~~SUBMITTED~~ GRANT

IN-47-CR
126016
6P.

DEPARTMENT OF STATISTICS

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

(NASA-CR-182531) ON THE DYNAMIC ESTIMATION OF RELATIVE WEIGHTS FOR OBSERVATION AND FORECAST IN NUMERICAL WEATHER PREDICTION (Wisconsin Univ.) 6 p CSCL 04B N88-18106 Unclas G3/47 0126016

TECHNICAL REPORT NO. 818

January, 1988

On the dynamic estimation of relative weights for observation and forecast in numerical weather prediction.

by

Grace Wahba

To appear in the Proceedings of the
International Workshop on
Remote Sensing Retrieval Methods,
Held at Williamsburg, VA Dec. 15-18, 1987
A. Deepak, Ed.

This research supported in part by NASA under Contract NAG5-315
and NSF under Grant ATM-840373.

ON THE DYNAMIC ESTIMATION OF RELATIVE WEIGHTS FOR OBSERVATIONS AND FORECAST IN NUMERICAL WEATHER PREDICTION

Grace Wahba
Department of Statistics
University of Wisconsin-Madison
1210 W. Dayton St.
Madison, Wisconsin 53706, USA

ABSTRACT

We look at the problem of merging direct and remotely sensed (indirect) data with forecast data to get an estimate of the present state of the atmosphere, for the purpose of numerical weather prediction. To carry out this merging optimally, it is necessary to provide an estimate of the relative weights to be given to the observations and forecast. It is possible to do this dynamically from the information to be merged, if the correlation structure of the errors from the various sources is sufficiently different. We describe some new statistical approaches to doing this and quantify conditions under which such estimates are likely to be good.

1. INTRODUCTION

We have been studying various aspects of the problem of simultaneously combining information from various sources, for the purpose of obtaining initial conditions of the atmosphere for numerical weather prediction. By information, we mean data from diverse instruments, information from a forecast, prior information concerning the atmosphere, and physical constraints. See Wahba(1981,1982a, 1985a). Various parts of this what might be called "multispectral" point of view, whereby data from different sources is combined, and several meteorological parameters retrieved simultaneously, is a major theme in several papers presented at this conference, in particular, see Isaacs et al. (1988), Smith et al.(1988), Westwater et al. (1988) and Rodgers (1988). See also Lorenc(1986).

In this paper we will first briefly review the variational prescription for combining data from different sources and prior information, and its relation to Gandin (Bayes) estimation. In this prescription are a number of "tuning" parameters, which we will divide into two classes. The first class will be called weighting parameters, and the second smoothing (also known as bandwidth) parameters. The weighting parameters are those which govern the relative weights to be given to various types of observational, forecast, and physical information, while the smoothing parameters control the relative amount of "information" which is to be assigned to "signal" and to "noise". All practical forecast models have many such tuning parameters, and in practice, they are chosen by trial and error, by the use of externally measured data on various sources of error and strength of signal, and in very simple cases, by Kalman filtering, which propagates estimates of covariances forward in the model.

it may model the relationship between directly measured quantities such as the horizontal wind field, and state variables in the forecast model such as the coefficients in a spherical harmonic expansion of stream function and velocity potential.

We will suppose that a reasonable model for the state variables (the mean having been subtracted off), is

$$Es = 0, \quad Ess' = b\Sigma$$

See Wahba(1982b), for further references and a discussion as to how Σ may be obtained from historical data.

This approach results in a mandate to find s as the minimizer of the variational problem:

$$(y^{(1)} - F^{(1)}(s))Q_1^{-1}(y^{(1)} - F^{(1)}(s)) + \frac{1}{r}(y^{(2)} - F^{(2)}(s))Q_2^{-1}(y^{(2)} - F^{(2)}(s))$$

$$+ \lambda s'\Sigma^{-1}s ,$$

where $r = \frac{\sigma_2^2}{\sigma_1^2}$ and $\lambda = \frac{\sigma_1^2}{b}$. See Wahba (1981, 1982a, 1985a), O'Sullivan and Wahba(1985). We note that physical constraints on the state vector s or functions of s can be inserted by solving the variational problem above subject to these constraints, see Villalobos and Wahba(1982), Svensson(1985). Svensson put constraints on the dry adiabatic lapse rate when solving a variational problem of this form with satellite radiance data to estimate vertical temperature profiles.

In the linear Normally distributed case the estimate of s which minimizes the above variational problem is the Gandin(Bayes) estimate of s, given the prior covariance of s and the covariance matrices of the errors. See, e. g. Kimeldorf and Wahba(1970). Kalman filter theory would give s as the minimizer of the variational problem with the term preceeded by λ absent. Here the inclusion of this term enters apriori (smoothness) information that may have been lost via model error.

If r, Q_1 and Q_2 are known, then λ and some parameters inside Σ may be estimated by the GCV See, for example, Wahba and Wendelberger(1980), Wahba(1982b), O'Sullivan and Wahba(1985), Merz(1980).

3. DYNAMIC ESTIMATION OF WEIGHTING PARAMETERS

3.1 A SIMPLE EXAMPLE - 500 mb RAOBS AND FORECAST

We first illustrate the method and results by letting $y^{(1)}$ be observed 500mb heights and $y^{(2)}$ be forecast 500mb heights. It is reasonable to take Q_1 to be I for the observations since these measurements can be assumed to be independent, with about the same variance from station to station. Hollingsworth and Lonnberg(1986) and Lonnberg and Hollingsworth(1986) (LH), have recently obtained estimates of r and $Q_f = Q_2$ from three months data from the European Center forecast model. We will use their example and results as an illustration, before going on to the general case.

Figure 1 shows an estimated 500 mb correlation function from LH. Let h_l^o and h_l^f be the observation and the forecast 500 mb height at station l, and ϵ_l^o and ϵ_l^f be the observation and forecast errors, (on a particular day). Let

$$\xi_l = h_l^o - h_l^f = e_l^o - e_l^f$$

Here, $y^{(1)} = h^o$ and $y^{(2)} = h^f$, thus both vectors represent the same meteorological quantity, that condition will be relaxed later. Letting τ_{lm} be the distance between stations l and m, LH assumed that

$$E\epsilon_l^f \epsilon_m^f = \sigma_f^2 \rho(\tau_{lm})$$

where $\rho(0) = 1$. If the e_l^o are independent and identically distributed zero mean random variables then

$$E\xi_l \xi_m = \sigma_1^2 \delta_{lm} + \sigma_f^2 \rho(\tau_{lm})$$

where $\delta_{lm} = 1$ if $l = m$ and 0 otherwise. LH collected 90 days of values of ξ_l for each station. Letting j index day, they used sample correlations computed from

$$\frac{1}{90} \sum_{j=1}^{90} \xi_l(j) \xi_m(j)$$

as an estimate of $\sigma_1^2 \delta_{lm} + \sigma_f^2 \rho(\tau_{lm})$. In the figure, sample values of $\dfrac{\sigma_f^2 \rho(\tau)}{\sigma_1^2 + \sigma_f^2 \rho(0)}$ are plotted, and $\dfrac{\sigma_f^2}{\sigma_f^2 + \sigma_1^2}$ is estimated by extrapolating the smooth part of the curve back to the origin by methods described in their paper (quite different than the methods to be discussed here.)

Figure 2 shows a one parameter family of (synthetic) correlation functions, defined by

$$\rho_L(\tau) = \frac{(1 - 2\theta(L)\cos(\frac{2\pi\tau}{R_O}) + \theta^2(L))^{-1/2} - (1 + \theta(L))^{-1}}{(1 - \theta(L))^{-1} - (1 + \theta(L))^{-1}}$$

where $\theta(L)$ is determined by

$$\frac{3}{2}\theta(L) - \frac{1}{2}\theta^3(L) = \cos\frac{2\pi L}{R_O}.$$

Here the parameter L in km is the distance τ for which $\rho_L(\tau) = \frac{1}{2}\rho_L(0) = \frac{1}{2}$, where R_O is the circumference of the earth, in km. This family of (isotropic) correlation functions on the sphere has been chosen here partly because of a superficial resemblance to some of the curves obtained by LH and partly for mathematical convenience. We wish to use this family as a moderately realistic example of the estimation of a single (important) parameter in Q_f, namely, the correlation half-distance. Further study is needed to determine if it is appropriate to include other factors, such as anisotropy, variation with latitude, etc. in this correlation function. With this model, letting $\xi = (\xi_1, \ldots, \xi_n)$ be a vector of one day's data, we have that the covariance matrix of ξ is $\sigma_1^2(I + rQ_f(L))$, where the lmth entry of $Q_f(L)$ is $\rho_L(\tau_{lm})$. The GCV estimate of r and L can be shown to be the values of r and L which minimize

We now consider the general (linear) case, where

$$y^{(i)} = F^{(i)}s + \varepsilon^{(i)}, \quad i=1,2$$

where $y^{(i)}$ is of dimension $n^{(i)}, i=1,2$ and the covariance of $\varepsilon^{(i)}$ is $\sigma_i^2 Q_i$. To estimate r, we must be able to construct two matrices $B^{(1)}$ and $B^{(2)}$ of dimension $n \times n^{(1)}$ and $n \times n^{(2)}$ respectively, which satisfy $B^{(1)}F^{(1)} = B^{(2)}F^{(2)}$. Let $z^{(i)}$ be defined by

$$z^{(i)} = B^{(i)}y^{(i)}, \quad i=1,2$$

and let w be defined by

$$w = z^{(1)} - z^{(2)} = B^{(1)}y^{(1)} - B^{(2)}y^{(2)} = B^{(1)}\varepsilon^{(1)} - B^{(2)}\varepsilon^{(2)}.$$

The covariance matrix of w is then

$$Ew'w = \sigma_1^2 B^{(1)}Q_1B^{(1)\prime} + \sigma_2^2 B^{(2)}Q_2B^{(2)\prime}.$$

Suppose $B^{(1)}Q_1B^{(1)\prime}$ is of full rank, then we can take the Cholesky decomposition LL' of $B^{(1)}Q_1B^{(1)}$, where L is lower triangular, and let $\xi = L^{-1}w$. Then the covariance matrix of ξ is

$$E\xi\xi' = \sigma_1^2(I + rQ),$$

where $Q = L^{-1}B^{(2)}Q_2B^{(2)}L^{-1\prime}$. The ML and GCV estimates are then given by the minimizers of V and M of Section 3.1, and the estimability of r depends on the properties of Q. Loosely speaking, the two vectors $z^{(1)}$ and $z^{(2)}$ need have to have their "energy" at different "wavenumbers".

3.4 COMPUTATIONAL CONSIDERATIONS

Computation of r and L minimizing quantities similar to M and V has been carried out in a relatively straightforward way using matrix decompositions for n up to a few hundred, on a VAX 11/720 in the Statistics Dept. at the University of Wisconsin- Madison, and with n up to about 800 on the Cray XMP at the San Diego Super Computer Center, without much optimization of the code, in under 150 seconds, using GCVPACK (Bates et al. (1985)). Research is continuing on more efficient methods for larger problems. It is probably true that large data sets will be required in practice.

3.5. POTENTIAL APPLICATIONS

A possible important application is to the comparison of satellite observed radiances to forecast. It is possible to observe certain gross features of the atmosphere in two dimensional plots of satellite (raw) radiances. Forecast errors frequently tend to be phase errors, with certain spatial features displaced in space. To compare forecast $y^{(2)}$ with satellite radiance data $y^{(1)}$ following the approach in this paper, one should compute from the forecast, radiances that would be seen by the satellite, that is, let $B^{(1)}$ be I and $B^{(2)}$ map forecast into radiance data as would be seen by the satellite. Assuming that realistic Q_1 and Q_2 can be established (remember, Q_2 is in the radiance observational domain), it is likely that r could be estimated, and used to help decide wether to trust the forecast or the radiance data in the event of a major discrepancy.

linear remote sensing experiments, *J. Comput. Physics,* vol. 59, pp. 441-455.

Parzen, E. (1963), Probability density functionals and reproducing kernel Hilbert spaces, in Proceedings of the Symposium on Time Series Analysis Held at Brown, M. Rosenblatt, ed, pp. 155-169, Wiley, N. Y.

Rodgers, C. (1988), A general error analysis for profile retrieval, in International Workskhop on Remote Sensing Retrieval Methods, ed. A. Deepak, A. Deepak Publishing, Hampton, VA.

Seaman, R. S. and M. F. Hutchinson (1985), Comparative real data tests of some objective analysis methods by withholding observations, *Aust. Met. Mag.,* vol. 33, pp. 37-46.

Smith, W. L., H. E. Revercomb, H. B. Howell, H. M. Woolf, and A. Huang (1988), Science applications of high resolution interferometer spectrometer measurements from the NASA U2 aircraft, in International Workshop on Remote Sensing Retrieval Methods, ed. A. Deepak, A. Deepak Publishing, Hampton, VA.

Svensson, J. (1985), A nonlinear inversion method for derivation of temperature profiles from TOVS data, Tech. Proc. Second International TOVS Study Conference, pp. 292-307, CIMMS, Space Scence and Engineering Center, Madison, WI, 1985.

Villalobos, M. and G. Wahba (1987), Inequality constrained multivariate smoothing splines with application to the estimation of posterior probabilities, *J. Am. Statist. Assoc.,* vol. 82, pp. 239-248.

Wahba, G. and J. Wendelberger (1980), Some new mathematical methods for variational objective analysis using splines and cross validation, *Monthly Weather Review,* vol. 108, pp. 1122- 1145.

Wahba, G. (1981), 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. (1982a), Variational methods in simultaneous optimum interpolation and initialization, in The Interaction Between Objective Analysis and Investigation, ed. D. Williamson, pp. 178-185, Atmospheric Analysis and Prediction Division, Na- tional Center for Atmospheric Research, Boulder, CO. Publication in Meteorology 127

Wahba, G. (1982b), Vector splines on the sphere, with application to the estimation of vorticity and divergence from discrete, noisy data, in Multivariate Approximation Theory, Vol. 2, ed. W. Schempp and K. Zeller, pp. 407-429, Birkhauser Verlag.

Wahba, G. (1985a), Variational methods for multidimensional inverse problems, in Advances in Remote Sensing Retrieval Methods, ed. M. Chahine, pp. 385-410, A. Deepak Publishing,