $(U_{\lambda,\delta},V_{\lambda,\delta})$.

3. On the Choice of $J^{(1)}$ and $J^{(2)}$

Let $\lambda_{\ell s} = |\sum_{j=0}^{m} \alpha_j[(-\ell)(\ell+1)]^j|^{-2}$ and suppose that $\lambda_{\ell s} > 0$ for $\ell = 1,2,\ldots$, $s = -\ell,\ldots,\ell$. It is not hard to see that

$$J(f) \doteq \sum_{\ell=1}^{\infty} \sum_{s=-\ell}^{\ell} \frac{f_{\ell s}^2}{\lambda_{\ell s}} = \int(\sum_{j=0}^{m} \alpha_j \Delta^j f)^2 dP \tag{3.1}$$

so that the choice of the $\lambda_{\ell s}$ can then be reduced to the choice of m and the $\{\alpha_j\}$. (If $\lambda_{\ell s} = 0$ for one or more s, the minimization problem can be handled by the methods described in KIMELDORF and WAHBA (1971), see also FREEDEN (1981a). In principle m and possibly α_{m-1} (with $\alpha_m=1$) can be chosen by cross validation (see WAHBA and WENDELBERGER (1980)), but it is undesireable to attempt to choose too many of these parameters from the data, see WAHBA (1981c).

In this section we will use the duality theorem which relates smoothing by splines to Bayesian estimation/Weiner filtering on stochastic processes to suggest how the J's may be chosen based on historical meteorological data.

To give the duality theorem we need some background, which we will give in a univariate context.

Let X(P), $P\varepsilon S$ be a (univariate) zero mean Gaussian stochastic process on the sphere with covariance R(P,Q) defined by

$$R(P,Q) = EX(P)X(Q),$$

where E is mathematical expectation. Following PARZEN (1961), CRAMER and LEADBETTER (1967) we can define the Hilbert space X spanned by X(P), $P\varepsilon S$, as all finite linear combinations of random variables (r.v.'s) of the form

$$Z_k = \sum_{j=1}^{n_k} \xi_{kj} X(P_{kj}) \tag{3.2}$$

and their quadratic mean (q.m.) limits. (A sequence $Z_1,Z_2,\ldots$, of r.v.'s has a q.m. limit if $\lim_{\ell,m\to\infty} E(Z_\ell-Z_m)^2 = 0$). The inner product in X is

$\langle X(P),X(Q)\rangle = EX(P)X(Q) = R(P,Q)$, and is extended by linearity to all r.v.'s of the form $Z_k = \sum_{j=1}^{n_k} \xi_{kj} X(P_{kj})$ and their q.m. limits. For example, letting L be a linear functional, the r.v. $LX = \frac{\partial}{\partial\phi}X(P_0)$ will be in X if the sequence of r.v.'s

$$Z_k = \frac{X(\lambda_0,\phi_0+h_k)-X(\lambda_0,\phi_0)}{h_k} \tag{3.3}$$

has a q.m. limit, as $h_k \to 0$, where $(\lambda_0,\phi_0) = P_0$. Then, it is not hard to show that the sequence $\{Z_k\}$ will have a q.m. limit $Z = \frac{\partial}{\partial\phi}X(P_0)$ if and only if

$$\frac{\partial^2}{\partial\phi\partial\phi'} R(P,P')_{P=P'=P_0} \tag{3.4}$$

is well defined and finite. Then the quantity in (3.4) is equal to $E(\frac{\partial}{\partial\phi}X(P_0))^2$, and furthermore

$$E(\frac{\partial}{\partial\phi}X(P_0))X(Q) = \frac{\partial}{\partial\phi_0} R(P_0,Q) \ .$$

More generally, let H_R be the reproducing kernel Hilbert space with reproducing kernel R. Then each random variable of the form $Z = LX$ can be identified with the bounded linear functional L on H_R, and vice versa. The argument is as follows. If $Z = LX$ is a r.v. in X it can be shown that $EZX(Q) = L_{(P)}R(P,Q) = \eta(Q)$, say, where $L_{(P)}$ means the linear functional L applied to R considered as a function of P. However, by the properties of reproducing kernels, it can be shown that $\eta(\cdot)$ is the representer of L in R, that is $Lf = \langle\eta,f\rangle_R$, where $\langle\cdot,\cdot\rangle_R$ is the inner product in H_R. We are now ready to state the

<u>Duality Theorem</u> (KIMELDORF and WAHBA (1970)).

Let $X(P)$, $P\epsilon S$ be a zero mean Gaussian stochastic process with covariance $bR(P,Q)$, and let H_R be the reproducing kernel Hilbert space with reproducing kernel H_R. Let

$$Y_i = L_i X + \varepsilon_i, \ i = 1,2,\ldots,n,$$

where $L_i X$, $i = 1,2,\ldots,n$ are n r.v.'s in X, and the $\varepsilon_1,\ldots,\varepsilon_n$ are independent, 0 mean Gaussian r.v.'s, independent of $X(P)$, $P \varepsilon S$, with common variance σ^2. Then the conditional expectation of $X(Q)$, given $Y_i = y_i$, $i = 1,2,\ldots,n$,

$$E\{X(Q)|Y_i = y_i,\ i = 1,2,\ldots,n\} \tag{3.5}$$

is given by $f_\lambda(Q)$, where $f_\lambda(\cdot)$ is the solution to the minimization problem: Find $f \varepsilon H_R$ to minimize

$$\frac{1}{n}\sum_{i=1}^{n}(L_i f - y_i)^2 + \lambda||f||_R^2,$$

and $\lambda = \sigma^2/nb$.

Proof: See KIMELDORF and WAHBA (1970,1971), WAHBA (1978). However, the proof proceeds by direct calculation of $f_\lambda(Q)$ and by using the facts that $E(L_i X)X(Q) = \eta_i(Q)$, where $\langle\eta_i, f\rangle_R = L_i f$.

Now let f be some atmospheric variable of interest. We will proceed as though the different realizations of f were sample functions from a zero mean Gaussian stochastic process with covariance $R(\cdot,\cdot)$. If repeated (independent!) observations on f were available, then various properties of R could be estimated from this data. We will discuss both "frequency domain" and "space domain" methods for doing this. Using the properties of reproducing kernel spaces (see, e.g. NASHED and WAHBA (1974)) it is not hard to show that if $J(f) = \sum_{\ell,s} \frac{f_{\ell s}^2}{\lambda_{\ell s}}$ is the norm on a reproducing kernel space H, then the reproducing kernel R for H is given by

$$R(P,Q) = \sum_{\ell,s} \lambda_{\ell s} Y_\ell^s(P) Y_\ell^s(Q). \tag{3.6}$$

To simplify the discussion, in this paper we are considering only R's whose eigenfunctions are the spherical harmonics. (Other eigenfunctions, i.e. those associated with Laplace's tidal equations, may well be reasonable in certain meteorological applications, sec WAHBA (1981b)).

If a stochastic process $X(P)$, $P\varepsilon S$, has covariance

$$\sum_{\ell,s} \lambda_{\ell s} Y_\ell^s(P) Y_\ell^s(Q)$$

then X may be modelled as a random linear combination of the spherical harmonics (Karhumen-Loeve expansion)

$$X(P) = \sum_{\ell s} X_{\ell s} Y_\ell^s(P) \tag{3.7}$$

where the $X_{\ell s}$ are random variables with

$$EX_{\ell s} X_{\ell' s'} = \lambda_{\ell s},\ \ell s = \ell' s',\ = 0,\ \ell s \neq \ell' s'.$$

(To see this, compute $EX(P)X(Q)$ from (3.7) to obtain (3.6).) We have

$$X_{\ell s} = \int X(P) Y_\ell^s(P)$$

and

$$\lambda_{\ell s} = EX_{\ell s}^2 = E(\int X(P) Y_\ell^s(P) dP)^2.$$

If K independent observations, $f^1,\dots,f^K$ of a meteorological variable of interest are available, this suggests choosing $\{\lambda_{\ell s}\}$ based on estimates

$$\hat{\lambda}_{\ell s} = \frac{1}{K} \sum_{k=1}^{K} (f_{\ell s}^k)^2$$

where the sample Fourier-Bessel coefficients $f_{\ell s}^k$, $k = 1,2,\dots,K$ are given by

$$f_{\ell s}^k = \int f^k(P) Y_\ell^s(P) dP.$$

Figure 1 gives a plot of February 1974 monthly averages of some atmospheric mean square sample Fourier Bessel Coefficients collected by STANFORD (1979) from Channels 2 and 4 of the Radiometer on NIMBUS-5. The radiation received by Channels 2 and 4 respectively can be used (crudely) to infer the temperature $T(P)$, $P\varepsilon S$ in the upper and lower stratosphere, respectively below the satellite. By piecing together data

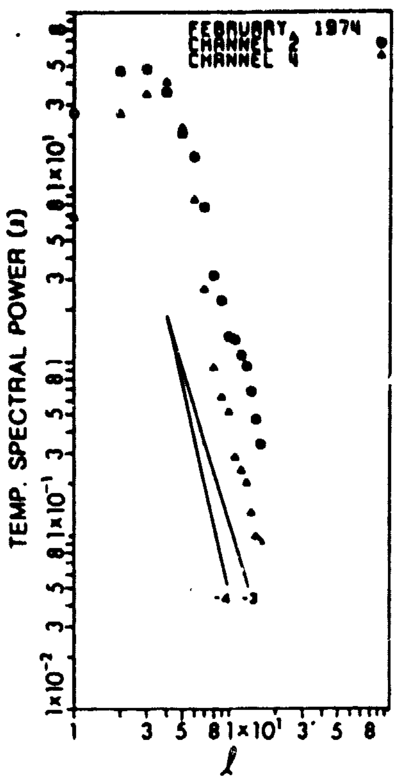

Figure 1: Temperature Spectral Power (ℓ).

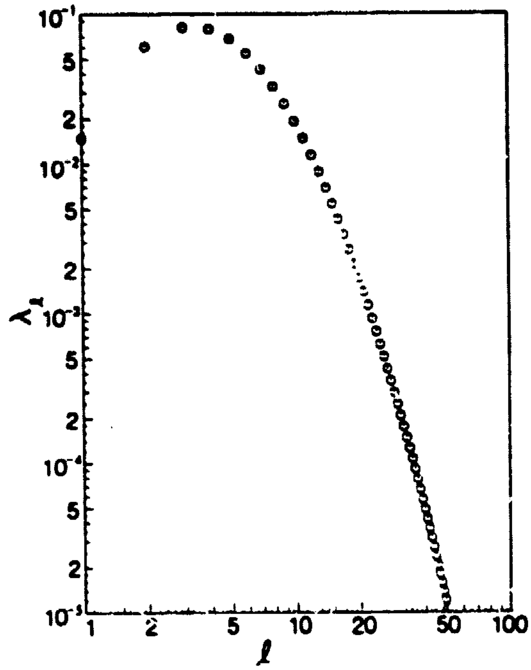

Figure 2: Idealized λ_ℓ.

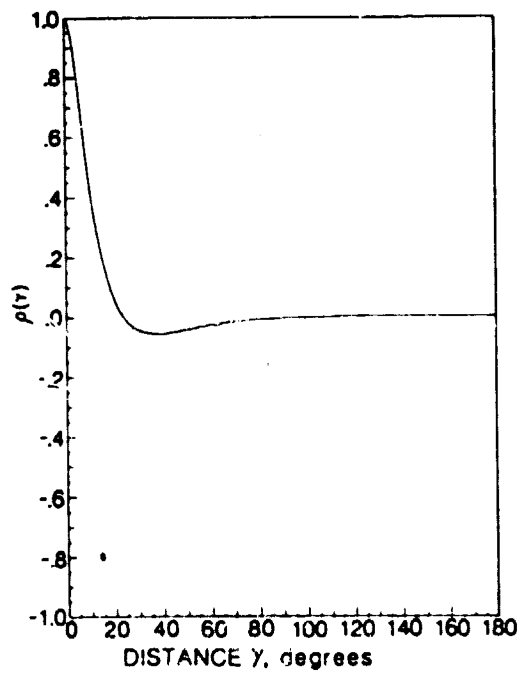

Figure 3: Correlation function for the (λ_ℓ) of fig. 2.

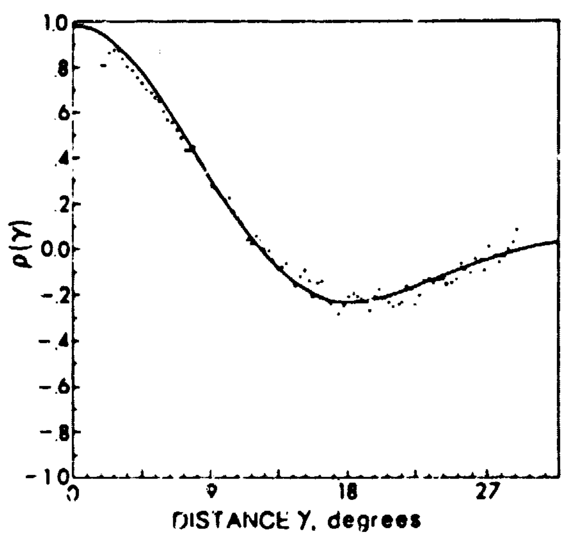

Figure 4: Sample Correlation Function.

from several orbits, (approximations to) $T_{\ell s}^{k} = \int T^{k}(P)Y_{\ell}^{s}(P)dP$ can be obtained. STANFORD has computed monthly mean square values $\bar{T}_{\ell s}^{2}$,

$$\bar{T}_{\ell s}^{2} = \frac{1}{K}\sum_{\ell=1}^{K}(T_{\ell s}^{k})^{2}.$$

What has actually been plotted in Figure 1 is the "TEMP SPECTRAL POWER" defined as

$$\text{TEMP SPECTRAL POWER }(\ell) = \frac{1}{5}\sum_{j=1}^{5}\bar{T}_{\ell+j,j}^{2}\ .$$

The energy spectrum in Temperature fields is related to the energy spectrum of other meteorological variables, i.e. wind and geopotential. We are not concerned here with the exact details of these pictures b.. rather that sequences $\{\lambda_{\ell s}\}$ can be fitted to this kind of data to provide meteorologically reasonable Hilbert Space norms. See KASAHARA (1976) for some plots of sample Fourier-Bessel coefficients with respect to the eigenfunctions of Laplace's Tidal equations for wind and geopotential. Figure 2 gives a plot of an idealized sequence $\lambda_{\ell s} = \lambda_{\ell}$, $\ell = 1,2,\ldots$, where λ_{ℓ} was obtained by fitting (by an ad hoc procedure), a function of the form

$$\lambda_{\ell s} = \left|\sum_{j=0}^{2}\alpha_{j}[-\ell(\ell+1)]\right|^{-2}$$

to some of the data behind Figure 1. If $\lambda_{\ell s}$ does not depend on s, $\lambda_{\ell s} = \lambda_{\ell}$, then the covariance

$$R(P,Q) = \sum_{\ell,s}\lambda_{\ell s}Y_{\ell}^{s}(P)Y_{\ell}^{s}(Q)$$

reduces by the addition formula for spherical harmonics, to

$$R(P,Q) = \frac{1}{4\pi}\sum_{\ell=1}^{\infty}(2\ell+1)\lambda_{\ell}\ P_{\ell}(\cos\gamma(P,Q)),$$

where $\gamma(P,Q)$ is the angle between P and Q. Figure 3 gives the function $\rho(\gamma)$ defined by

$$\rho(\gamma) = \sum_{\ell=1}^{\infty}(2\ell+1)\lambda_{\ell}P_{\ell}(\cos\gamma)/\sum_{\ell=1}^{\infty}(2\ell+1)\lambda_{\ell}P_{\ell}(\cos 0)$$

which is associated with the $\{\lambda_\ell\}$ of Figure 2. Figure 4 gives an estimate for $\rho(\gamma)$ for $f(P)$ = the 500 millibar (geopotential) height obtained by JULIAN and THIEBAUX (1975) from sample covariances from data from a network of 51 North American weather stations for the winters of 1966 and 1967. In estimating $\rho(\gamma)$, an isotropic covariance function was assumed. The purpose of providing Figures 1 and 4 here is to convince the reader that historical collected or collectable meteorological data may be used to choose the norm on H, although the particular data sets exhibited here may or may not be the most appropriate. In the numerical experiments to be described we have taken the $\{\lambda_{\ell s}(1)\}$ and $\{\lambda_{\ell s}(2)\}$ both as in Figure 2.

4. Numerical Methods. The Generalized Cross-Validation Estimates of λ and δ.

Given λ,δ, $\{\lambda_{\ell s}(1), \lambda_{\ell s}(2)\}$ and the data $\{(U_i,V_i)\}$, an approximate minimizer (Ψ,Φ) of (2.8) can be obtained in the form

$$\Psi = \sum_{\ell=1}^{N} \sum_{s=-\ell}^{\ell} \alpha_{\ell s} Y_\ell^s \tag{4.1}$$

$$\Phi = \sum_{\ell=1}^{N} \sum_{s=-\ell} \beta_{\ell s} Y_\ell^s \tag{4.2}$$

where N is sufficiently large. For other numerical approaches to the minimization of (2.8) see WAHBA (1980,1981a), WENDELBERGER (1982). Let $\tilde{N} = \sum_{\ell=1}^{N} \sum_{s=-\ell}^{\ell} 1 = N^2-1$ and renumber the indices (ℓ,s), $s = -\ell,\dots,\ell$, $\ell = 1,\dots,N$, as $1,2,\dots,\tilde{N}$. Let X_ϕ be the $n\times\tilde{N}$ matrix with $(i,\ell s)$th entry

$$\frac{1}{a}\frac{\partial}{\partial\phi}Y_\ell^s(P_i)$$

and X_λ be the $n\times\tilde{N}$ matrix with $(i,\ell s)$th entry

$$\frac{1}{a}\frac{1}{\cos\phi_i}\frac{\partial}{\partial\lambda}Y_\ell^s(P_i)$$

and let X be the $2n\times2\tilde{N}$ matrix

$$X = \begin{pmatrix} -X_\phi & X_\lambda \\ X_\lambda & X_\phi \end{pmatrix} \tag{4.3}$$

Let D_δ be the $2\tilde{N}\times 2\tilde{N}$ matrix

$$D_\delta = \begin{pmatrix} D_1 & 0 \\ 0 & \delta D_2 \end{pmatrix} \tag{4.4}$$

where D_i is the $\tilde{N}\times\tilde{N}$ diagonal matrix with $\ell s,\ell s$th entry $\lambda_{\ell s}(i)$, $i = 1,2$. Letting $z = (U_1,\ldots,U_n,V_1,\ldots,V_n)$, $\gamma = (\alpha_1,\ldots,\alpha_{\tilde{N}},\beta_1,\ldots,\beta_{\tilde{N}})$, it is seen by substituting (4.1) into (2.8) that we have to find γ which minimizes

$$\frac{1}{n}||z - X\gamma||^2 + \lambda\gamma' D_\delta^{-1}\gamma.$$

The minimizer is

$$\gamma = (X'X+n\lambda D_\delta^{-1})^{-1}X'z. \tag{4.5}$$

By the use of (2.3a) and (4.3), it follows that the estimated wind field $(U_{\lambda,\delta},V_{\lambda,\delta})$ at the data points satisfies

$$\begin{pmatrix} U_{\lambda,\delta}(P_1) \\ \vdots \\ U_{\lambda,\delta}(P_n) \\ V_{\lambda,\delta}(P_1) \\ \vdots \\ V_{\lambda,\delta}(P_n) \end{pmatrix} = A(\lambda)z \tag{4.6}$$

where $A(\lambda)$ is the $2n\times 2n$ "influence" matrix

$$A(\lambda) = X(X'X+n\lambda D_\delta^{-1})^{-1}X'$$

The generalized cross validation (GCV) estimate of (λ,δ) is the minimizer of the cross validation function $V(\lambda,\delta)$ defined by

$$V(\lambda,\delta) = \frac{\frac{1}{n}||(I-A(\lambda,\delta))z||^2}{\frac{1}{n}[\mathrm{Trace}(I-A(\lambda,\delta))]^2} \quad . \tag{4.7}$$

This method for estimating smoothing parameters in regularization problems was proposed in CRAVEN and WAHBA (1979), GOLUB, HEATH and WAHBA (1979) and WAHBA (1977b) and its numerical and theoretical properties have been studied in various places, see for example UTRERAS (1981). We only note here the useful property of the GCV estimate of λ and δ. Let the predictive mean square error $R(\lambda,\delta)$, when λ and δ are used be defined by

$$R(\lambda,\delta) = \frac{1}{n} \sum_{i=1}^{n} (U_{\lambda,\delta}(P_i)-U(P_i))^2$$

$$+ \frac{1}{n} \sum_{i=1}^{n} (V_{\lambda,\delta}(P_i)-V(P_i))^2, \tag{4.8}$$

where $U(P_i)$, $V(P_i)$ is the true (but unknown) wind vector, and suppose the measurement errors ε_i^U and ε_i^V are independent identically distributed zero mean normally distributed random variables. Then under rather general conditions, for large n the minimizer $(\hat{\lambda},\hat{\delta})$ of $V(\lambda,\delta)$ provides a good estimate of the minimizer of $R(\lambda,\delta)$. V is not guaranteed to have a unique, or even a finite minimizer. Practical difficulties in minimizing V though possible appear to be moderately rare when the assumptions are reasonably well satisfied. Various diagnostic tools are available in troublesome cases and will be discussed elsewhere.

The numerical experiment reported in Section 4 was performed on the Amdahl at Goddard Space Flight Center, with $2n = 228$, $N = 15$, $2\tilde{N} = 448$. We outline the calculations used. Let $W_\delta = XD^{1/2}$, and let the singular value decomposition (SVD) of W_δ be

$$W_\delta = UD_W V' \tag{4.9}$$

where $UU' = U'U = I_{2n\times 2n} = V'V$ and D_W is a diagonal matrix with entries $b_1,\ldots,b_{2n}$. $U,\{b_i\}$ and V' are computed using LINPACK. Letting

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_{2n} \end{pmatrix} = U'z,$$

then

$$V(\lambda,\delta) = \frac{\frac{1}{2n}\sum_{i=1}^{2n} z_i^2 - \frac{1}{2n}\sum_{\nu=1}^{2n} w_\nu^2\left(\frac{b_\nu^2}{b_\nu^2+n\lambda}\right)^2}{\left(1-\frac{1}{2n}\sum_{\nu=1}^{2n}\frac{b_\nu^2}{b_\nu^2+n\lambda}\right)^2} \tag{4.10}$$

$$\gamma = D_\delta^{1/2} V \begin{pmatrix} \frac{b_1}{b_1^2+n\lambda} & & 0 \\ & \ddots & \\ 0 & & \frac{b_{2n}}{b_{2n}^2+n\lambda} \end{pmatrix} w. \tag{4.11}$$

For fixed $\delta, \hat{\lambda}(\delta)$, the minimizer of (4.10), is easily found by a global search in increments of $\log\lambda$. Then $V(\hat{\lambda}(\delta),\delta)$ was plotted for 8 values of δ chosen in powers of 1/6, and the minimum was readily evident. No doubt more efficient and automatic search procedures can be found.

For large n, N, and W_δ poorly conditioned, computing the SVD can be expensive, or it can fail to converge in a reasonable time. Some shortcut methods which alleviate this problem somewhat and use less storage have been developed. (BATES and WAHBA, (1982) in preparation.)

5. A realistic Monte Carlo test of the method

A number of techniques for estimating divergence of the upper atmosphere from radiosonde data have been proposed in the atmospheric sciences literature. For example, see SCHMIDT and JOHNSON (1972). In an attempt to determine how well the proposed method might work in practice a Monte Carlo experiment simulating realistic measured wind data from "model" stream functions and velocity potentials has been coded, and various experiments run. We describe one such experiment.

We obtained a model streamfunction and velocity potential of the form

$$\Psi = C_1 \sum_{\ell=1}^{N} \sum_{s=-\ell}^{\ell} a_{\ell s} Y_\ell^s$$

$$\Phi = C_2 \sum_{\ell=1} \sum_{s=-\ell}^{\ell} b_{\ell s} Y_\ell^s \tag{5.1}$$

by choosing $a_{\ell s}$ and $b_{\ell s}$ as normally distributed pseudo-random numbers with mean 0 and variances $\lambda_{\ell s}(1) = \lambda_{\ell s}(2) = \lambda_{\ell s}$ given in Figure 2. C_1 and C_2 were scale factors chosen so that the simulated $\zeta = \Delta\Psi$ and $D = \Delta\Phi$ had magnitudes typical of real atmospheres.
$(\int\zeta^2 dP)^{1/2} = 6 \times 10^{-5}$/sec., $(\int D^2 dP)^{1/2} = 1 \times 10^{-5}$/sec. Model wind vectors $(U(P_i), V(P_i))$ were computed from the model (Ψ, Φ) of (5.1) for $\{P_i\}$ corresponding to $n = 114$ North American radiosonde stations. The data $z = (U_1,\ldots,U_n, V_1,\ldots,V_n)$, where $U_i = U(P_i) + \varepsilon_i^U$, $V_i = V(P_i) + \varepsilon_i^V$, were constructed by adding the measurement errors ε_i^U, ε_i^V as normally distributed pseudo random numbers with mean 0 and standard deviation $\sigma = 2.5$ meters/sec., a realistic value for the measurement error standard deviation. Since the ability to estimate divergence will depend on the signal to noise ratio, it is necessary that the values of "signal" and "noise" be chosen realistically. The results reported here can be expected to be rosier than that obtainable in practice, however, primarily to the extent that wave numbers $\ell > N$ occur in practice but are not simulated here, and (secondarily) because in practice $J^{(1)}$ and $J^{(2)}$ cannot be so precisely matched to the "truth" as they are in this experiment.

Figure 5 shows the simulated wind vectors. Figure 6 shows the estimate of the true wind field, plotted on a 5° × 5° grid in latitude and longitude. Figures 7 and 8 show the model and estimated vorticity and divergence, respectively. Figure 9 shows $V(\hat{\lambda}(\delta),\delta)$ and $R(\hat{\lambda}(\delta),\delta)$, (of 4.8) plotted as a function of δ. In Figures 6 - 8 $\hat{\delta} = 1/36$ was used. It can be seen that the minimizer of $V(\hat{\lambda}(\delta),\delta)$ was a good estimate of the minimizer of $R(\hat{\lambda}(\delta),\delta)$. Figure 10 gives $MSE(\zeta_{\hat{\lambda}(\delta),\delta})$ and $MSE(D_{\hat{\lambda}(\delta),\delta})$ and their sum, where

$$MSE(\zeta_{\lambda,\delta}) = \frac{1}{K}\sum_{k=1}^{K}(\zeta_{\lambda,\delta}(P_k)-\zeta(P_k))^2$$

$$MSE(D_{\lambda,\delta}) = \frac{1}{K}\sum_{k=1}^{K}(D_{\lambda,\delta}(P_k)-D(P_k))^2.$$

The $\{P_k\}$ constitute a regular grid inside the U.S. It can be seen from Figure 10 that if δ is taken as too small (i.e. divergence is suppressed), then the mean square error in the estimated vorticity becomes large.

ORIGINAL PAGE IS
OF POOR QUALITY

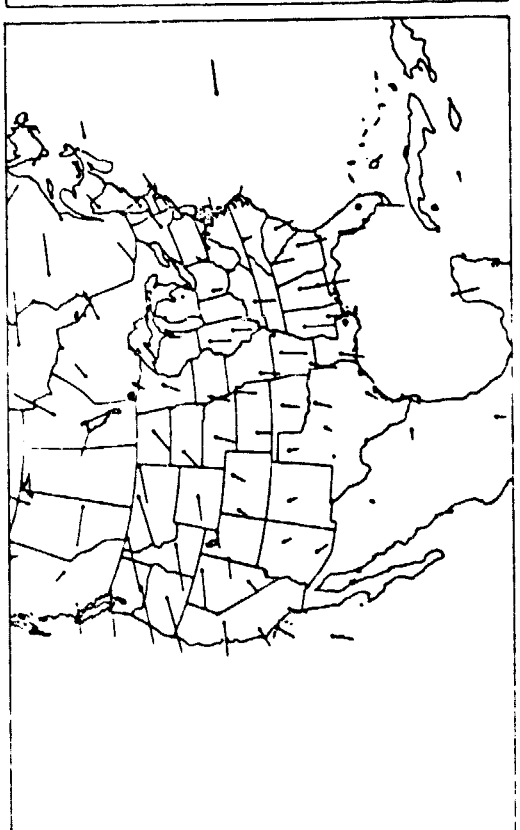

Figure 5: Simulated Wind Data.

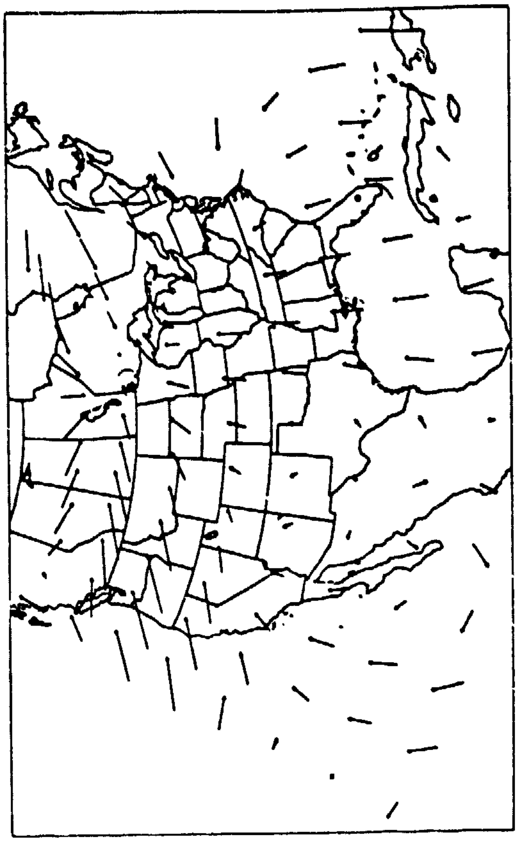

Figure 6: Estimated Wind Field.

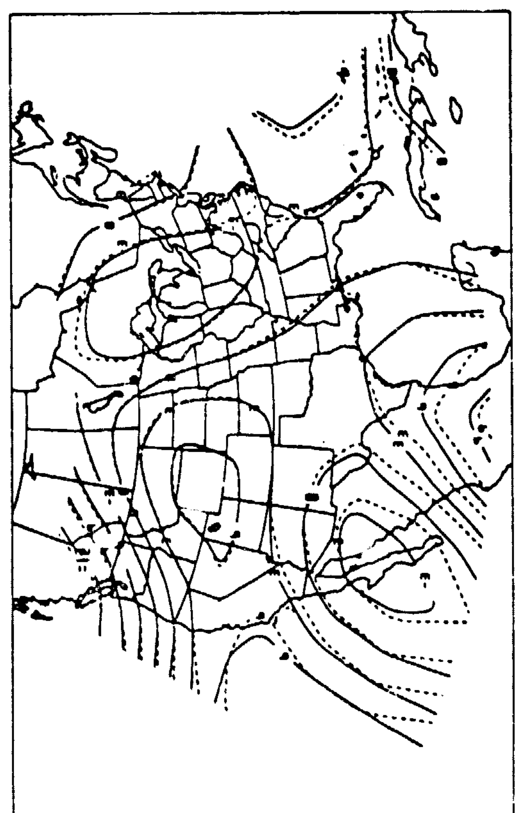

Figure 7: Model and Estimated Vorticity, X 10^{-5}/sec.

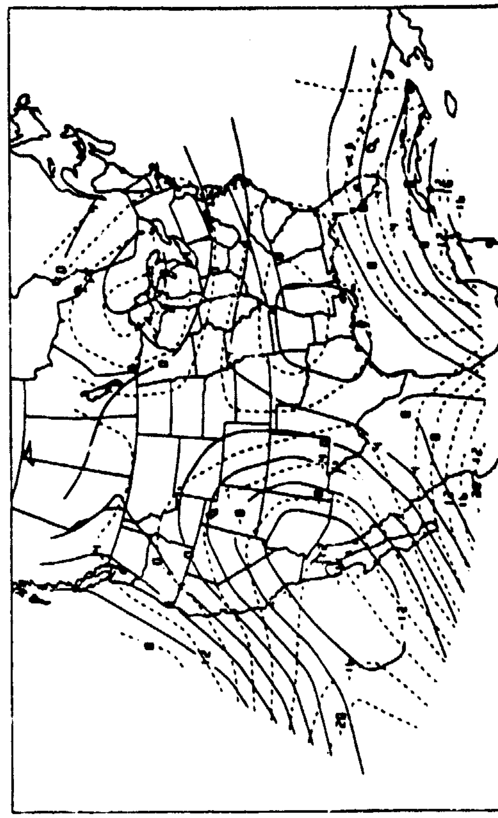

Figure 8: Model and Estimated Divergence, X 10^{-6}/sec.

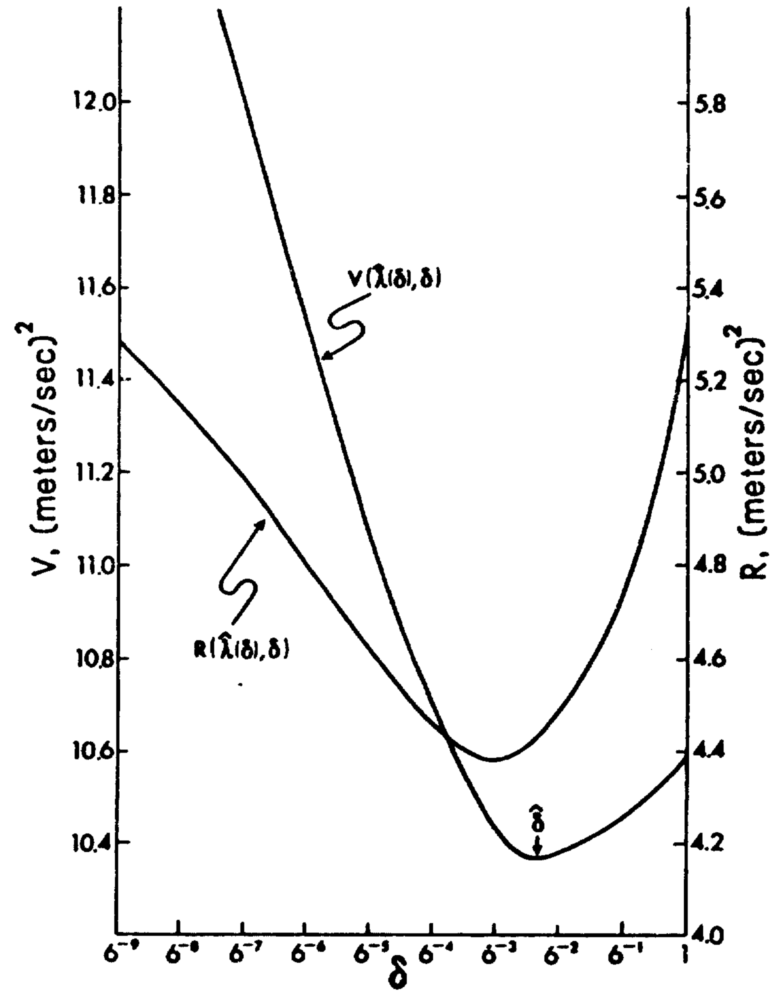

Figure 9: The cross validation function V and the predictive mean square error R of the wind field.

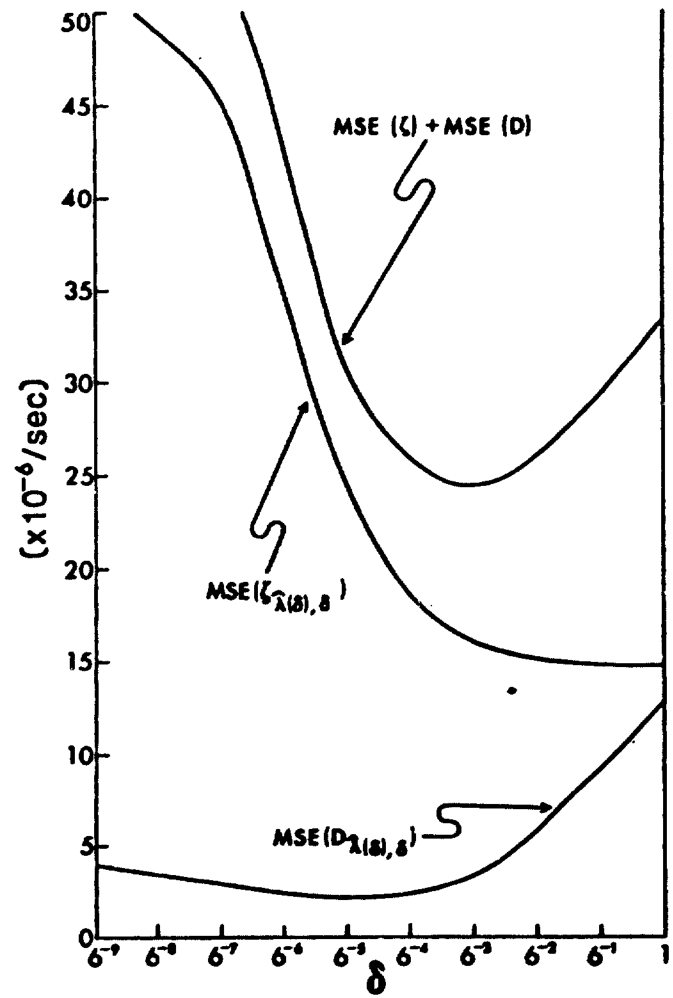

Figure 10: Mean square error in the estimated vorticity and divergence.

An estimate $\hat{\sigma}^2$ for the variance of the measurement error is available as

$$\hat{\sigma}^2(\hat{\lambda},\hat{\delta}) = \frac{||(I-A(\hat{\lambda},\hat{\delta})||^2}{Tr(I-A(\hat{\lambda},\hat{\delta}))}$$

since the numerator is the residual sum of squares and the denominator is the equivalent degrees of freedom for error. In this example $\hat{\sigma}$ was 2.58m/sec., very close to the "true" value of 2.5 meters/sec. In those occasional sticky cases encountered in practice where $V(\lambda,\delta)$ has multiple minima, if the order of magnitude of σ is known apriori, the examination of $\hat{\sigma}$ can usually be used to resolve ambiguity. See WAHBA (1981d), WENDELBERGER (1982). Bayesian confidence intervals are also available for these estimates, see Wahba (1981d).

We have concluded that this approach has much promise for applications.

6. References

Bates, D., and Wahba, G. (1982) in preparation.

Cramer, H., and Leadbetter (1967). Stationary and related stochastic processes, Chapter V. Wiley, New York.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. Numer. Math., 31, 377.

Freeden, W. (1981a). On spherical spline interpolation and approximation, Math. Meth. in The Appl. Sci. 3, 551-575.

Freeden, W. (1981b). On approximation by harmonic splines. Manuscripta Geodaetica, 6, 193-244.

Golub, G., Heath, M. and Wahba, G. (1979), Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics 21, 215-223.

Julian, P.R., and Thiebaux, M. Jean (1975), On some properties of correlation functions used in optimum interpolation schemes, Monthly Weather Review, 103, 7, pp. 605-616.

Kasahara, A. (1976). Normal modes of ultra-long waves in the atmosphere. Monthly Weather Review 104, 6, 669-690.

Kimeldorf, G., and Wahba, G. (1970), A correspondence between Bayesian estimation of stochastic processes and smoothing by splines, Ann. Math. Statist., 41, 2.

Kimeldorf, G., and Wahba, G. (1971), Some results on Tchebycheffian spline functions, J. Math. Anal. and Applic., 33, 1.

Micchelli, C., and Wahba, G. (1981), Design problems for optimal surface interpolation in "Approximation Theory and Applications: Z. Ziegler, ed., Academic Press.

Nashed, M.Z. and Wahba, G. (1974), Generalized inverses in reproducing kernel spaces: an approach to regularization of linear operator equations, SIAM J. Math. Analysis, 5, 6.

Parzen, E. (1961), An approach to time series analysis. Ann. Math. Statist. 32, 951-989.

Schmidt, P.J., and Johnson, D.R. (1972), Use of approximating polynomials to estimate profiles of wind, divergence, and vertical motion. Monthly Weather Review, 100, 5, 249-353.

Shure, L., Parker, R.L., and Backus, G.E. (1981), Harmonic splines for geomagnetic modelling, to appear, PEPI.

Stanford, J. (1979), Latitudinal-Wavenumber power spectra of stratospheric temperature fluctuations, J. Atmospheric Sciences, 36, 5, pp. 921-931.

Utreras, F. (1981), Optimal smoothing of noisy data using spline functions, SIAM J. Sci. Stat. Comput. 2, 3, 349-362.

Wahba, G. (1977a) in invited discussion to Consistent nonparametric regression, C.J. Stone, Ann. Stat., 5, 4, 637-645.

Wahba, G. (1977b), Practical approximate solutions to linear operator equations when the data are noisy, SIAM J. Numerical Analysis, 14, 4.

Wahba, G. (1978), Improper priors, spline smoothing and the problem of guarding against model errors in regression, J. Roy. Stat. Soc. Ser. B., 40, 3.

Wahba, G. (1980), Spline bases, regularization, and generalized cross validation for solving approximation problems with large quantities of noisy data. Proceedings of the International Conference on Approximation Theory in Honor or George Lorenz, Jan. 8-11, 1980, Austin, Texas, Ward Cheney, ed. Academic Press (1980).

Wahba, G. (1981a), Spline interpolation and smoothing on the sphere, SIAM J. Scientific and Statistical Computing, 2, 1.

Wahba, G. (1981b), Some new techniques for variational objective analysis on the sphere using splines, Hough functions, and sample spectral data. Preprints of the Seventh Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society.

Wahba, G. (1981c), Data-based optimal smoothing of orthogonal series density estimates. Ann. Statist., 9, 1, 146-156.

Wahba, G. (1981d), Bayesian confidence intervals for the cross validated smoothing spline. University of Wisconsin-Madison Statistics Department Technical Report No. 645, Submitted.

Wahba, G., and Wendelberger, J. (1980), Some new mathematical methods for variational objective analysis using splines and cross-validation, Monthly Weather Review 108, 36-57.

Wendelberger, J. (1982), Ph.D. thesis, in preparation.

We thank Christopher Sheridan, who wrote the computer program, and Donald R. Johnson for many helpful discussions.

Professor Grace Wahba, Department of Statistics, University of Wisconsin-Madison Madison, Wisconsin, 53706, USA.