Estimation of Errors in Objectively Analyzed Fields and Sensitivity to Number and Spacing of Stations

M.J. Leach, J.J. Yio, and R.T. Cederwall
Atmospheric Science Division
Lawrence Livermore National Laboratory
Livermore, California

Motivation

Single-column models (SCMs) are designed to test parameterizations for radiative fluxes, cloud processes, and surface exchange that are used in general circulation models (GCMs). The SCM is a vertical column of GCM grid cells exercised in isolation from the GCM. The lateral information normally supplied to the column by neighboring columns in a GCM must be supplied externally by estimates of the temperature and moisture advection and the horizontal wind divergence to force the physical processes within the isolated column. In the Atmospheric Radiation Measurement (ARM) Program, these required forcing terms are derived from observations using objective analysis. We seek to quantify the errors in the objectively analyzed fields and to investigate the sensitivity of these errors to the number and spacing of observation stations.

Methodology

Recent investigators (Zamora et al. 1987; Davies-Jones 1993; Michael 1994) have used triangulation techniques and line integrals or Cramer's Rule to estimate fields and derivatives with observations from a few stations. Our work builds on that of Bretherton et al. (1976) in using optimal analysis to estimate the error in both the field and derivatives. Errors in both the field estimates and the derivative estimates is a function of signal wavelength relative to the station spacing and also the number of observations used in creating the estimates.

The Gauss-Markov theorem states that the optimum linear estimator for $\theta(x,y)$ is

$$\theta_x = \sum_{r=1}^{N} C_{xr}\left(\sum_{s=1}^{N} A_{rs}^{-1}\phi_s\right) \qquad (1)$$

The error variance is

$$(\theta - \theta_x)^2 = C_{xx} - \sum_{r,s=1}^{N} C_{xr}C_{xs}A_{rs}^{-1} \qquad (2)$$

where C_{xr} is the covariance between the estimated quantity and the *r*th measurement, and A_{rs} is the covariance between pairs of measurements. In this study, we assume covariance of the form

$$F(x) = \exp(-x^2 / R_c^2) + e_v \exp(-x^2 / R_n^2) \qquad (3)$$

The first term in (3) represents signal and the second term represents noise.

The existing observation network at the Southern Great Plains (SGP) site used in ARM for the objective analysis is illustrated in Figure 1. The dots in Figure 1 represent the locations of the SGP Central Facility and Boundary Facilities where radiosondes are launched. The triangles represent locations of the National Weather Service (NWS) 404-MHZ wind profilers, which are used to enhance the ARM observations in the estimation of derived fields. The siting scenarios considered in this study, as illustrated in Figure 2, approximate the SGP network. The main difference is that we used evenly spaced 'optimal' siting, given the number of stations. A few siting scenarios that had more stations than we expect to have in the SGP network were included.

We averaged the error internal to regularly spaced stations interior to the domain. That is, we added the error calculated by (2) at all internal grid points, then divided by the number of grid points. We calculated this average error for various smoothing length scales, R_c (or equivalently, resolved wavelengths), noise length scales R_n and noise to signal ratio e_v.

Results

A number of signal and noise scales and noise amplitudes were investigated. The amplitude of the signal was always considered as unity. The noise amplitude e_v is expressed as a fraction of the signal. In our non-dimensional space, the signal scale R_c and the noise scale R_n are represented as a fraction of the size of the entire grid. A composite of the results for the various scenarios is presented in Figure 3. In

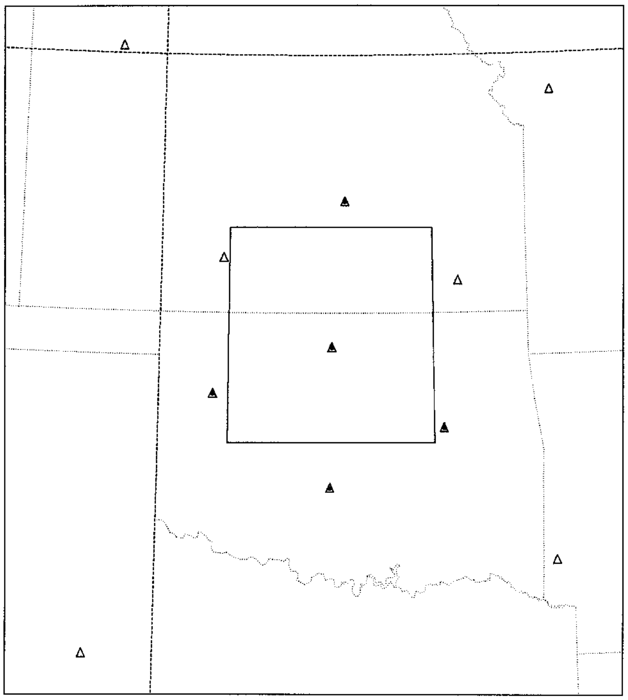

Figure 1. Location of ARM radiosonde launch points (dots) and National Weather Service 404-MHZ wind profilers (triangles) at the Southern Great Plains site.

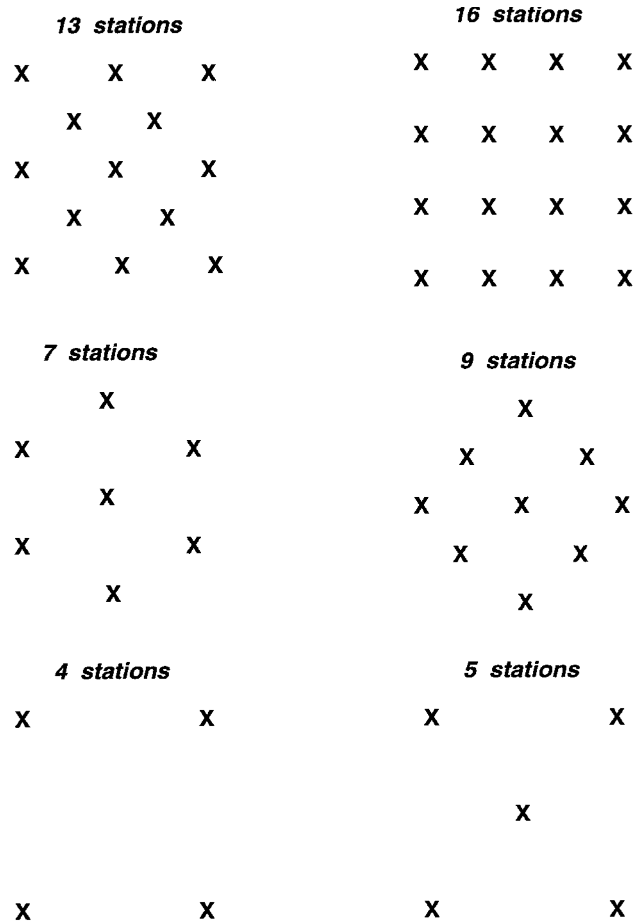

Figure 2. Siting scenarios considered in this study.

each case, the ordinate is the noise amplitude relative to the signal, and the abscissa is the scale length of the signal. In these cases, the scale length of the noise is constant at R_n=0.1. The contours are the average error for all points in the interior grid. As expected, the error decreases as the number of sampling stations increases. In fact, when the number of stations is increased by a factor of four (from 4 to 16), the error is decreased by a factor of two. This is consistent with fundamental sampling theory; decreasing station spacing by a factor of two in each direction resolves wavelengths that are smaller by a factor of two.

In Figure 4, a different perspective of the same issue is illustrated. In this figure, the ordinate is the number of stations; the error amplitude is held constant at 0.1. The contours are again the average error. In all cases, a monotonic decrease in the error results as stations are added, effectively sampling shorter wavelengths of the signal. The consistency with sampling theory is further illustrated. The error decreases by a factor of two for a four-fold increase in the number of stations. Because we are examining two-dimensional fields, the four-fold increase in station number is equivalent to a doubling of the station density.

Summary

Using the techniques of Optimal Interpolation, we have estimated the errors introduced in representing continuous fields with discrete measurements. The error depends on the number of stations, the station spacing, and the wavelength of the atmospheric forcing relative to the half width of the Gaussian filter used in the interpolation. The results suggest that there is no critical number of stations necessary below which one must abandon traditional objective analysis techniques based on weighting functions in favor of the non-filtering techniques of line integrals or Cramer's Rule.

We are continuing efforts to assimilate data from ARM SGP instruments and working to incorporate more data from the NWS Wind Profiler Demonstration Network, as well as the NWS synoptic radiosonde network. In addition, we are continuing to evaluate the adequacy of the available data for calculating means and gradients required to drive SCMs.

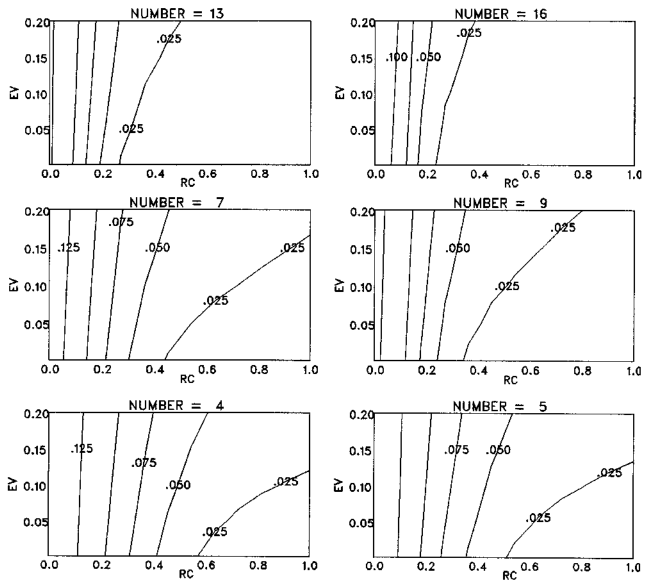

Figure 3. Sensitivity of error variance to noise level (e_v) and signal scale (R_c), for different station siting scenarios.

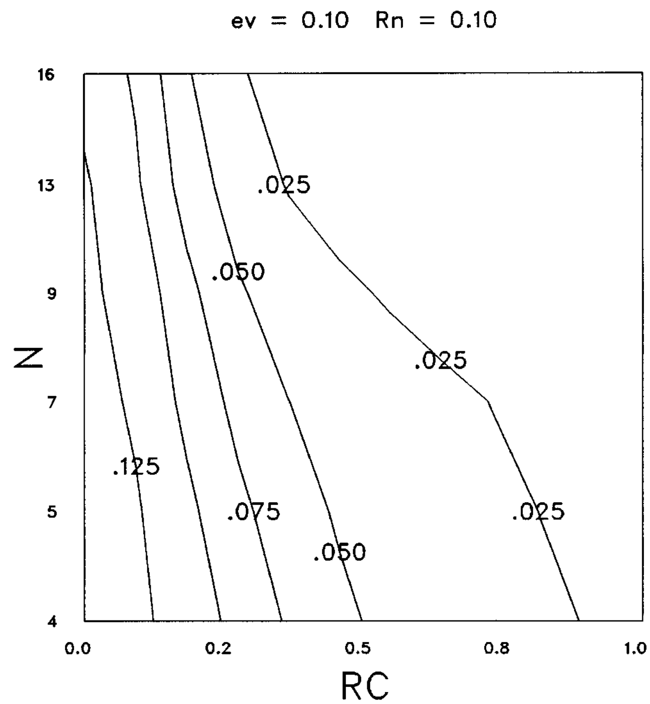

Figure 4. Sensitivity of error variance to the number of stations (N) and signal (R_c), for a given noise characterized by relative magnitude (e_v) and scale (R_n), for different station siting scenarios.

References

Bretherton, F.P., R.E. Davis, and C.B. Fandry, 1976: A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Research*, **23**, 559-582.

Davies-Jones, R., 1993: Useful formulas for computing divergence, vorticity, and their errors from three or more stations. *Mon. Wea. Rev*., **121**, 713-725.

Michael, P., 1994: Estimating advective tendencies from field measurements. *Mon. Wea. Rev*., **122**, 2202-2209.

Zamora, R.J., M.A. Shapiro, and C.A. Doswell III, 1987: The diagnosis of upper-tropospheric divergence and ageostrophic wind using profiler wind observations. *Mon. Wea. Rev*., **115**, 871-884.