3.6 Regional analysis

3.6.1 Introduction

In June 2003, JMA has also implemented the 4-Dimensional Variational data assimilation (4D-Var) system for the Regional Spectral Model (RSM) instead of the 3-Dimensional Optimal Interpolation (3D-OI) system with a physical initialization for assimilating Radar-Raingauge Analyzed Precipitation data. To support a short-term forecast up to 2 days, RSM makes forecasts over East Asia area with 20 km horizontal resolution and 40 vertical levels up to 10 hPa and also provide the lateral boundary condition for MSM. The 4D-Var system for RSM has 6-hour assimilation window, i.e., ± 3 hours at 4 initial times (00, 06, 12, and 18UTC). Assimilated data are SYNOP, TEMP, PILOT, Wind Profiler, SHIP, BUOY, Aircraft, AMW, SATEM, ATOVS and Radar-Raingauge Analyzed Precipitation.

The design of the 4D-Var cycle is illustrated in Fig. 3.6.1.

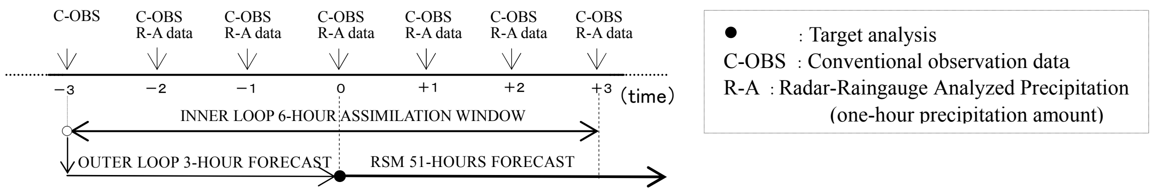

Fig. 3.6.1. Schematic illustration of the 4D-Var 6hour cycle design.

3.6.2 Basic formulation

4D-Var seeks the model trajectory in phase space that minimises the difference between model and observations in an assimilation window. The difference is measured by a cost function. The cost function is given by

$$\begin{aligned} J(\mathbf{x_0}) = & \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right) \\ & + \frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}\right)^{\mathrm{T}} \mathbf{O}^{-1}\left(H(\mathbf{x}) - \mathbf{y}\right) + J_c \end{aligned} \tag{3.6.1}$$

where the superscript T indicates the transpose of vectors or matrices, $\mathbf{x}_0^b$ denotes the first guess for the model state variables at the beginning of the assimilation window and the lateral boundary condition in the assimilation window, $\mathbf{y}$ the column vector consisting of observational data available in the assimilation window. $\mathbf{x}$ is the column vector consisting of model state variables at all time levels in the assimilation window, which is determined by the model, the initial condition and lateral boundary condition $\mathbf{x}_0$.

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{pmatrix} = M(\mathbf{x}_0) \tag{3.6.2}$$

where $\{x_i\}$ is the column vector consisting of the model state variables at time level i, and $M(\cdot)$ is the prediction equation of the model as a function of initial condition and lateral boundary condition. The total number of the time levels in the assimilation window is N+1. $H(\cdot)$ is the observation operator that converts model state variables to observed variables and interpolates from model grid points to observation points. $\mathbf{B}$ and $\mathbf{O}$ are the error covariance matrices for $\mathbf{x}_0^{\mathrm{b}}$ and $\mathbf{y}$ called the background error covariance matrices and the observation error covariance matrix, respectively, and J_c is the penalty term for suppressing gravity wave noise. The value of the cost function is calculated by a forward integration of the model. The penalty term is given as

$$J_c = \frac{1}{2} r \left(\frac{\partial D'}{\partial t} \right)^{\mathrm{T}} \frac{\partial D'}{\partial t} \tag{3.6.3}$$

where r denotes the penalty parameter, $(\partial D'/\partial t)^{\mathrm{T}}$ the column vector consisting of the perturbation of the time tendency of the horizontal divergence at all model grid points and all time levels. The time tendency is approximated by a two-time level finite difference scheme.

Minimization of the cost function using a smooth optimization algorithm requires the gradient of the cost function with respect to the initial condition and the boundary condition $\mathbf{x}_0$. It is written as follows.

$$\nabla_{\mathbf{x}_0} J = \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right) + \left(\frac{\partial M}{\partial \mathbf{x}_0} \right)^{\mathrm{T}} \left\{ \left(\frac{\partial H}{\partial \mathbf{x}} \right)^{\mathrm{T}} \mathbf{O}^{-1} \left[H(\mathbf{x}) - \mathrm{y} \right] + r \left(\frac{\partial}{\partial \mathbf{x}} \frac{\partial D}{\partial t} \right)^{\mathrm{T}} \frac{\partial D}{\partial t} \right\} \tag{3.6.4}$$

The transpose of the Jacobian matrices of the model, $(\partial M/\partial \mathbf{x}_0)^{\mathrm{T}}$, are the propagator matrices of the adjoint model. The basic state at each time level, which is necessary for the backward integration of the adjoint model, is provided by a forward integration of the original model Eqs. (3.6.2).

3.6.3 Background error covariance

(a) Control variable

RSM is a hydrostatic spectral model with a sigma-pressure hybrid vertical coordinate. The predicted variables are wind, virtual temperature, surface pressure and specific humidity in spectral space. The control variables of the 4D-Var system are unbalanced wind (u_{U}, v_{U}), virtual temperature T_v, surface pressure p_s, and specific humidity q in grid space at the beginning of the assimilation window, and the boundary conditions for them at the beginning and end of the assimilation window. The unbalanced wind is calculated by subtracting a wind component explained by pressure gradient force from the full wind as follows.

$$\begin{pmatrix} u_{\mathrm{U}} \\ v_{\mathrm{U}} \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} r_{xx} & r_{xy} \\ r_{yx} & r_{yy} \end{pmatrix} \begin{pmatrix} u_g \\ v_g \end{pmatrix} \tag{3.6.5}$$

where (u_g, v_g) denote the geostrophic wind and $\{r_{ij}\}$ are regression coefficients. It is assumed that the control variables $\{u_U, v_U, (T_v, p_s), q\}$ are uncorrelated with each other.

(b) Vertical correlation

The control variables normalized by their background error standard deviations are expanded in the eigenvectors of the background error vertical correlation matrices that are assumed to be homogeneous in the horizontal direction:

$$\begin{pmatrix} x_1/\sigma_1 \\ \vdots \\ x_K/\sigma_K \end{pmatrix} = \mathbf{U} \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_K \end{pmatrix} \tag{3.6.6}$$

where x_k denotes one of the control variables at a grid point at the k-th level, σ_k the background error standard deviation of x_k, $\mathbf{U}$ the orthogonal matrix consisting of the eigenvectors, $\tilde{x}_k$ the k-th expansion coefficient, and K the number of vertical levels. It is assumed that the background error standard deviations are horizontally homogeneous.

(c) Horizontal correlation

The horizontal correlations of the expansion coefficients are also assumed to be horizontally homogeneous. Then the horizontal correlation matrix of the k-th coefficient is written as

$$\mathbf{C}^{(k)} = \begin{pmatrix} \mathbf{C}_1^{(k)} & \mathbf{C}_2^{(k)} & \cdots & \mathbf{C}_N^{(k)} \\ \mathbf{C}_2^{(k)} & \mathbf{C}_1^{(k)} & & \mathbf{C}_{N-1}^{(k)} \\ \vdots & & \ddots & \vdots \\ \mathbf{C}_N^{(k)} & \mathbf{C}_{N-1}^{(k)} & \cdots & \mathbf{C}_1^{(k)} \end{pmatrix} \tag{3.6.7}$$

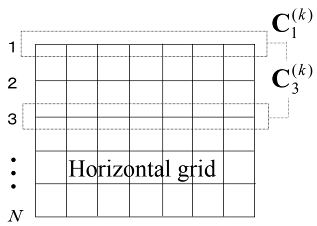

Fig.3.6.2. Homogeneous covariance

where $\mathbf{C}_j^{(k)}$ is the $M \times M$ - correlation matrix between the k-th expansion coefficients at the grid points in two rows that lie in the zonal direction and are separated by $(j-1)\Delta y$ in the meridional direction. M and N denote the numbers of grid points in the zonal and meridional directions, respectively, and Δy the meridional grid interval.

Horizontal correlation functions are assumed to be of a Gaussian type. Then the sub correlation matrices $\{\mathbf{C}_j^{(k)}\}$ are proportional to $\mathbf{C}_1^{(k)}$:

$$\mathbf{C}_j^{(k)} = \exp\left[-\frac{1}{2}\left(\frac{(j-1)\Delta y}{\eta^{(k)}}\right)\right]\mathbf{C}_1^{(k)} \qquad (j = 1, 2, \cdots, N) \tag{3.6.8}$$

$$= \varepsilon_j^{(k)} \mathbf{C}_1^{(k)}, \qquad \varepsilon_j^{(k)} = \exp\left[-\frac{1}{2}\left(\frac{(j-1)\Delta y}{\eta^{(k)}}\right)\right] \tag{3.6.9}$$

where $\eta^{(k)}$ is the meridional correlation length of the k-th expansion coefficient. The Cholesky decomposition matrix of $\mathbf{C}^{(k)}$ is also written by the Cholesky decomposition matrix of $\mathbf{C}_1^{(k)}$:

$$\mathbf{C}^{(k)} = \mathbf{L}^{(k)}\mathbf{L}^{(k)\mathrm{T}}, \qquad \mathbf{L}^{(k)} = \begin{pmatrix} a_{11}^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ a_{N1}^{(k)} & \cdots & a_{NN}^{(k)} \end{pmatrix} \otimes \mathbf{L}_1^{(k)} \tag{3.6.10}$$

where $\left\{a_{jj}^{(k)}\right\}$ are scalars calculated from the scalar coefficient in Eq. (3.6.10),

$$a_{j1}^{(k)} = \varepsilon_j^{(k)} \qquad (j = 1,2,\cdots,N) \tag{3.6.11}$$

$$a_{jj}^{(k)} = \sqrt{1 - \sum_{j'=1}^{j-1}\left(a_{jj'}^{(k)}\right)^2} \qquad (j = 2,3,\cdots,N) \tag{3.6.12}$$

$$a_{j+l\,j}^{(k)} = \frac{1}{a_{jj}}\left(\varepsilon_{l+1}^{(k)} - \sum_{j'=1}^{j-1} a_{jj'}^{(k)} \cdot a_{j+l\,j'}^{(k)}\right) \qquad (j = 2,3,\cdots,N-1,\ \ l = 1,2,\cdots,N-j) \tag{3.6.13}$$

$\otimes$ the tensor product, and $\mathbf{L}_1^{(k)}$ the $M \times M$ lower triangular matrix that satisfies the following equation.

$$\mathbf{C}_1^{(k)} = \mathbf{L}_1^{(k)}\mathbf{L}_1^{(k)\mathrm{T}} \tag{3.6.14}$$

The background error statistics were obtained from differences between 18- and 6-hour forecasts for the same valid time using the NMC method (Parrish and Derber, 1992). The elements of the correlation matrix $\mathbf{C}^{(k)}$ of which absolute values were less than 0.0001 were neglected to save computational time.

3.6.3 Adjoint model

While forward mode model in the 4D-VAR has all physics included, the adjoint model currently has dynamical processes, horizontal diffusion, and physics for four processes: large scale condensation and evaporation, moist convective adjustment, simplified vertical diffusion, and long-wave radiation, but convective parameterization is not implemented.

The algorithm of 4D-Var system is designed for the parallel computer with distributed memory. An increment approach (Courtier et al., 1994) will be used to save computational time and memory of a computer. The outer loop model is RSM with a horizontal resolution of 20km and 40 vertical levels. The inner loop model for calculating analysis increments is a 40km version of RSM, which is a reasonable expense for operational application.

3.6.4 Observations to be assimilated

Assimilated observational typical data are radiosonde, pilot balloon, wind profiler, aircraft, ship, buoy, geostationary satellite, polar orbit satellite and Radar-Raingauge Analyzed Precipitation. Their elements are surface pressure, (wind velocity u, v-components), temperature, relative humidity, and one-hour precipitation amount.

Table 3.6.1 shows the observation error tables used in the regional analysis (top 10hPa). The table means direct observations (a) and remote-sensing observations (b). The error at an arbitrary pressure level is linearly interpolated in the logarithm of pressure (log(*p*)).

Table 3.6.1 The observation error tables used in the operational regional analysis. *Ps*, *u*, *v*, *T* and *RH* denote surface pressure, (wind velocity *u*, *v*-components), temperature and relative humidity respectively.

(a) observational error (direct observation)					
element / *level*	*Ps* (hPa)	*u* (m/s)	*v* (m/s)	*T* (K)	*RH* (%)
1000hPa	0.50	2.10	2.30	1.30	8.9
850hPa		2.70	2.60	1.30	13.5
700hPa		2.70	2.60	1.00	15.2
500hPa		2.80	2.80	1.00	16.5
300hPa		3.60	3.60	1.10	17.0
200hPa		3.10	3.10	1.30	20.4
100hPa		2.80	2.50	1.40	13.0
50hPa		2.60	2.20	1.60	10.4
30hPa		2.70	2.50	1.80	10.2
10hPa		3.80	3.20	3.30	10.5

(b) observational error (remote sensing)					
element / *level*	*Ps* (hPa)	*u* (m/s)	*v* (m/s)	*T* (K)	*RH* (%)
1000hPa	1.87	3.40	3.30	3.40	13.0
850hPa		3.40	3.30	1.60	16.2
700hPa		3.10	3.20	1.20	20.5
500hPa		3.40	3.30	1.00	22.3
300hPa		3.70	3.10	1.10	23.7
200hPa		2.40	2.30	1.10	22.3
100hPa		2.80	2.80	1.00	19.6
50hPa		2.20	2.40	0.60	15.7
30hPa		3.20	2.90	1.00	15.7
10hPa		3.90	3.40	1.40	15.7

Observation error of precipitation σ_o is given as follows.

$$\sigma_o \equiv \begin{cases} \sigma_1 & (r \le r_o) \\ 3\sigma_1 & (r > r_o) \end{cases} \quad , \quad \sigma_1 \approx \begin{cases} 1\,mm/h & (r_o \le 1\,mm/h) \\ r_o & (r_o > 1\,mm/h) \end{cases} \qquad (3.6.15)$$

where r is one-hour precipitation amount provided by the model and r_o is observed one-hour precipitation amount.

<u>Reference</u>

Courtier, P., J.-N. Thepaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, 120, 1367–1387

Parrish, D. and J. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, 120, 1747–1763.

3.7 Meso-scale analysis

3.7.1 Introduction

The 4-Dimensional Variational data assimilation (4D-Var) system for the JMA Meso-Scale Model (MSM) had been developed since 1997 and has been operational in March 2002 with 3-hour assimilation window.

The 4D-Var system of meso-scale analysis has the identical framework of regional analysis except for horizontal resolution and analysis domain. The outer loop model of 4D-Var is RSM with a horizontal resolution of 10km and 40 vertical levels, and the inner loop model is a 20km version of RSM. The analysis domain is the same as the forecast domain of the MSM (see Section 4.5.2).

Formerly, initial conditions for MSM were prepared by the 3-hour PRE-RUN, in which 3D-OI for assimilating conventional data during the 3-hour period just before the initial time, and physical initialization for assimilating Radar-Raingauge Analyzed Precipitation data were conducted at one-hour intervals. PRE-RUN was replaced with the 4D-Var, which conducts 3-hour cycle analyses to prepare initial conditions for MSM.

In March 2006, the MSM was executed to produce 15 hour forecast every 3 hours at 8 times (00, 03, 06, 09, 12, 15, 18 and 21UTC initial time). The meso-scale analysis frequency was increased from 4 times to 8 times to conduct 3-hour cycle with 6-hour assimilation window.

3.7.2 Observations to be assimilated

Assimilated observational typical data are radiosonde, pilot balloon, wind profiler, aircraft, ship, buoy, geostationary satellite, polar orbit satellite and Radar-Raingauge Analyzed Precipitation. More and more new data, total column water vapor (TCWV) and precipitation retrievals from radiances of microwave radiometer (MWR) such as SSM/I, TMI, and AMSR-E, surface wind data from scattrometers such as QuikSCAT and doppler radars have been assimilated to enhance an accuracy of the initial condition. Their elements are surface pressure, (wind velocity u, v-components), temperature, relative humidity, one-hour precipitation amount, rain rate, precipitable water and radial velocity. Rain rate and total column precipitable water data retrieved from TMI, SSM/I and AMSR-E data. TMI, SSM/I and AMSR-E are satellite-borne microwave imagers. The radial velocity data of operational doppler radars are assimilated directly into the meso-scale analysis.

Table 3.7.1 shows the observation error tables used in the meso-scale analysis (top 10hPa). The table means direct observations (a) and remote-sensing observations (b). The error at an arbitrary pressure level is linearly interpolated in the logarithm of pressure (log(p)).

Table 3.7.1 The observation error tables used in the operational meso-scale analysis. *Ps*, *u*, *v*, *T* and *RH* denote surface pressure, (wind velocity *u*, *v*-components), temperature and relative humidity respectively.

(a) observational error (direct observation)				
element / *level*	*Ps* (hPa)	*u*, *v* (m/s)	*T* (K)	*RH* (%)
1000hPa	0.62	1.00	0.25	2.80
850hPa		1.00	0.31	4.59
700hPa		1.07	0.33	5.79
500hPa		1.20	0.32	7.30
300hPa		1.52	0.58	13.42
200hPa		1.77	0.68	13.82
100hPa		2.20	0.71	7.23
50hPa		2.29	0.59	3.50
30hPa		2.36	0.59	3.50
10hPa		2.50	0.57	3.50

(b) observational error (remote sensing)				
element / *level*	*Ps* (hPa)	*u*, *v* (m/s)	*T* (K)	*RH* (%)
1000hPa	1.87	3.00	0.50	4.90
850hPa		3.00	0.77	10.14
700hPa		3.00	0.82	11.03
500hPa		3.00	0.81	16.98
300hPa		3.48	1.31	22.38
200hPa		3.85	1.43	19.29
100hPa		4.50	1.38	5.25
50hPa		4.65	1.09	5.25
30hPa		4.76	1.04	5.25
10hPa		5.00	0.95	5.25

In order to provide forecast within one and half hours from analysis times, data cut-off time is set to fifty minutes. Because of the short data cut-off time, several conventional observation data such as most of the overseas upper air reports and most of the satellite observations, are not used on just analysis time. But the meso-scale analysis can use them in next 3-hour cycle with 6-hour assimilation window.