Chapter 3

Data assimilation on the SPEEDY primitive-equation model

3.1 Introduction

3.1.1 Overview

In Chapter 2, we described theoretical review and applied 3DVAR, serial EnSRF, LEKF, and full KF on the Lorenz-96 model. All KF methods (serial EnSRF, LEKF, and full KF) show similar results on the Lorenz-96 model, outperforming 3DVAR. In this chapter, we describe the implementation and numerical experiments of three data assimilation methods (3DVAR, serial EnSRF, and LEKF) on the SPEEDY primitive-equation model. We explore the following questions in this chapter:

1. How does each method work on the SPEEDY model?
2. What is the relative advantages and disadvantages between serial EnSRF and LEKF?
3. How sensitive is the data assimilation to experimental settings?
4. What are the characteristics of the analysis and forecast error fields?

A goal of this chapter is to investigate and compare the performance of the three methods. In addition, we perform sensitivity experiments with different observational networks, moisture observations, vertical error correlations, error covariance localization, and random perturbation addition to ensemble members. These sensitivity experiments investigate robustness of our results to the experimental settings. Moreover, they suggest possible ways to improve the filter performance. We also see the characteristics of the

analysis and forecast error fields, suggesting how EnKF works.

3.1.2 The SPEEDY model

The SPEEDY model (Molteni 2003) is a recently developed atmospheric general circulation model (AGCM) with a spectral primitive-equation dynamic core and a set of simplified physical parameterization schemes (SPEEDY stands for Simplified Parameterizations, primitivE-Equation DYnamics). The goal of this model is to achieve computational efficiency while maintaining characteristics similar to the state-of-the-art AGCMs with complex physics. The resolution of the model is T30L7 (horizontal spectral truncation of 30 wave numbers and 7 vertical levels), the computational cost is one order of magnitude less than that of state-of-the-art AGCMs at similar horizontal resolution. According to Molteni (2003), the SPEEDY model simulates the general structure of global atmospheric circulation fairly well, and some aspects of the systematic errors are similar to many AGCMs, though the error amplitude is larger than state-of-the-art models.

The SPEEDY model includes basic components of physical parameterizations used in more complex GCMs, such as convection (a simplified mass-flux scheme), large-scale condensation, clouds, short-wave radiation (two spectral bands), long wave radiation (four spectral bands), surface fluxes of momentum and energy (bulk aerodynamic formula), and vertical diffusion. Details of the simplified physical parameterization schemes of the SPEEDY model can be found in Molteni (2003), especially in its Appendix which is available on the website: "http://www.ictp.trieste.it/~moltenif/speedy-doc.html". The boundary conditions of the SPEEDY model includes topographic height and land-sea mask, which are constant, and sea surface temperature (SST), sea ice fraction, surface temperature in the top soil layer, moisture in the top soil layer and the root-zone layer,

and snow depth, all of which are specified by monthly means, and bare-surface albedo and fraction of land-surface covered by vegetation, which are specified by annual-mean fields. The lower boundary conditions such as SST are obtained by ECMWF's reanalysis in the period 1981-90. The incoming solar radiation flux and the boundary conditions (SST etc.), except bare-surface albedo and vegetation fraction, are updated daily.

Since the SPEEDY model is designed for long-term climate variability studies that require ensembles of long-term integrations, the original code contains only time-mean outputs. Thus, the input/output processes of a grid point value at intermittent time steps had to be developed in order to enable short-term integrations of the forecast-analysis cycle. The prognostic variables are zonal and meridional wind velocity components (u, v), temperature (T), specific humidity (q), and surface pressure (p_s). The grid point value output is in the physical space with the grid size of $96 \times 48 \times 7$. Inputs are only taken in sigma levels, but outputs are in both sigma levels and pressure levels. The specific heights of the vertical levels are shown in Table 3.1.

As for the timing results of running the model, a 3-month integration took about 6 minutes on a Linux PC with a 2.7GHz Intel Celeron (Northwood) processor, whereas it took around 8 minutes for 24-hour integration of the Japanese operational global weather prediction model with a reduced resolution (T42) on the same PC. For 6-hour cycle experiments, a 6-hour forecast of the SPEEDY model requires about 2 seconds on the same PC. Thus, a 2-month cycle experiment requires about 8-minute forecast computations for each ensemble member.

Level index	Sigma heights (σ)	Pressure heights (hPa)
1	0.950	925
2	0.835	850
3	0.685	700
4	0.510	500
5	0.340	300
6	0.200	200
7	0.080	100

Table 3.1: Vertical levels of the SPEEDY model outputs. Sigma levels are also used as model levels.

3.2 3DVAR implementation

3.2.1 Theory for practical implementation

The variational formulation has been introduced in Section 2.2.8. Practically, it is impossible to implement the variational formulation explicitly because of the large number of degrees of freedom of the system. Usually, NWP models have prognostic variables with dimensions of $\mathcal{O}(10^7)$, thus, the matrix $\mathbf{B}$ has $\mathcal{O}(10^{14})$ elements, which requires at least $\sim$ 10TB of memory just for storing the matrix. However, most components of the matrix $\mathbf{B}$ are very close to 0, and we can simplify $\mathbf{B}$ very much under reasonable assumptions. We can understand this fact by considering the forecast error of temperature at College Park, for example, has nothing to do with the forecast error of wind in Tokyo.

Parrish and Derber (1992) assumed zero spatial error correlation in the spectral space. With the spectral transformation, they reduced most components of $\mathbf{B}$. Alter-

natively, Barker et al. (2004) did not use spectral transformation in their 3DVAR built for the nonhydrostatic fifth-generation Pennsylvania State University/National Center for Atmospheric Research Mesoscale Model (MM5). We follow a similar approach as Barker et al.

We define a variable transformation $\mathbf{U}$ as follows:

$$\delta\mathbf{x} = \mathbf{U}\delta\mathbf{v} \tag{3.1}$$

Here, $\delta\mathbf{x}$ is defined by eq.(2.75), $\delta\mathbf{v}$ is a vector with the same dimension as $\delta\mathbf{x}$. Substituting eq.(3.1) into eq.(2.73), we get

$$J(\delta\mathbf{v}) = \frac{1}{2}\delta\mathbf{v}^\top\mathbf{U}^\top\mathbf{B}^{-1}\mathbf{U}\delta\mathbf{v} + \frac{1}{2}(\mathbf{HU}\delta\mathbf{v} - \mathbf{d})^\top\mathbf{R}^{-1}(\mathbf{HU}\delta\mathbf{v} - \mathbf{d}) \tag{3.2}$$

In order for the covariance matrix to be the identity,

$$\mathbf{B} = \mathbf{U}\mathbf{U}^\top \tag{3.3}$$

should be satisfied. Then, eq.(3.2) can be rewritten as

$$J(\delta\mathbf{v}) = \frac{1}{2}\delta\mathbf{v}^\top\delta\mathbf{v} + \frac{1}{2}(\mathbf{HU}\delta\mathbf{v} - \mathbf{d})^\top\mathbf{R}^{-1}(\mathbf{HU}\delta\mathbf{v} - \mathbf{d}) \tag{3.4}$$

Differencing eq.(3.4) with respect to $\delta\mathbf{v}$, we get the gradient of the cost function

$$\nabla J(\delta\mathbf{v}) = \delta\mathbf{v} + \mathbf{U}^\top\mathbf{H}^\top\mathbf{R}^{-1}(\mathbf{HU}\delta\mathbf{v} - \mathbf{d}) \tag{3.5}$$

At this point, we can solve the 3DVAR problem for $\delta\mathbf{v}$. Since we start the minimizing process from the background state, initially $\delta\mathbf{v} = 0$. Given the cost function (eq.(3.4)) and its gradient (eq.(3.5)), a quasi-Newton minimizer finds the solution for $\delta\mathbf{v}$, and eventually, eq.(3.1) converts $\delta\mathbf{v}$ to $\delta\mathbf{x}$, the analysis state in the real space. Note that only $\mathbf{U}$ and $\mathbf{U}^\top$ are used in this process; $\mathbf{U}^{-1}$ is not required.

The construction of the variable transformation **U** is the key in the 3DVAR algorithm. Usually it is assumed that **U** can be separated into a spatial correlation component and an inter-variable correlation component, that is,

$$\mathbf{U} = \mathbf{VCA} \tag{3.6}$$

where **A**, **C** and **V** stand for the error standard deviation, the spatial error correlation and the inter-variable error correlation, respectively. Moreover, it is assumed **C** can be separated into horizontal and vertical correlations. Usually, the horizontal correlation is assumed to be a Gaussian shape. Only correlation length scale parameter needs to be stored. Furthermore, the length scale is assumed to be constant in all points at the same vertical level. Quasi-isotropy is also assumed. Full vertical correlation could be considered, but it also can be assumed to be Gaussian.

An important part of 3DVAR is how to express the inter-variable correlation **V**, which transforms prognostic variables into control variables that are assumed to be independent of each other. Some prognostic variables are strongly dependent on each other mostly because of the geostrophic balance. Usually in global models, only the geostrophic balance is considered since it is the strongest and the most important balance in the system.

3.2.2 Implementation on the SPEEDY model

In the previous section, we described a general idea to implement 3DVAR on a practical atmospheric model following the approach by Barker et al. (2004). Here, we describe the design of the variable transformation **U** in the present implementation.

The first part is the error standard deviation **A** in eq.(3.6). Since the present 3DVAR analyzes not in the spectral space but in the physical grid space, we consider the full spatial

dependence of the error standard deviation. Almost no additional computation is required by considering the full spatial dependence.

For the spatial error correlation $\mathbf{C}$ in eq.(3.6), the spatial correlation is separated into vertical and horizontal correlations. The vertical resolution of the SPEEDY model is so coarse that vertical background error correlation is not considered (experiments including 3 levels yielded worse results than one layer). Horizontal error correlation is assumed to be Gaussian, and the 4th order recursive filter (RF) technique (Purser et al. 2003) has been implemented to estimate the Gaussian correlation shape. The 4th order RF is a quasi-Gaussian filter composed of a pair of forwarding and backwarding processes

$$B_i = \beta A_i + \alpha_1 B_{i-1} + \alpha_2 B_{i-2} + \alpha_3 B_{i-3} + \alpha_4 B_{i-4} \tag{3.7}$$

$$C_i = \beta B_i + \alpha_1 C_{i+1} + \alpha_2 C_{i+2} + \alpha_3 C_{i+3} + \alpha_4 C_{i+4} \tag{3.8}$$

where the subscript i denotes the space index of grid points in one direction. In eq.(3.7), that is the forwarding process in the ascending order of i, A is input and the filtered value B is obtained. In eq.(3.8), that is the backwarding process in the descending order of i, B is input and the filtered value C is obtained. Eq.(3.7) and eq.(3.8) are mutually adjoint, so eq.(3.7) may be used in $\mathbf{U}$ whereas eq.(3.8) may be used in $\mathbf{U}^\top$. The coefficients β and α can be obtained by solving a 4th degree equation according to the correlation length scale. The details of the recursive filter technique are described in Appendix B.

For the inter-variable error correlation $\mathbf{V}$ in eq.(3.6), we consider the geostrophic balance and introduce control variables as follows:

$$u_u = u - r_1 u_g(p_s, T) \tag{3.9}$$

$$v_u = v - r_2 v_g(p_s, T) \tag{3.10}$$

where r_1 and r_2 are regression coefficients determined from statistics, and u_g and v_g are

zonal and meridional components of the geostrophic wind computed from p_s and T using the geostrophic balance equation on the sigma coordinate system (cf. Kalnay 2003, p.65, eq.(2.6.27)):

$$f\mathbf{k} \times \mathbf{v}_g = -RT\nabla \ln p_s - \nabla\phi \tag{3.11}$$

where f, $\mathbf{k}$, R, and ϕ denote Coriolis parameter, vertical unit vector, gas constant, and geopotential height, respectively. Other variables (T, q, and p_s) are assumed to be independent of the unbalanced wind (u_u and v_u). Thus, only the geostrophic balance is considered for inter-variable correlation.

3.2.3 Background error statistics - NMC method

The design of variable transformation $\mathbf{U}$ described in the previous section requires background error statistics. We need to specify the error standard deviation, the length scale of horizontal error correlations, and the regression coefficients from the statistics. Background error statistics were computed using the NMC method (Parrish and Derber 1992), accumulating the difference between 24-hour forecast and 18-hour forecast as a sample of the background error. In the present experiment, 87 samples from January 10 to January 31 were accumulated for the statistics. The analysis fields for the first time are obtained using 3DVAR with a reasonable guess of background error statistics, such as 500km for the horizontal background error correlation length scale and the background error standard deviation with the same magnitude as observational error standard deviation. The NMC method provides better statistics. Once we get a more reasonable guess of background error statistics using the NMC method, this process may be repeated.

Fig.3.1 shows the zonal mean of the regression coefficients r_1 with respect to u-wind (eq.(3.9)). Almost no geostrophic balance is observed in the tropics, whereas strong

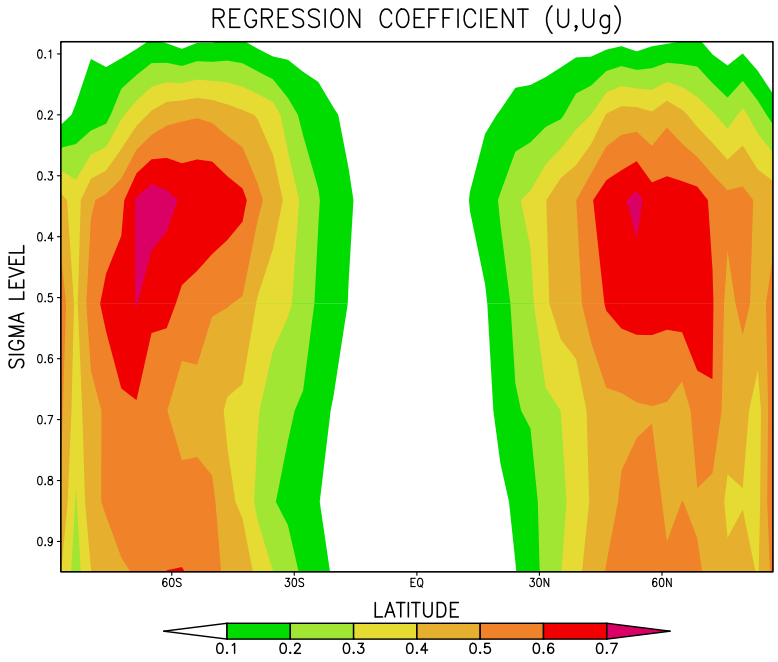

Figure 3.1: Zonal mean of the regression coefficients r_1 of u. The vertical and horizontal axes show height and latitude respectively. The coefficients reflect statistical strength of the geostrophic relationship. The larger values, i.e. stronger geostrophic balance, are observed in mid-latitudes.

geostrophy is observed in the mid-latitudes especially at the jet levels. The regression coefficients r_2 with respect to v-wind show a similar structure. Fig.3.2 shows the error standard deviation of p_s (top panel) and u at the 4th level (bottom panel), indicating strong spatial dependence, especially in zonal structures in error standard deviation. Fig.3.3 shows the vertical structure and latitudinal dependence of horizontal error correlations. There is almost no vertical dependence (top panel), all levels except the top show almost no correlation beyond 2 grid points. There is large latitudinal dependence in grid spacing (bottom panel). This reflects the skewness of the global map, grid spacing is physically more dense in higher latitudes.

The figures shown as background error statistics so far are only small parts of the

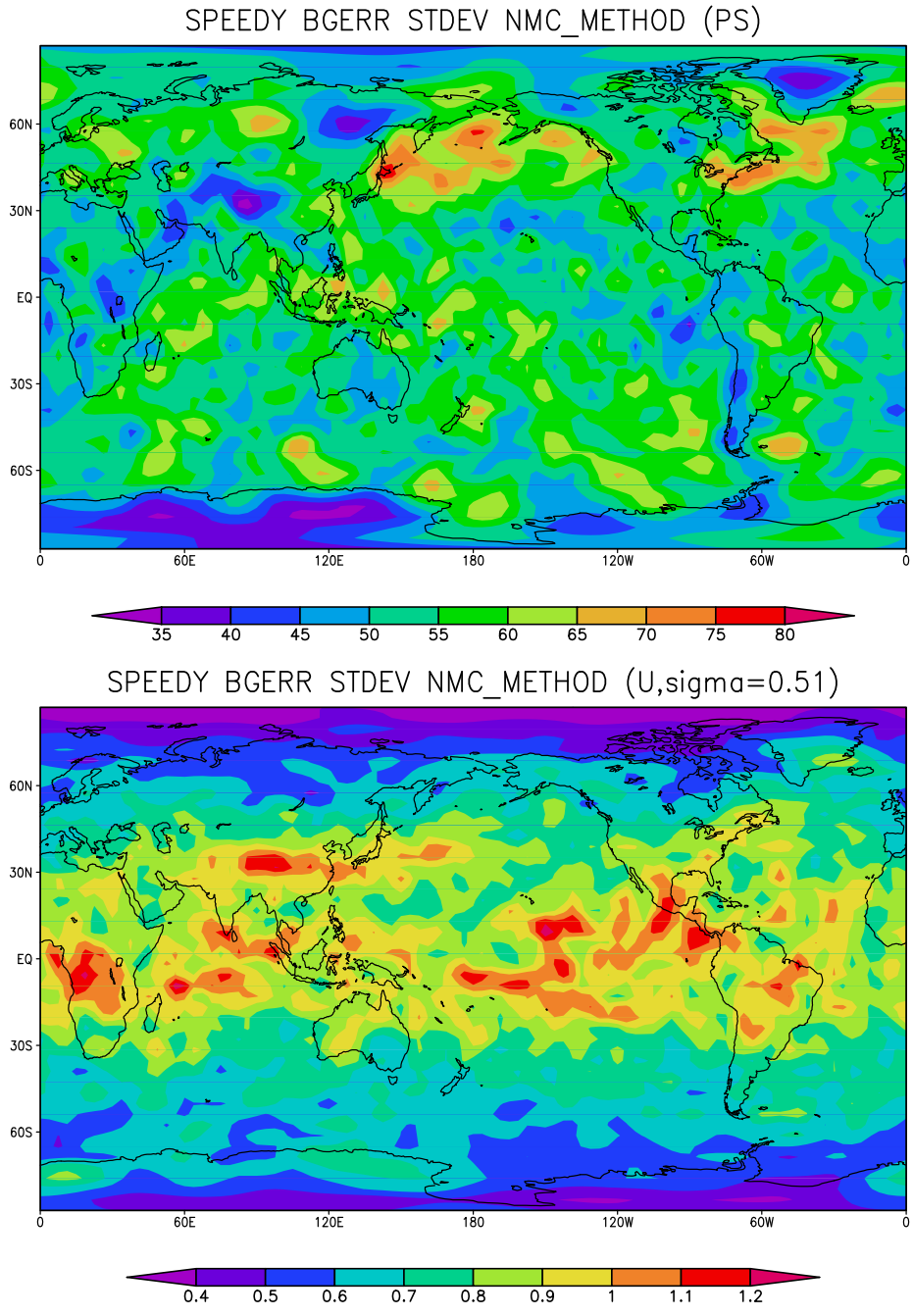

Figure 3.2: Background error standard deviation of p_s (top panel) and u at the 4th level (bottom panel) that are estimated using the NMC method. The units for p_s and u are Pa and m/s, respectively. Strong spatial dependence especially in zonal structures is observed, but their noisiness indicates sampling errors.

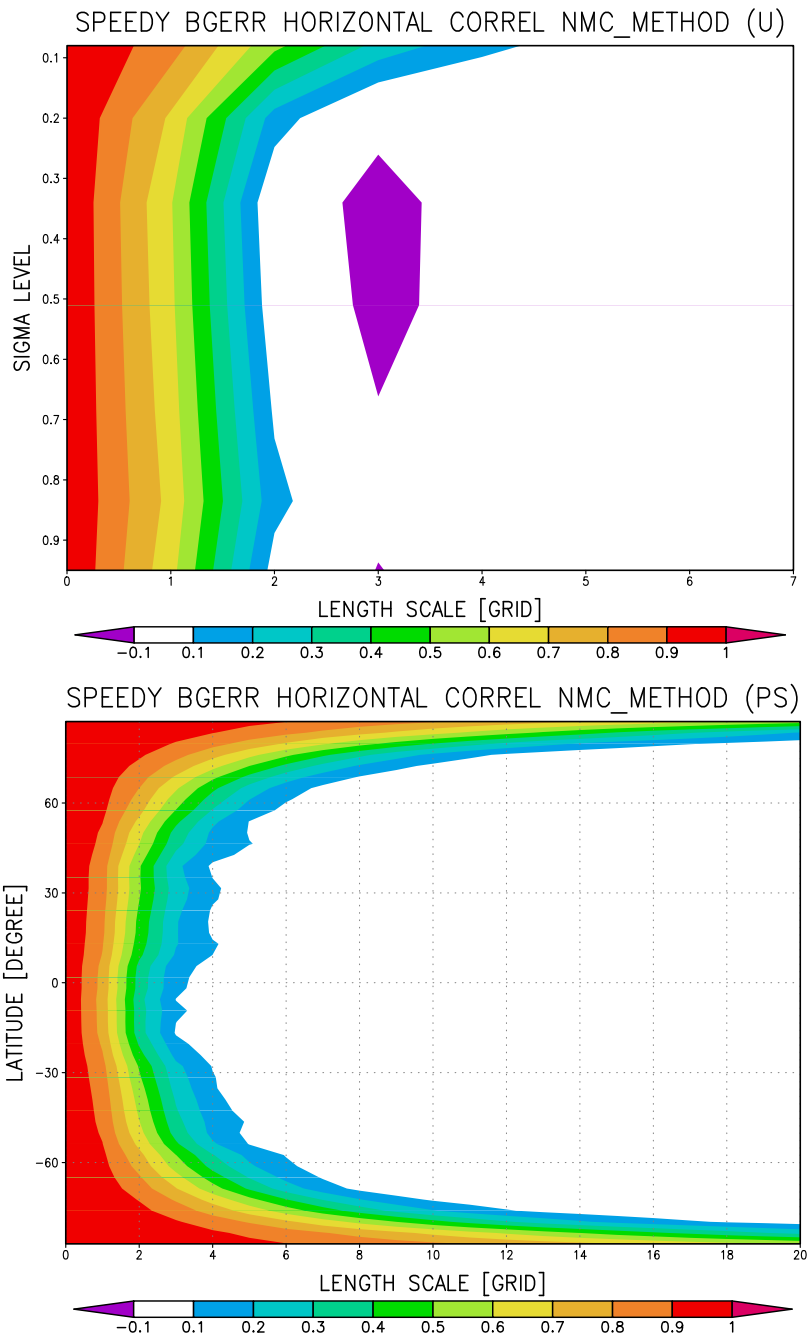

Figure 3.3: Background error horizontal correlations. The top panel shows vertical structure of u in y-direction, where the vertical axis shows sigma levels. The bottom panel shows the latitudinal structure of p_s in x-direction, where the vertical axis shows latitudes. The horizontal axis shows horizontal length scale in grid spacing. The shades show correlation.

background error information that are considered in the present 3DVAR. The present 3DVAR considers all components of standard deviation, the standard deviation has the same size as a grid point value. As horizontal spatial correlation, zonal and meridional components are stored separately. The zonal components are allowed to vary in latitude and height, whereas meridional components are allowed to vary only in height. The regression coefficients measure the strength of the geostrophic balance and are allowed to vary in height and latitudes. The regression coefficients for u (r_1) and v (r_2) are separately stored, although they should be equal in theory.

In order to examine our assumption of vertical independence, we compute vertical correlations averaged horizontally. Fig.3.4 shows vertical correlations using the NMC method. Some vertical correlations are seen at lower levels, but almost no correlation is observed at upper levels. Except for moisture, only the first and second levels have large correlations (about 0.4), other levels show less than 0.2 correlations. Thus, the vertical independence is not a bad assumption in 3DVAR. Moisture, shown by long-short dashed lines with black square, has largest vertical correlation at levels 2, 3, and 4, which may be related to vertical moisture transport by large-scale convective parameterization.

3.2.4 Response tests

We perform a response test using a single observation of u with the observational increment of 1.0m/s. Fig.3.5 shows responses when the observational station is located at mid-latitude (top panel) and high-latitude (bottom panel) at the 4th level. It shows how RF works, and it seems RF expands the observation signal to the Gaussian shape as expected. In the higher latitude, the skewness of the map is considered correctly. Fig.3.6 shows analysis increments of wind vector and T (top panel) and p_s (bottom panel) when

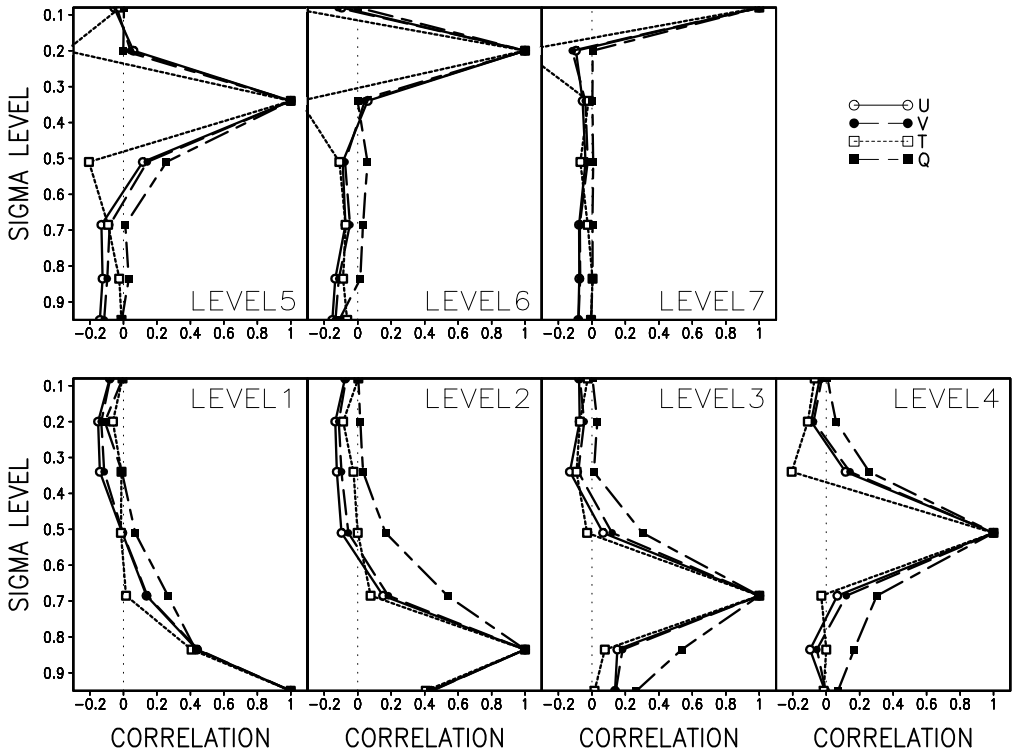

Figure 3.4: Vertical background error correlations of u-wind (solid lines with white circle), v-wind (dashed lines with black circle), temperature (short dashed lines with white square), and moisture (long-short dashed lines with black square). The correlations are averaged horizontally.

the observation is at mid-latitude at the 4th level. Although only u is observed, T and p_s are analyzed through the inter-variable correlation.

Since the single observation response tests show reasonable results, the 3DVAR is tested using the dense observational network (Fig.3.9). Fig.3.7 shows analysis increment (top panel) and the difference between truth and first guess (bottom panel). Analysis increment has a very similar structure as the difference between truth and first guess, which means the 3DVAR scheme succeeds as expected.

3.3 EnKF implementation

3.3.1 Serial EnSRF

We now apply the same core modules used for serial EnSRF in the Lorenz-96 model (Section 2.3) to the SPEEDY model. Appendix A describes the core modules. The core subroutines require one-dimensional arrays for inputs and outputs, all the variables are combined together to form a one-dimensional array. All inter-variable correlations are considered.

As described in Section 2.2.4, localization is very important to avoid sampling errors caused by the limited ensemble size. As for horizontal localization, distance is computed using the grid spacing unit. Thus, skewness of the map in the physical space is not considered. The boundary is cyclic in longitudinal direction, whereas in latitudinal direction, the boundary is treated as a solid wall. The Schur product with the Gaussian-like weighting function (eq.(2.49)) is applied in this grid spacing distance.

As in 3DVAR, errors in different vertical levels are considered to be independent. Surface pressure (p_s) is assumed to be correlated only with variables at the bottom level ($\sigma = 0.950$). Thus, p_s changes because of observations of p_s itself and all other variables