8.1

IMPACT OF DATA RESOLUTION ON METEOROLOGICAL FIELDS OVER THE POLAR REGIONS

Serguei Ivanov and Jeffrey S Tilley
Geophysical Institute, University of Alaska Fairbanks, Fairbanks AK 99775-7320

1. Introduction

Improvements to numerical weather prediction (NWP) models are actively pursued in the arenas of improved data inputs (through data assimilation), model physics (through field campaigns and process studies) and improved numerics. Additionally, it is often assumed that increasing the model resolution will also lead to improved analyses and forecasts as a matter of course, especially for severe weather events.

Consequently, it is particularly interesting to examine how a combination of present observations with a model solution can be combined to utilise various informational sources. The variational problem associated with the estimation of state vector, $\mathbf{x}^t$, from the background vector, $\mathbf{x}^b$, and observation vector, $\mathbf{y}^o$, corresponds to the minimisation of the objective function (Talagrand 1997, T97 in the following)

$$J(\mathbf{x}) \equiv [\mathbf{x}-\mathbf{x}^b]^T(\mathbf{P}^b)^{-1}[\mathbf{x}-\mathbf{x}^b]+[\mathbf{H}\mathbf{x}-\mathbf{y}^o]^T\mathbf{R}^{-1}[\mathbf{H}\mathbf{x}-\mathbf{y}^o] \quad (1.1)$$

where $\mathbf{x}$ is the n-dimensional vector to be found. The objective function is the sum of two terms, one measuring the distance to the background, $\mathbf{x}^b$, the other measuring the distance to the observation vector, $\mathbf{y}^o$. These two terms are weighted by the inverse covariance matrices of the corresponding errors, $\mathbf{P}^b$ and $\mathbf{R}$.

If the relevant observations are taken at points in space-time different from the points at which estimates are sought, the observation operator, $\mathbf{H}$, will represent some space-time interpolation. If the observations are "indirect" functions of the parameters to be estimated and do not directly represent all physical quantities, then the observation operator, $\mathbf{H}$, will represent an appropriate linearisation of the physical and statistical relationship linking state vector, $\mathbf{x}$, and observation vector, $\mathbf{y}^o$. As an example, satellite-derived radiometers measure the radiative flux emitted by the Earth to space, while what is desired are estimates of the atmospheric temperature and humidity fields. The measured fluxes are functions of these fields as well as of other quantities, such as cloud amount, cloud top pressure, and surface emissivity, through the radiative transfer equations or an equivalent operator.

Such an approach in a data assimilation context requires knowledge of correct error statistics of both a background field and the observations to be used in obtaining the optimal state estimation. For an ideal case, weighting coefficients should be prescribed in an inverse ratio to the standard deviation of the error of the corresponding data source. However, these error statistics are still far from perfect.

While instrumental observation error or the accuracy of an equivalent transfer operator function for indirect measurements can often be reasonably evaluated, the representativeness error of the observations (Lorenc 1986) as well as the model errors are still poorly known. The representativeness error, which is thought to introduce a spatial correlation into the observation error, is difficult and expensive to estimate and/or specify. The model errors are a sum total of our inadequate description of physical processes within the governing equations plus the transformation of these equations to a finite difference or spectral representation. Thus, the analyzed initial conditions as well as the ensuing forecast are very dependent on the quality of the information sources. An additional uncertainty arises when the data is distributed heterogeneously in time or space. This is the case of traditional observation networks covering most of the globe, including the polar regions. High resolution satellite data should also be treated as heterogeneous in the vertical direction as well as the horizontal directions for polar orbiting satellites.

2. Experimental framework

2.1. General Considerations

Most assimilation algorithms that have been used operationally or in research, can be described as more or less simplified forms of a least-squares statistical linear estimation. This is a classical tool, whose basic principles are straightforward, even though practical implementation on large dimension systems can raise many problems.

Sequential data assimilation schemes, such as optimal interpolation, nudging, adaptive filtering, or the Bayesian approach disturb the physical consistency of the solutions through so-called "updating" of the model states when correcting it by the observations.

Variational assimilation approaches, on the other hand,

aim at globally adjusting a model solution to all the observations available over the assimilation period. As the adjustment is simultaneous, the adjusted states at all times are influenced by all the observations over the assimilation period, thereby avoiding the difficulty mentioned for sequential methods (T97). Another feature of variational assimilation is a requirement that the accuracy (or error) of each informational source be known. As was mentioned in the Introduction, the errors and their statistics are not well understood.

Often, some assumptions concerning the accuracy of both the model and the observations are used to compensate for such a gap in knowledge. Such assumptions are mostly made in the context of homogeneous data distribution over the domain to be studied. Within the past few years several authors (Schyberg 2002, Tanguay 2002, Majumdar 2001, Dee 2001, Molteni 2001, Marsigli 2001, Desroziers 2001a) have paid more attention to the problem of the impact from how the data is distributed on variational assimilation systems. A heterogeneously distributed observing network would be expected to have a significant and irregular impact on the observation error distribution, if one only assumes that the representativeness error is related to the corresponding scale to be captured by the network. Such an assumption has a strong basis due to the fact that atmospheric processes possess continuous spectra with amplitude descending towards the smaller scales. As such, the representativeness error should be proportional to the amplitude at each scale, and the observation error is therefore a function of the scales given by the distances between observation locations as such scales fluctuate over the heterogeneous net.

2.2. Test in data assimilation scheme

To test the potential impact of observation data resolution on a set of analysis fields, an experiment in which observations are generated by a model itself has been designed. This is a general approach (Desroziers 2001b, Rabier 2002) which, given knowledge of the true answer, allows for a clear and simple evaluation of the performance of the developed procedure. In this experiment a domain over Alaska, shown in Figure 1, was considered. The intent of the approach is to simulate observations corresponding to a given set of observation errors produced by using the operational error for variances.

Two simulations were completed with the MM5 4Dvar data assimilation model (e.g., Zou 1998). In the first simulation, the observations are given by a simulated set of measurements of the same variables as in the

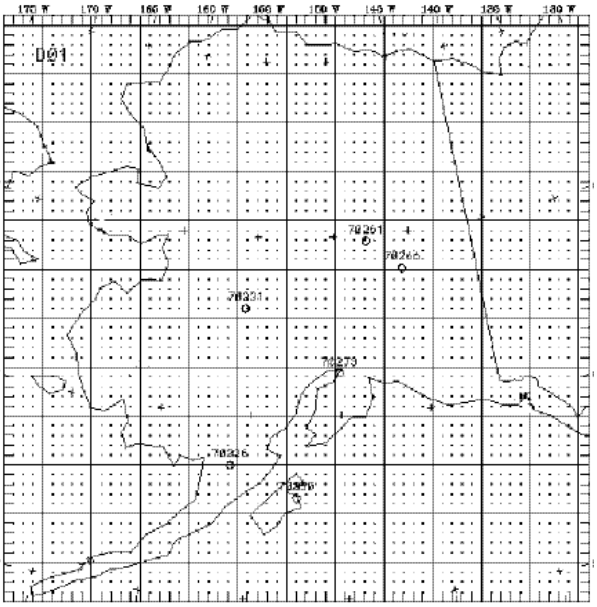

Figure 1. Study domain, with model grid (dots) overlain. 5-digit numbers indicate observation locations used in the experimental simulations.

model, for the model grid point locations. The observation errors are produced by using a random noise generator applied to the operational error variance for the atmospheric variables. The second simulation was performed in the same manner, except the number of the observational locations has been restricted to six and placed at the model grid points nearest to the real observations stations with the numbers 231, 261, 266, 273, 326, 350 (see Figure 1). This procedure of shifting the observation locations onto the gridpoints allows us to avoid the use of an observation operator, which when used can be non-linear and result in additional error.

2.3 Field representation in physical and spectral space

.

The fact that an arbitrary mathematical function can be represented by its trigonometric series was established long ago by J. Fourier, in 1807. It is well known that some restrictions must be placed on an arbitrary function in order for it to be expanded in a Fourier decomposition. This restriction is reflected in the Riesz-Fischer theorem, which states that if the spectrum of the function has a finite energy, then the Fourier coefficients provide a one-to-one mapping between the Fourier expansion and the continuous field. This mapping preserves the energy, and the associated Fourier series converges to the function in the sense that the mean square error of the energy tends to zero as the number of the coefficients tends to infinity.

Thus, an infinite number of Fourier coefficients is required term for a correct transform as well as a complete description of a field in spectral space. But in nearly all atmospheric applications, discrete fields over a finite time period are of concern. This gives rise to the necessity to take into account possible misrepresentations of the original fields by numerical operations. Apparently, these distortions take place witin both physical and the spectral spaces, linked by the appropriate relations. However, to classify and describe such variations is simpler and more convenient in spectral space. Until now, the representativeness error, has been treated within the context of an estimation of its possible limit value. However, the complexity of the derived sources of data and the sensitivity of data assimilation to the form of any operator relating the atmospheric state to measured quantities can be seen through the evaluation of the representativeness error in spectral space. This approach also gives an opportunity to describe a potential impact from using nonuniform carriers for digitising a continuous signal. These carrier forms posses rather complicated responses even if the number of sampling points is not very much.

The primary applicable rules of the Fourier transform and the description of the typical errors that arise due to digitising the continuous signal will be discussed below.

2.4. Errors due to finite field representations.

Referring to Findlay (1978), the complex convolution theorem for the spectrum of a field is given by:

$$X_L(f) = 1/2\pi \int X(f)*A(f-\theta)d\ddot{y} \qquad (2.3)$$

where $X_L(f)$, is the convolution of the original spectrum, X(f), with a spectral width, A(f). (2.3) states that the discretized spectrum, $X_L(f)$, is an altered version of the original spectrum, X(f). The alteration is described in part by the spectral width A(f), which in turn is determined by the spatial domain of interest, a(**x**), in physical space. As a(**x**) increases, we see that the departure of the discretized spectrum from the original spectrum decreases, and asymptotically approaches the delta-function when the domain becomes infinitely large.

Another aspect that should be noted is often termed "leakage", and refers to the effect of interaction among physical processes that operate over similar but not equal scales. Such “leakage” leads to an incorrect estimation of the physical fields. From a numerical point of view, this is an outgrowth of the fact that a spectral representation of a field can only recognize components consistent with the maximum resolvable scale of a domain. So, a larger domain ensures a more precise extraction of the appropriate scales.

The types and potential sources of the errors in the spectral space can be transferred into the physical space. Formally, they are linked by the appropriate properties in the physical space by the Fourier transform.

The choice of domain size is often governed by the available computing resources or the forecast problem of interes. It is possible and probably necessary to evaluate the value and distribution of the potential error raised due to an insufficient amount of information in the scales covered within a given domain. This error grows from the improper definition of the contributions from different scales, and can be classified as a representativeness error.

2.5. Aliasing and the error of discretising.

A numerical field representaiton of an atmospheric quantity, defined stepwise in time and space by steps Δt (or Δx) and over a limited domain size N can be linked to the spectral space by a Fourier series even though the field cannot be considered periodic. The corresponding Fourier coefficients exist only for integer indeces and contain in their core a fundamental frequency $f=1/N$ (or $\omega=2\pi/N$). That is, the field can be partially represented as being periodic in space with period N. Thus, we may only consider the Fourier coefficients for one period, that is, for indices $k=0,1,2,...,N-1$. In doing so, we establish a consistency between the finite domain representation (physical space) and the corresponding finite Fourier coefficient domain representation (spectral space). Moreover, we need only consider the harmonic frequencies $f_k=1/N$ for $k=0,1,2,...,N/2$, since harmonic frequencies for $k>N/2$ represent redundant information. Thus, all the necessary information provided by a discrete Fourier transform is contained in the interval $0<k<N/2$.

However, for the case where $f_N<2F$, where F is the highest value of the frequency from the band limited spectrum, $X_L(f)$ becomes distorted due to the overlapping harmonic components. In fact, the upper frequencies in X(f) are reflected into the lower frequencies in $X_L(f)$ and the high frequency component in X(f) cannot be resolved. This effect is known as aliasing.
Consequently, proper sampling requires that the samples be taken at a frequency of at least $f_N=2F$. This upper limit for the resolved frequency, known as the Nyquist frequency, is determined numerically as

$f_N=1/2\Delta t$ (or Δx). Here, Δt and Δx represent the time step and grid size, respectively.

In the practice of numerical computation, it is necessary to sample at a rate much higher than that given by the theoretical minimum. But as a rule, certain factors restrict or eliminate possibilities for the choice of the sampling frequency. Also, all physical signals found in the real world contain components covering a full frequency range with amplitudes diminishing towards the highest band. All of these factors make the exact reproduction of a continuous signal from the sampled signal impossible. On the other hand, *a priori* evaluation of the error of such reproduction via knowledge about the sampling is possible and highly recommended.

2.6. The spatial space windows for the test experiment.

As mentioned earlier, we consider, for the MM5 4DVAR model, six observation points in the central area of the domain Figure 1. For the sake of simplicity, their coordinates were shifted to the nearest model grid points. This operation allows us to specify the simulated observations and model variables at the same location and to avoid using an observational operator that could confound the studied effects.

Figure 2 shows the two-dimensional (in a horizontal plane) spectral space window corresponding to the spatial distribution of the six observing sites. Here the distance and spatial frequency (inverse distance value) are given in terms of the model Δx and Δy. The corresponding Nyquist domain has a rectangular shape, which regionally determines the area for valid analysis. Only the lower frequencies should be considered for the analysis, because the rest contain wrong information that disturb the original signal (of the field of interest). This distortion primarily happens on the smallest scales corresponding to two to three model grid steps.

However, close inspection of Figure 2 shows that there is a still smaller area within this main Nyquist domain which contains the scales undisturbed through incompatibilities of model grid and observation network resolutions. Near the main peak, one can see two side lobes with a spectral power that reaches 80% of the main one. These lobes correspond to the scales of order 400-600 km (12-20 model grid units) and, consequently, should mostly affect structures of such size. Thus, the spectral transform method applied to the continuous atmospheric fields ensures valid estimations only for the large scale processes which correspond to the area of smaller frequencies. At the same time, the false peaks which arise at the intermediate frequency

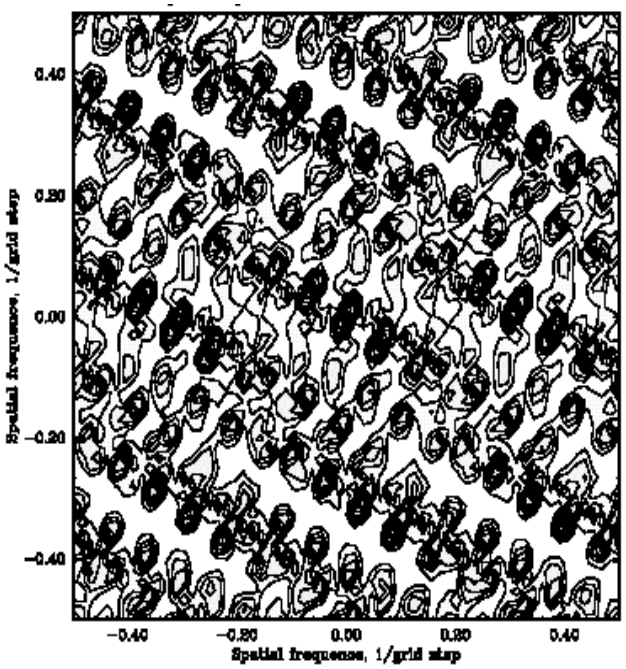

Figure 2. The two-dimensional spectral space window corresponding to the present spatial distribution of six observation locations indicated on Figure 1.

band (i.e. the mesoscale) have a numerical origin and distort the corresponding mesoscales.

3. Numerical results

The model and data assimilation runs have simulated a series of atmospheric fields on the various model vertical levels. The analysis of their statistics as well as a description of the particular atmospheric pattern are the subject for later investigation. Here we will only perform an analysis of the features directly related to the impact coming from the irregular data distribution on the analysis fields. Figure 3 shows the first guess temperature field for a mid-tropospheric sigma level (Figure 3a), along with the temperature field obtained as a set of optimal initial conditions when both the full simulated observational set (for each model grid point) (Figure 3b) and the "degraded" set (with only six observational locations) (Figure 3c) were used in the data assimilation scheme.

Figure 4 shows the differences between the first guess and the corresponding optimal initial conditions obtained with the use of the two data sets. Even a cursory glance establishes that these differences are significantly dissimilar. When the full set of observations is utilized, the difference is mainly concentrated in the area of the high gradients in the original temperature field. There are also a number of

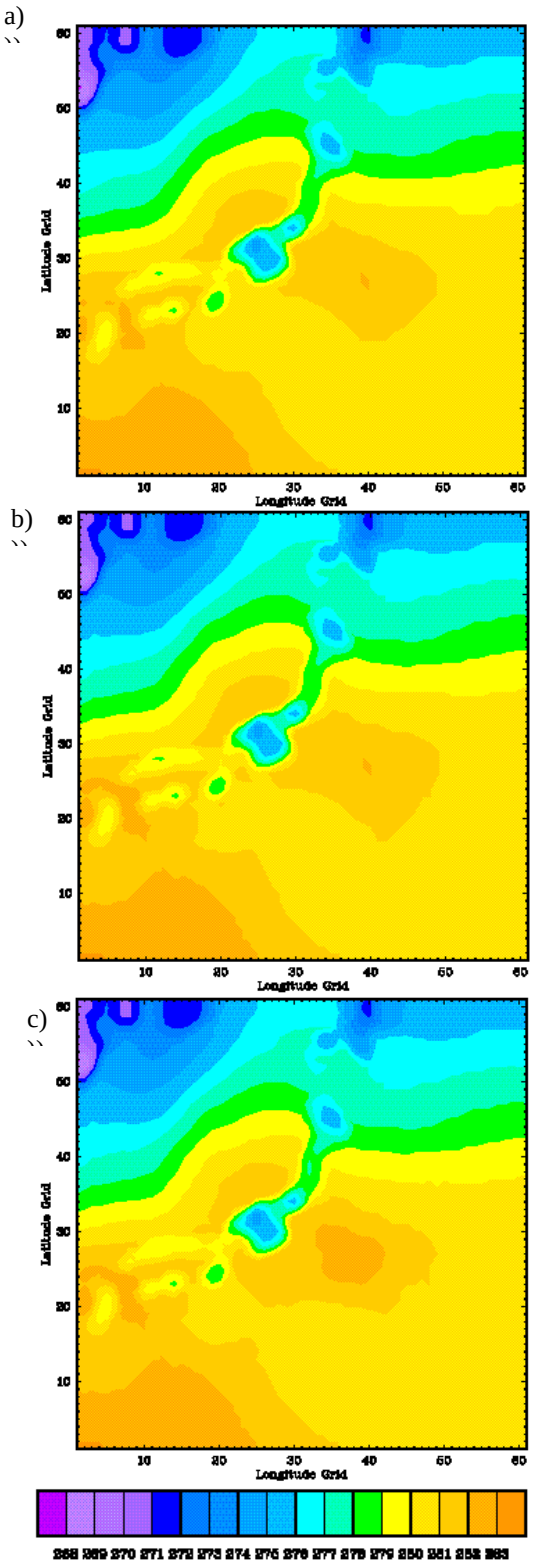

Figure 3. The temperature fields for a mid-tropospheric sigma level corresponding to (a) first guess, (b) optimal initial conditions obtained with using full simulated observational set in each model grid point , and (c) optimal initial conditions obtained with using "degraded" six-point observation set.

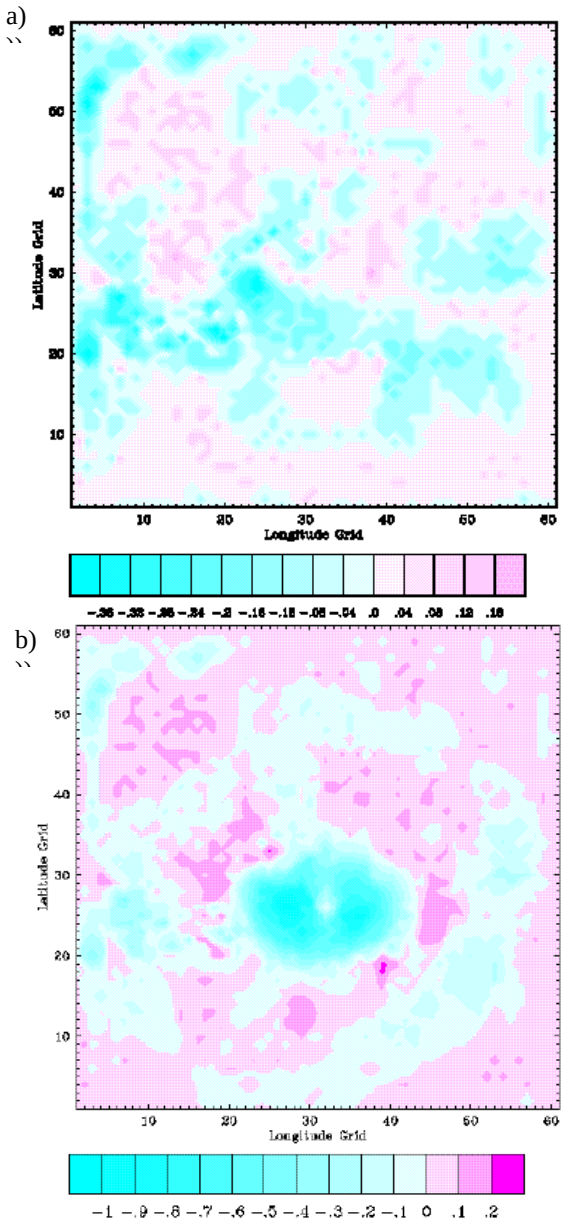

Figure 4. The differences in temperature fields between first guess and initial conditions obtained: (a)with using the full simulated observational set in each model grid point and (b) with optimal initial conditions obtained via using the "degraded" six-point observation set.

structures with various scales and minor amplitudes spread over the whole domain. The other difference field contains a considerably larger deviation in the obtained initial conditions, predominantly around the area where the observations have been provided in the assimilation scheme. Additionally, there is a quasi-circle or elliptical structure propagating out from the area containing data in the form of a dispersive wave trail. Finally, the difference field shows the presence of structures with a scale of order 300-600 km (10-20 model grid steps). A description of such structures in

the physical space is not immediately apparent. This is in contrast to the spectral space, where complete information contained in the physical space can be found concentrated in the corresponding frequencies.

Figure 5 compares the spectra for two cross-sections passing through the approximate core of the data area. It clearly shows that these spectra are very far from each other; primarily, the differences are related to the spatial frequencies that roughly correspond to the false side lobes seen in Figure 2. This confirms that the various informational sources are not being utilized within the data assmilation scheme appropriately weighted for their spatio-temporal distribution, resulting in structures that do not likely have a physical basis.

4. Conclusions

The problem of optimal utilization of various informational sources in data assimilation is examined in this manuscript. It is shown that the representativeness error of observations, in addition to the instrumental error as well as the model error should be considered through proper weighting. In the case of an absence or incorrect accounting of different types of data, an imbalance in the data assimilation can arise and lead to an appearance of new structures in the atmospheric fields. These patterns have no physical basis, instead being born as a result of the numerical features of the forward and assimilation models.

Simplified study allows specification, in general, of the main features of such a contribution with respect to the distribution of the observations that play a role of the forcing. In our case, the response of the data assimilation model on the new informational source is reflected in the fluctuations extending out from the source and asymptotically extinguishing toward the boundaries. This is a specific feature of grid point models that is recognized in sensitivity analysis (Langland 1995). The critical point is that the characteristics of the wave trail are mainly determined by the model parameters, such as the model grid size and time step. Thus, the new structure in the atmospheric field which corresponds to the leading mode of the particular numerical model is purely of a numerical origin. Certainly, for spectral models one should anticipate another type of response, because all of the forcing energy will be distributed among modes corresponding to the spectral coefficients, and thus each mode could posses a smaller relative portion of the energy compared to the dominant mode within a gridpoint model.

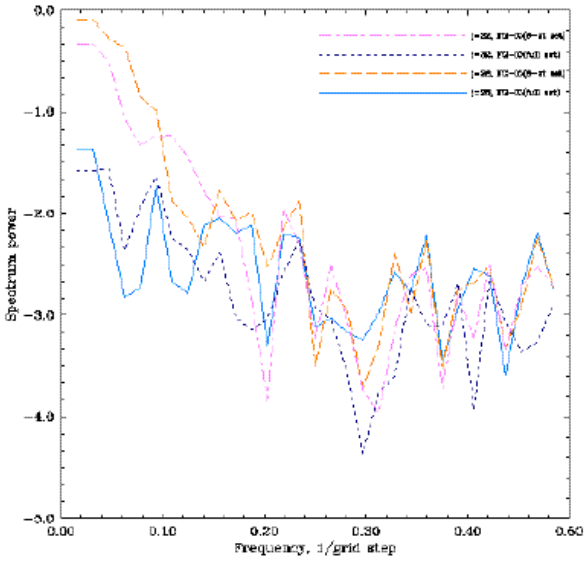

Figure 5. The spatial spectra of field differences through the approximate core of the data area. Numbers denote a corresponding model grid index. FG - first guess, IC - initial conditions.

Another feature of the response is the mesoscale structures whose sizes roughly correspond to the false side lobes in the spectral window obtained for the specified observation network. They appear due to the fact that an irregular distributed network is able to amplify or suppress certain scales from the continuous set (Kay 1993, Rabiner 1975). Alternatively, this means that each particular observation network acts as an antenna establishing priority to selective frequencies (scales) and giving them, in default, higher weighting coefficients when used within a data assimilation context.

Although application of any data assimilation model is intended to obtain "optimal" initial conditions on the basis of all available information, the result does not necessarily lead directly to an improved description of the atmosphere at the initial time. This occurs because the so-called "optimal" initial conditions are determined through the local minimum in a cost function under the constraints imposed. In this case, the closeness between "optimal" and "perfect" initial conditions will, of course, be related to the accuracy of various informational sources and the correctness in prescribing the weighting coefficients for them.

Although the described approach has been applied to a degraded observation net in a horizontal plane, we anticipate the same numerical effects in the vertical plane or in 3-dimensional space, when, for example, horizontally high resolution satellite data is provided on discrete vertical levels.

References

Dee D.P., Rukhovets L., Todling R., Da Silva A.M., Larson J.W. 2001. An adaptive buddy check for observational quality control. *Q.J.R.Meteorol.Soc.*, **127**, 2451-2471.

Desroziers G., Brachemi O., Hamadache B. 2001a. Estimation of the representativeness error caused by the incremental formulation of variational data assimilation. *Q.J.R.Meteorol.Soc.*,, **127**, 1775-1794.

Desroziers G., Ivanov S. 2001b. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q.J.R.Meteorol.Soc.*,, **127**, 1433-1452.

Findley, David F. 1978. Applied time series analysis. *Academic Press, N.Y., 345pp.*

Kay S. 1993 Fundamentals of statistical signal processing. *Prentice-Hall, Englewood Cliffs, N.J.*

Langland R.H., Elsberry R.L., Errico R.M. 1995. Evaluation of physical processes in an idealised extratropical cyclone using adjoint sensitivity. *Q.J.R.Meteorol.Soc.*,, **121**, 1349-1386.

Liu Z.-Q., Rabier F. 2002. The interaction between model resolution, observation resolution and observation density in data assimilation: An one-dimensional study. *Q.J.R.Meteorol.Soc.*, **128**, 1367-1386.

Lorenc A.C. 1986. Analysis methods for numerical weather prediction. *Q.J.R.Meteorol.Soc.*, **112**, 1177-1194.

Majumdar S.J., Bishop C.H., Etherton B.J., Szunyogh I., Toth Z. 2001. Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Q.J.R.Meteorol.Soc.*, **127**, 2803-2820.

Marsigli C., Montani A., Nerozzi F., Paccagnella T., Tibaldi S., Molteni F., Buizza R., 2001. A strategy for high-resolution ensemble prediction. Part II: Limited-area experiment in four Alpine flood events. *Q.J.R.Meteorol.Soc.*, **127**, 2095-2115.

Molteni F., Buizza R., Marsigli C., Montani A., Nerozzi F., Paccagnella T. 2001. A strategy for high-resolution ensemble prediction. Part I: Definition of representative members and global-model experiments. *Q.J.R.Meteorol.Soc.*, **127**, 2069-2094.

Peebles P. 1993. Probability, random variables and random signal principles. *McGraw Hill, N.Y.*

Rabiner L.R., 1975. Gold B. Theory and applications of digital signal processing. *Prentice-Hall, Englewood Cliffs, N.J.*

Schyberg H., Breivik L.-A. 2002. Objective analysis combining observation errors in physical space and observation space. *Q.J.R.Meteorol.Soc.*, **128**, 695-711.

Talagrand O. 1997. Assimilation of observations, an introduction. *J. of the Met. Soc. of Japan*, **75**, **1B**, 191-209.

Tanguay M., Laroche S. 2002. Grid-point response to the incremental strategy for variational applications. 2002. *Q.J.R.Meteorol.Soc.*, **128**, 385-397.

Zou X., Huang W., Xiao Q. 1998. A user's guide to the MM5 adjoint modelling system. *NCAR technical note.* NCAR TN-437+IA.