이학박사학위논문

Data Assimilation of Radar and AWS Observations for Improving Heavy Rainfall Prediction

(집중호우의 예보 향상을 위한 레이더 및 AWS 관측 자료의 동화 연구)

2012월 8월

서울대학교 대학원
지구환경과학부
하 지 현

Data Assimilation of Radar and AWS Observations for Improving Heavy Rainfall Prediction

(집중호우의 예보 향상을 위한 레이더 및 AWS 관측 자료의 동화 연구)

지도교수 임 규 호

이 논문을 이학박사 학위논문으로 제출함.
2012월 4월

서울대학교 대학원
지구환경과학부
하 지 현

하지현의 이학박사 학위논문을 인준함.
2012년 6월

위 원 장 _____(인)
부위원장 _____(인)
위 원 _____(인)
위 원 _____(인)
위 원 _____(인)
위 원 _____(인)

Abstract

This study investigated the impact of multiple-Doppler radar and AWS surface data assimilation for improving the accuracy of heavy rainfall forecast; the Weather Research and Forecasting (WRF) and its three-dimensional variational data assimilation (3DVAR) were used for this purpose. In the data assimilation, the WRF 3DVAR cycling mode with incremental analysis updates (IAU) was used to remove the high-frequency gravity wave. To evaluate the impact of the data assimilation, a heavy rainfall case on 11-12 July 2006 associated with the back-building mesoscale convective systems (MCSs) was chosen. Using the IAU method, the gravity wave fluctuation was greatly reduced and the noise was effectively removed, which help to reduce aliasing in subsequent analyses. Prior to the investigation for the impact of radar and surface data assimilation, it was firstly assessed that the assimilation of multiple-Doppler radar and surface data assured the improvement in the accuracy of heavy rainfall forecast. The assimilation of both radar and surface data showed the best agreement with the observations in terms of location and amount of rainfall, and had a more positive impact on the quantitative precipitation forecasting (QPF) than the assimilation of either radar data or surface data only. In addition, the back-building characteristic was successfully forecasted. Based on the data assimilation experiments, the radar data helped forecast the development of convective storms responsible for the heavy rainfall in the early hours of forecast, and the

surface data contributed to the occurrence of intensified low-level winds. Further, the surface data played a significant role in enhancing the thermal gradient and modulating the planetary boundary layer of the model, which resulted in favorable conditions for convection.

In the assimilation of high-resolution surface data, National Meteorological Center (NMC) method estimate of background error tended to exaggerate the length scale that determined the shape and extent to which observed information spreads out. For effective use of surface data to improve forecast accuracy of the heavy rainfall, the NMC method estimate of background error was tuned by comparing with independent estimates from accumulated observation minus background (O-B) data. A comparison revealed that the length scale of the NMC method should be halved in order to better assimilate the surface data with that of O-B. However, the correlation between NMC method and O-B statistics was still poor even using the half of the length scale of the NMC method, therefore, in this study, we examined a double iteration method with two different scales representing the large and small lengths.

The resulting assimilation clearly showed that the analysis with the tuned length scale was able to reproduce the small-scale features of the ideal field effectively. Further, the analysis using the double iteration method reflected the large- and small-scale features of observed information in the model fields, allowing the 3DVAR system to extract high-resolution observed information more effectively. The precipitation forecast using this double iteration with two different length scales for

the heavy rainfall case was in good agreement with the observations in terms of rainfall distribution and amount. The improved forecast resulted from the development of well-identified MCSs by intensified low-level winds and their consequent convergence near the rainfall area. In addition, we assessed the impact of the background error by the double iteration method on the improvement of the analysis through the assimilation of surface data during a one-month period, in comparison with the background error estimated by the NMC method. The statistics for the one-month period indicated that the 3DVAR analysis using the double iteration improved the root-mean-square-errors (RMSEs) verified against the surface observations. These results indicate that the prediction of the heavy rainfall can be improved by designing a suitable strategy of the background error for assimilating the surface data.

Regarding to the surface observations, the surface observation impact in improving 6 hour forecast was evaluated for optimal use of surface observations and forecast skill improvement. The observation impact was evaluated during the warm season, with variant formula of third-order approximation of forecast error variation using WRF adjoint model. It was concluded that wind observations showed larger impact in improving the 6 hour forecast than thermodynamic observations. Among the wind observations, the meridional wind showed the largest impact in reducing 6 hour forecast error.

Keyword: radar and surface data, data assimilation, 3DVAR, mesoscale

convective system, heavy rainfall, tuning

Student number: 2005-30149

Table of Contents

List of tables

List of figures

Chapter 1 Introduction

1.1 Motivation and Objectives

A significant portion (normally 55 %, or about 700 mm) of annual precipitation on the Korean Peninsula is due to heavy rainfalls during summer (KMA, 2001). Heavy rainfall during the summer season is one of the most significant factors in natural disasters on the Korean Peninsula. According to the 10-year statistics (2000-2009), the damage to properties caused by the heavy rainfall approximately equals 0.7 billions US dollars and 61 % of the total damage to properties by natural disasters (National Emergency Management Agency, 2009). Loss of life and property damage has increased yearly owing to the development of a socio-economy and the increases in the population. Most heavy rainfall events result from the development of intense mesoscale convective systems (MCSs) (Zipser, 1982) under certain synoptic-scale environmental conditions; examples of such systems include a surface frontal system accompanying an upper-level trough during the *Changma* (referred to as the *biau* in Japan and the *meiyu* in China) from mid-June to mid-July or strong instability in the vicinity of the subtropical Pacific high in the post-*Changma* (Lee et al., 1998). Observation and numerical simulation studies show that the dynamic and physical process of heavy rainfall over the Korean Peninsula can be indentified and MCSs accompanying heavy rainfall are predictable (Lee et al. 1998; Wee 1999; Sun and Lee 2002).

However, it remains a challenge for operational forecasters to identify the location and evolution of MCSs, and the mechanisms. Therefore, one of the greatest challenges in operational forecasting is to predict the evolution and development of MCSs that lead to heavy rainfall and to understand the mechanisms governing their evolution and development processes.

Radar data provide effective information for studying convective systems and are widely used to investigate MCSs characteristics. In studies of MCSs since the 1970s, Doppler radar imagery has played an increasingly important role (Burgess and Ray 1986). Bluestein and Jain (1985) studied mesoscale convective-line development in the United States on the basis of reflectivity data, and suggested the existence of line formation classes. Schumacher and Johnson (2005) investigated the structure and evolution of extreme rain events using composite radar reflectivity data, and observed two MCS organization patterns occurring frequently in the United States: training line/adjoining stratiform systems and back-building/quasi-stationary MCSs. In Korea, Chun and Oh (1991) retrieved the horizontal wind field from a single Doppler radar image and showed that maximum convergence occurred prior to the period of maximum rainfall intensity and that the maximum convergence corresponded with the rainfall intensity zone. Kim and Lee (2006) investigated the characteristics of MCSs accompanying heavy rainfall using a single data of Weather Surveillance Radar-1988 Doppler (WSR-88D). They demonstrated that low-level vertical wind shear played an

important role in the development of quasi-stationary and multicell storms, and that the convective systems merged and stagnated due to blocking by mountain ranges in the Korean Peninsula. However, Rinehart (1997) pointed out that the application of single radar data was limited to investigate entire MCSs, and suggested the use of routinely combined data from multiple radars to cover entire events. In this context, Gao et al. (1999) proposed a variation method for the analysis of three-dimensional wind fields from two Doppler radars in order to gain a more complete understanding. Park and Lee (2009) retrieved high-resolution wind fields by combining data obtained from a total of 18 operational radars over the Southern Korean Peninsula.

Numerical studies of MCSs have been performed to complement observational studies, which are insufficient on their own. Kato (1998) successfully simulated a convective system incorporating a back-building system by using a non-hydrostatic model with a horizontal grid of 2 km; it was shown that the quasi-stationary convergence line near the surface and the strong wind shear in the lower layer played significant roles in the successive generation of convective cells upstream of the rainband. Sun and Lee (2002) simulated an insensible long-lived convergence formed rapidly on the upwind side of the convection initiation point, after which convective cells developed and moved along this line. Parker (2007) performed real and ideal simulations to investigate the basic kinematics and precipitation features of a parallel stratiform convective system. However, several simulations showed that the initiation of

convective cells was not successfully simulated as shown in the Doppler radar data; this was because the initial and boundary data for the models were limited to identify the storms due to coarse temporal and spatial resolution of used data. Therefore, it was suggested that asynoptic data assimilation was needed to improve convective cell prediction and that more experimental studies on the assimilation of asynoptic data should be conducted (Wee, 1999; Lee and Lee, 2003). Also it is generally accepted that the use of data with a high spatial and temporal resolution has led to significant improvements in the accuracy of short-range forecast and flash flood predictions in the United States (Serafin and Wilson, 2000).

With the technological advances, various sources of data with high spatial and temporal resolutions have become available, such as Doppler radars, high-frequency automatic weather stations (AWSs), wind profilers, and space-borne sensors. Among these observational sources, multiple-Doppler radars and AWSs have been constructed as an operational network to observe and analyze weather conditions in the southern Korean Peninsula with the high spatial and temporal resolution (Fig. 1.1). Several studies on data assimilation using radar and surface data have been conducted, and the results of these studies have demonstrated that the prediction with the assimilation of these data can potentially be improved. Crook and Sun (2002) described a variational scheme for the analysis of low-level wind data by assimilating data from two Doppler radars, a surface mesonet, and a boundary layer profiler for cases of sea breezes, southerly changes, and a severe tornadic hailstorm,

and showed a reasonable agreement between the analysis wind and the dual-Doppler radar winds, particularly in terms of wind direction. Gu et al. (2005) indicated that assimilation of the Korean AWS surface data improved the landfall-typhoon track and its precipitation. Xiao et al. (2005) assimilated the radar radial velocity data by using the fifth-generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model (MM5) and its three-dimensional variational data system, and demonstrated potential ability for the radar observations to improve the rainfall forecasts. Lee et al. (2006) showed the application of the incremental analysis updates to improve initialization performance and short-range forecast using AWS surface data. Xiao et al. (2007) investigated the impact of multiple-Doppler radar data assimilation on quantitative precipitation forecast (QPF) of a squall line, and demonstrated that assimilation of both radial velocity and reflectivity had more positive impact on the QPF skill with multiple-radar data assimilation than the assimilation of either radial velocity or reflectivity only. In addition, they demonstrated that the more radar data in the temporal and spatial dimension were assimilated, the more positive was the impact on the QPF skill. Although these studies have shown promising results, there are no reports of numerical studies on heavy rainfall events involving the assimilation of both radar and surface data, and the way in which each data source contributes to the improvement of heavy rainfall forecast has not yet been studied. While AWSs collect data for various variables at the surface level only, the

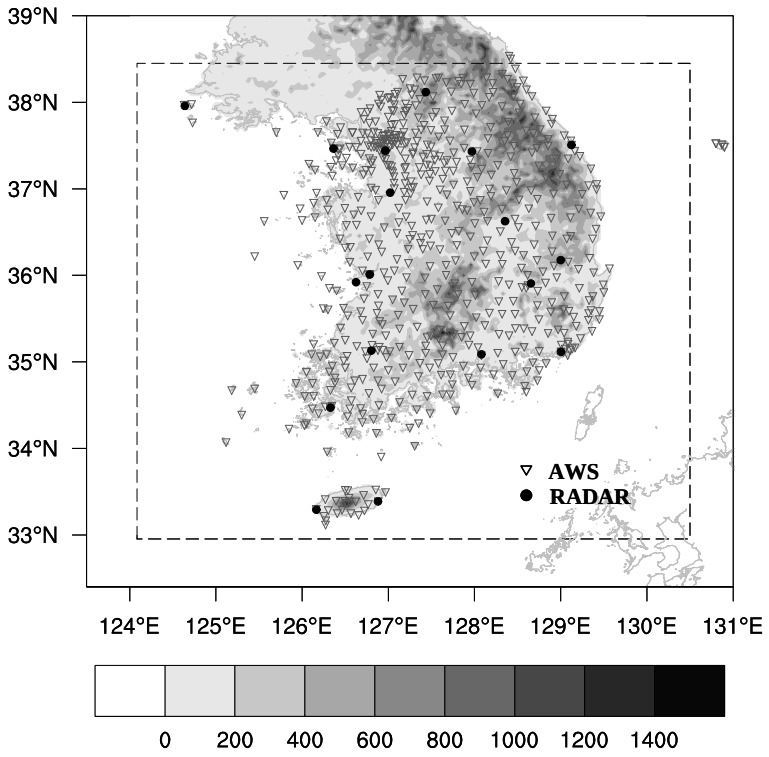

Figure 1.1 Topography (m, shaded), and locations of the operational AWS (triangles) and radars (dots) sites over the southern Korean Peninsula. The dashed box denotes the area analyzed in Fig. 5.13 and 5.14.

radars collect data on a limited number of three-dimensional variables such as radial velocity and reflectivity of hydrometeors. Therefore, the first objective of this study is to investigate the evolution and development of MCSs responsible for a heavy rainfall event using a Doppler radar dataset that combined data from 18 radars, and to investigate the impact of radar and surface data assimilation for improving the accuracy of heavy rainfall forecasts.

Besides the observations, another important input for data assimilation systems is the error statistics because a data assimilation system relies on specified error statistics to determine the optimal combination of observation and background. Observation data are insufficient, in terms of the number of grid points times the number of variables, to describe the atmosphere. Therefore, in the assimilation system, estimates of atmospheric variables from observation data must be supplemented by information from a background or first guess, detailed background and observation error statistics, and the law of physics. Among the error statistics, the background error statistics determine the respective weights given to each variable of the analysis and define the extent to which the observed information propagates horizontally and vertically in these variables. Therefore, correct specification of the background error statistics remains a major challenge in the assimilation system.

In order to create these background error statistics, the National Meteorological Center (NMC) method (Parrish and Derber, 1992) is

widely used in three-dimensional (3DVAR) and four-dimensional variatioinal (4DVAR) systems. This method provides a climatological estimate of background error statistics assuming that it is well approximated by averaged forecast difference statistics (e.g., month-long series of 24 h minus 12 h forecast valid at the same time).

$$B = \overline{(X^b - X^t)(X^b - X^t)^T} = \overline{\varepsilon_b \varepsilon_b^T} \sim \overline{[X^f(T+24) - X^f(T+12)][X^f(T+24) - X^f(T+12)]^T} \quad (1)$$

where, X^b is the background, X^t is the true atmospheric state and ε_b is the background error. $X^f(T+24)$ and $X^f(T+12)$ are, respectively, 24 h and 12 h forecasts valid at the same time and the overbar denotes an average over time and/or space. This approach is very practical, but the background error covariance seems to differ in several ways from the true background error covariance. The main shortcoming of the NMC method is probably the evolution of the statistics of the forecast error in the short-range forecast (12-24 hour or 24-48 hour). Further, it is well known that the background error variance for winds tends to be overestimated, and therefore, the spatial correlation scales are excessively large (Lee et al., 2006). This implied that that the small-scale observed details tended to be filtered out in the analysis step, and that locally observed information was propagated over large spatial distances (Daley, 1991).

A number of numerical studies for tuning the error statistics have been performed to complement the background error statistics of the NMC method. Ingleby (2001) showed that the NMC method statistics

required an additional tuning in order to improve model performance. The *a posteriori* diagnosis approach of Desroziers and Ivanov (2001) tuned the background and observation error weighting parameters computed from analysis residuals. Gu et al. (2005) studied the tuned background error statistics by comparing between 3DVAR analysis and observations, and indicated that the predictions of typhoon track and its precipitation were improved by background error statistics tuning. Guo et al. (2006) performed an assimilation of bogus typhoon data with tuned background error statistics, and reported an improvement in the typhoon analysis in terms of its track and intensity. Skamarock et al. (2008) noted that, in variatainoal systems where the background error covariance was computed off-line, significant tuning was required to optimize the performance for any particular application. Lee et al. (2010) showed that quantitative features of the heavy rainfall case were improved by tuning of background and observation error statistics in the assimilated radar retrieval wind data. Besides the above research, several more studies on data assimilation with tuned error statistics have been conducted (Wu et al., 2002; Barker et al., 2004; Chapnik et al., 2004). However, there are currently no reports on the use of high-resolution surface observation data to tune the background error statistics of heavy rainfall events. The effective use of the AWS data to improve forecast accuracy of heavy rainfall is, therefore, a challenge, and it is necessary to examine the approach of tuned background error statistics in order to improve the assimilation of high-resolution surface observation and model forecast

fields. Therefore, another objective of this study is to investigate the effect of the tuned background error statistics calculated by the NMC method in a 3DVAR assimilation of high-resolution surface data, and assess their impact on the model for a heavy rainfall case.

The purpose of this study is to investigate the evolution and development of MCSs responsible for a heavy rainfall event using a Doppler radar dataset that combined data from 18 radars, and to investigate the impact of radar and surface data, and the tuned background error statistics in assimilation of surface data for improving the accuracy of forecast. The heavy rainfall event is simulated using the Weather Research and Forecasting (WRF) model by assimilating the radar and AWS data. The rest of this paper is organized as follows. The WRF 3DVAR system and the proposed method for tuning of background error statistics are presented in section 2. Section 3 describes the data used for analysis and data assimilation. The synoptic situation and the storm environment by the Doppler radar data are explained in section 4. In section 5, the numerical simulations and results are presented. Finally, a summary and conclusions are given in section 6.

Chapter 2 WRF 3DVAR system and tuning of background error correlation length scale

2.1 WRF 3DVAR assimilation system

The numerical model used in this study is the WRF model. WRF model has been developed as a next-generation mesosclae numerical weather prediction system for high-impact weather features across scales ranging from cloud to synoptic, with priority emphasis on horizontal girds of several kilometers. WRF equations are in terms of a terrain-following hydrostatics pressure vertical coordinate. The WRF dynamical core uses Eulerian finite-differencing to integrate the fully compressible non-hydrostatic equations in mass-coordinate, scalar-conserving flux form using a time-split small step for acoustic modes. Large time steps utilize a third-order Runge-Kutta technique. The horizontal staggering is an Arakawa C grid. The WRF physical processes for real-data mesoscale forecasts include the selections of explicit microphysics schemes, cumulus convective parameterization, planetary boundary layer parameterizations, and longwave and shortwave radiation schemes.

The WRF 3DVAR originated and evolved from MM5 3DVAR (Barker et al. 2004), but the basic software interface and coordinate framework are fully updated for the WRF model (Skamarock et al. 2008). The basic aim of the WRF 3DVAR system is to determine an optimal estimate of the true atmospheric state at the time of analysis by means of an iterative

solution of the following prescribed cost function $J(x)$:

$$J(x) = J^b + J^o = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - y_o)^T O^{-1}(y - y_o) \quad (2)$$

where, x is the analysis vector, x_b is the background vector, y_o is the observation vector and y is the analysis observation vector transformed by the observation operator, B and O are the background and observation error covariance matrices, respectively, and $(\bullet)^T$ and $(\bullet)^{-1}$ denote the matrix transpose and inverse, respectively. In Eq. (2), the analysis $x = x^a$ represents the posteriori maximum likelihood estimate of the true state of the atmosphere given two sources of data: the background (previous forecast) x_b and observation y_o (Lorenc 1986). The analysis fit to this data is weighted by estimates of background and observation error covariances. The cost function in (2) assumes that observation and background error are described using Gaussian probability density functions (PDFs) with 0 mean error. Correlations between observation and background errors are neglected in (2) as is typical in 3/4DVAR system (Parrish and Derber 1992; Zou et al., 1997; Lorenc et al. 2000). To obtain the background error statistics in this study, we carried out 12 h and 24 h forecasts using WRF from 0000 UTC 1 July to 1200 UTC 31 July 2006 at 12 h intervals. The model configuration for the background error statistics was identical with the one in the numerical experiments as mentioned in section 5.1. The averaged forecast difference (month-long series of 24 h minus 12 h forecast valid at the same time) was used to estimate the background errors by the NMC method.

Given a model state x with n degrees of freedom, calculation of full background J^b term in Eq. (2) requires $\sim O(n^2)$ calculations. For a typical numerical weather prediction model with $n^2 \sim 10^{12} \sim 10^{14}$, direct solution is not feasible in the time slot allotted for data assimilation in operational applications. One practical solution to reduce computational cost is to calculate J^b in terms of control variables defined via the relationship $x' = Uv$, where $x' = x - x^b$ is the analysis increment. The U transform is well designed by a series of operations $U = U_p U_v U_h$ so as that $B = UU^T$ (Lorenc et al. 2000). In addition, the U transform is designed to nondimensionalize the variational problem and also to permit use of efficient filtering techniques that approximate the full background error statistics. Using the incremental formulation (Courtier et al., 1994) and the control variable transform, Eq. (2) can be rewritten as:

$$J(v) = \frac{1}{2} v^T v + \frac{1}{2} (d - H' Uv)^T O^{-1} (d - H' Uv) \tag{3}$$

where $d = y_o - \mathrm{H}(x_b)$ is the innovation vector and H' is the linearization of the observation operator.

The horizontal transform of background error (U_h) is performed using recursive filters (Hayden and Purser, 1995; Purser et al., 2003a). The first-order recursive filter consists of an advancing step:

$$F_i = (1-\alpha) D_i + \alpha F_{i-1} \tag{4}$$

where D is the input and F is the output of the advancing step. In Eq. (4), i is the grid point index and must be treated in increasing order so that the quantities on the right-hand side of the equation are already

known. The advancing step is followed by the backing step:

$$R_i = (1-\alpha)F_i + \alpha R_{i+1} \tag{5}$$

where R is the final output from the recursive filter, and i must be treated in decreasing order. The smoothing parameter α is the filter coefficient, and has the following relationship with a new parameter e:

$$\alpha = 1 + e - \sqrt{e(e+2)}\text{, where } e = N(\Delta x)^2 / 4s^2 \tag{6}$$

where N is the number of iterations of the filter (as $N \to \infty$ the response approximates a Gaussian), Δx is the horizontal resolution of the forecast model, and s is the length scale of the filter. In the recursive filter, N and Δx are constant, and therefore the value of e varies according to the length scale. The smoothing parameter α, which lies between 0 and 1, is related to the length scale of the recursive filter. Barker et al. (2004) showed that a value of $N = 6$ was the minimum number of passes required to remove unphysical correlations in the wind field, and therefore we used this number in all applications. In this study, the length scale was estimated using the accumulated forecast difference data from the NMC method, processed as a function of grid-point separation. A least square fit of the resulting curve to a Gaussian function was then used to estimate the recursive filter length scale. The vertical dependence of length scale in a horizontal transform of background error for control variables is illustrated in Fig. 2.1 Figure 2.1 showed a general trend of increasing length scale as a function of decreasing vertical level, which represented the dominant error of the large scale (synoptic scale)

at the low level. The small-scale nature of humidity errors relative to wind errors was noticeable. A valid question was whether the small-scale forecast difference truly represented background error features or due to artifacts of the numerical forecast, for example, boundary conditions, noise, etc.

The vertical component of background error (U_v) is applied via an empirical orthogonal function (EOF) decomposition on model levels. Given a domain/time-averaged estimate of vertical component of background error, an eigendecomposition is performed to compute eigenvectors and eigenvalues. The projection onto orthogonal eigenvectors reduces the number of calculations required in the vertical component of background error from $O(K^2)$ to $O(K)$. By definition, the leading eigenvector contains the largest contribution to the background error, while, trailing eigenvectors contain the least information and may be removed to reduce the computational cost of 3DVAR. Barker et al. (2004) illustrated that for a sample 3DVAR analysis, truncation of vertical modes to retain 99.9 % of the background error variance resulted in a significant reduction in CPU and memory requirements. Furthermore, there was little impact on the final results. The truncation of vertical modes results in a significant cost reduction with negligible scientific impact and hence, in this study, the truncation of vertical component of background error is used.

The physical variable transformation (U_p) involves the conversion of control variables (streamfunction, unbalanced velocity potential,

unbalanced temperature, relative humidity, and unbalanced pressure) to model variables (u, v, t q, p_s) increment. Further details are available in Barker et al. (2003 and 2004).

Figure 2.2 shows the standard deviation of the background errors as a function of the vertical level. It indicated a general increase in background errors in summer (July-August 2006) relative to autumn (September-October 2006). It implies that seasonal error tuning factors for background error is necessary. On the other hand, the AWS observation errors used in the WRF 3DVAR system at the low level is 1.1 m s^{-1}, 1.1 m s^{-1}, 2.0 K and 10 % for zonal-, meridional-wind, temperature and relative humidity, respectively. Generally for the short range analysis, the background error was smaller than the observation error (Jarvinen, 2001). In the NMC method, the values of background errors were larger than the observation errors in the wind field, which might leads over-fitting of surface observations.

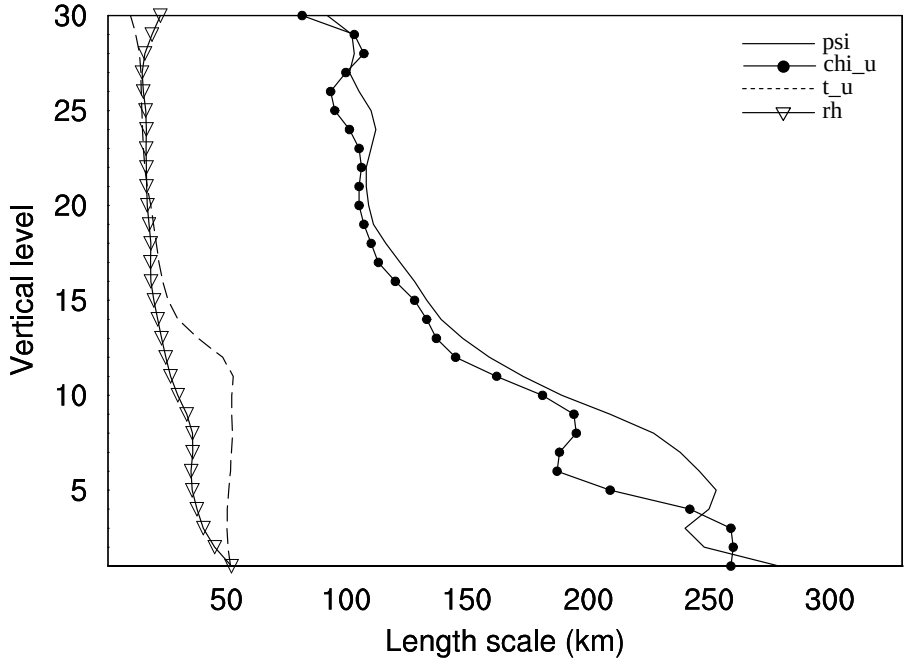

Figure 2.1 Length scale of background error for control variables.

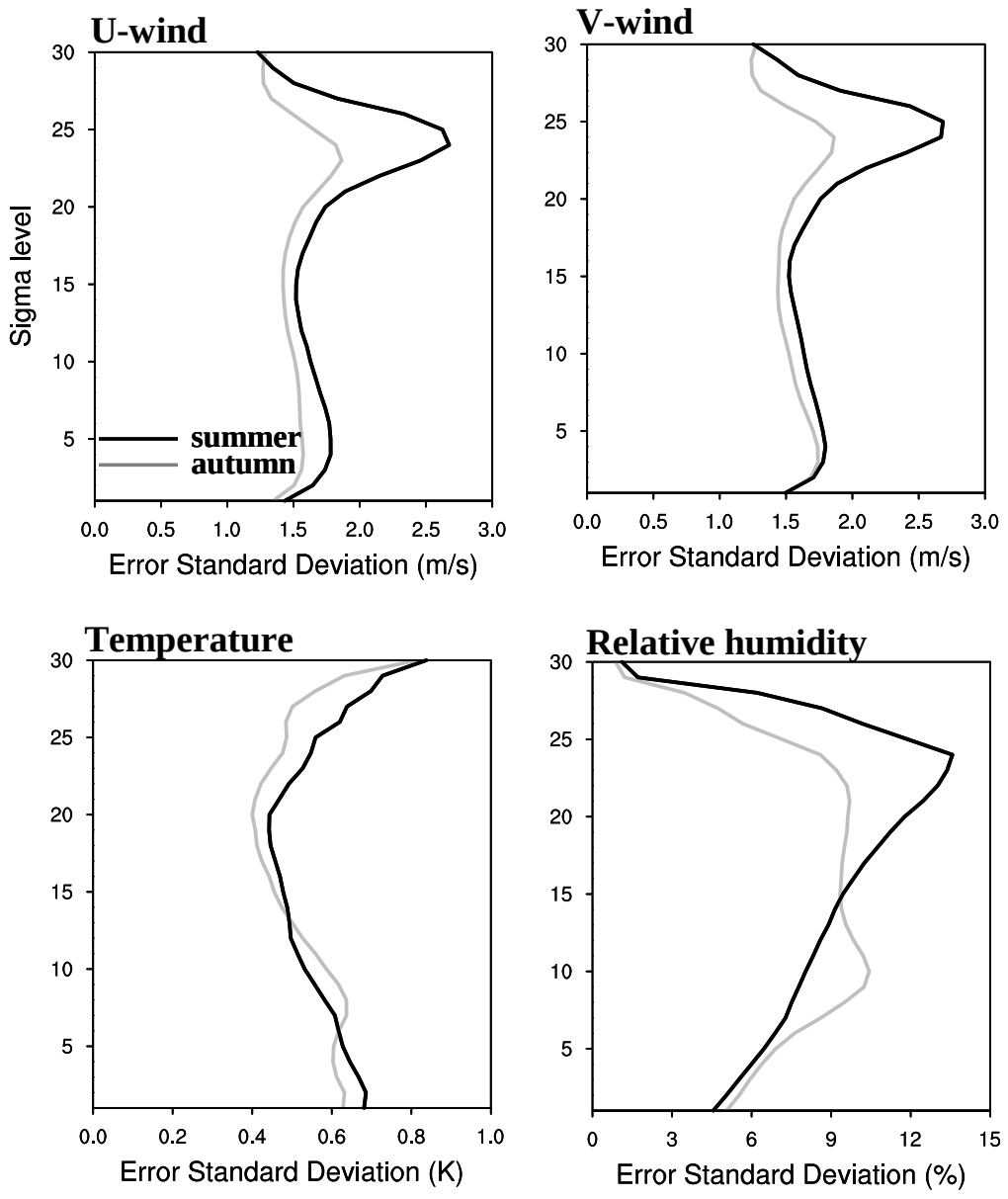

Figure 2.2 Vertical distribution of estimated background error from the NMC method for zonal, meridional-wind, temperature and relative humidity.

2.2 Tuning of the background error length scale for AWS surface data

In order to find an optimal background error statistics, one important thing we can do is to estimate the recursive filter length scale because the length scale of the control variable determines the shape and extent to which observed information is spread, resulting in significant impact on the analysis of WRF 3DVAR (Lee et al. 2010). In addition, diagnostics of the length scale of background error are often used as an approximate indicator of the degree of spatial smoothing. As mentioned above, the length scales were estimated using the NMC method's accumulated forecast difference data processed as a function of grid point separation. Figure 2.3 shows the WRF 3DVAR analysis increment in response to a single surface data point with the length scale calculated by the NMC method. The single observation test is an efficient way to determine how the observed information propagates to its vicinity via the established correlation among the 3DVAR variables. We discuss the results of a single AWS observation test using the WRF forecast as the first guess and assimilating a single surface observation at (36.5 °N, 127.5 °E; 126.0 m). The u and v increments (Figs. 2.3a and b) indicated cyclonic circulation on the southwest side of the AWS data, and anticyclonic circulation on the northeast side of the AWS data. Over the cyclonic region, there was a positive vertical velocity, and vice versa, but with a much smaller length scale than for the horizontal wind (Fig. 2.3c). It was also noted that these

patterns came mainly from the large-scale response, which should typically be maintained in the mesocale system. Because of the WRF 3DVAR multivariate nature, the balance used in the 3DVAR increments could be used to explain why the cyclonic wind increments were produced with negative temperature increments (Fig. 2.3d). When an air column has temperature decreased (negative increments), it will reduce the depth of the air column and the isobaric surface will fall in the column. The geostrophic balance embedded in the 3DVAR increments will force the wind to produce cyclonic increments. The temperature response to the single AWS data assimilation was small. In addition, the influence radius of increments was large (approximately 300-500 km) considering AWS data density (~18 km). This implied that using the length scale calculated by the NMC method would broaden and exaggerated the impact of the surface observations.

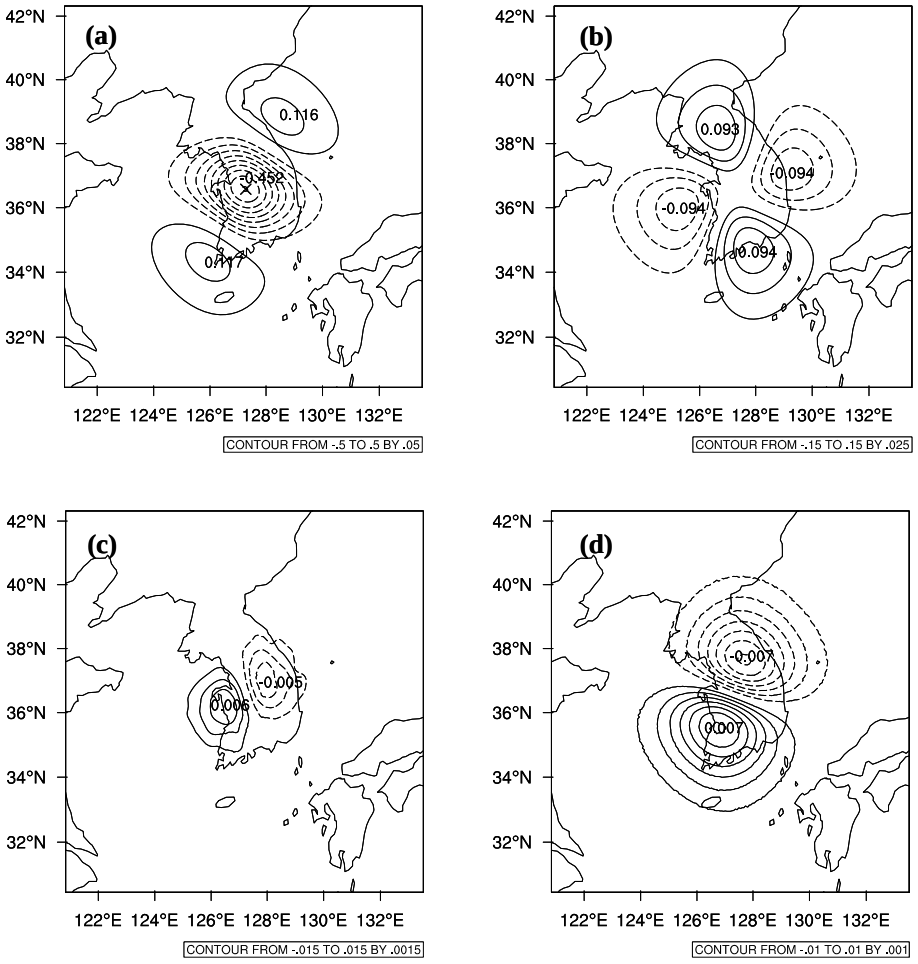

Figure 2.3 The WRF 3DVAR analysis incremental response to a single observation of AWS at 36.5 °N and 127.5 °E (126.0 m): (a) zonal wind, (b) meridional wind, (c) vertical wind, and (d) temperature. [Positive (negative) isoline is solid (dashed), and the zero isoline is omitted.] The cross in (a) denotes the location of the single observation point.

To estimate the length scale of the recursive filter for use of the high-resolution surface data, we compared the correlations from the recursive filter, with a length scale via the NMC method, to the averaged correlation from the difference between the observations and background (hereafter referred to as O-B), binned as a function of station separation. Hollingsworth and Lönnberg (1986) used the averaged correlation of O-B to estimate climatological characteristics of observation and background error. Hollingsworth and Lönnberg examined the statistics of O-B associated with radiosonde observations over North America. By assuming that observation errors for radiosonde were spatially uncorrelated, they were able to assign the spatial correlation of the O-B exclusively to background error. Binning the O-B statistics as a function of the distance between pairs of observations could estimate the spatial correlation of background error as a function of distance. Furthermore, by extrapolating the covariance of the innovations to zero separation, they were able to estimate the relative contribution of background and observation errors to the variance of the innovations. The theoretical basis of estimating the length scales of the surface data is that the O-B correlations represent spatial characteristics of observations when the number of data is sufficiently large. In addition, the appropriate length scale for the high-resolution surface data could be gained by the Gaussian function of O-B data,

$$c(\delta x) = \exp(-\frac{\delta x^2}{2s^2}) \tag{7}$$

where $c(\delta x)$ is the O-B covariance at distance δx between grid points. Eq. (7) can be inverted to extract the appropriate length scale for the surface data. In this study, we used AWS observation data from the warm season (July–August 2006). The AWS observations contain wind speed and direction, temperature, pressure, relative humidity, and hourly rainfall. The temperature, wind, and rainfall are the primary measurements, with pressure and humidity less frequently reported.

Figure 2.4a shows the averaged correlation of O-B for zonal wind and temperature. The shape of the correlations was similar; however, the averaged correlation for temperature was larger than that for winds. Figure 2.4b shows a comparison between the averaged correlation of O-B and the lowest level correlations from the recursive filter of several length scales with tuning factors of 1, 0.7, 0.5, and 0.3 for zonal wind. As mentioned above, the NMC method tends to exaggerate the spatial correlation of winds (Lee et al., 2006). Therefore, this study focuses on tuning the length scale for winds. The length scale from the NMC method was considerably larger than that from the O-B data, which was consistent with the results of the single observation test in Fig. 2.3. This implies that using untuned NMC method's length scale in the surface data assimilation would result in erroneous analysis increment. The result showed that a localized length scale correlation was more consistent with the O-B correlation. The correlation from the recursive filter with a length scale calculated by the Gaussian function indicated that the length scale from the NMC method should be halved for the

assimilation of the surface data. However, even with the half of the length scale, the shape of the correlation from the recursive filter was different from the O-B correlation. The correlation from the recursive filter dropped off faster than that from the O-B at distances near the observation point, and vice versa. A more sophisticated formulation of background error is thus required to fit the data more closely. A Cressman-type objective analysis method (Cressman, 1959) used the different influence radii in a successive correction to account for the different scales of the analysis. Wu et al. (2002) used two different scales in its recursive filter to represent a fat-tailed spectrum for background error in order to improve the analysis on the smaller scales. However, the WRF 3DVAR system did not consider multiple scales in the application of the recursive filter. Hence, to solve the problem of ill-fitting data, a double iteration with two different length scales was applied. In this study, the process of the double iteration with two different length scales progressed as follows:

1) Compute the innovation vectors $d = y_o - H(x_b)$

2) Find the analysis increments $x'^{a}(t_0)$ that minimize the cost function with the large length scale of the recursive filter

3) Update the first guess $x'(t_0) = x(t_0) + x'^{a}(t_0)$

4) Using the updated first guess, repeat steps 1–2 with the small length scale of the recursive filter in order to represent the correlation at distances near the observation point

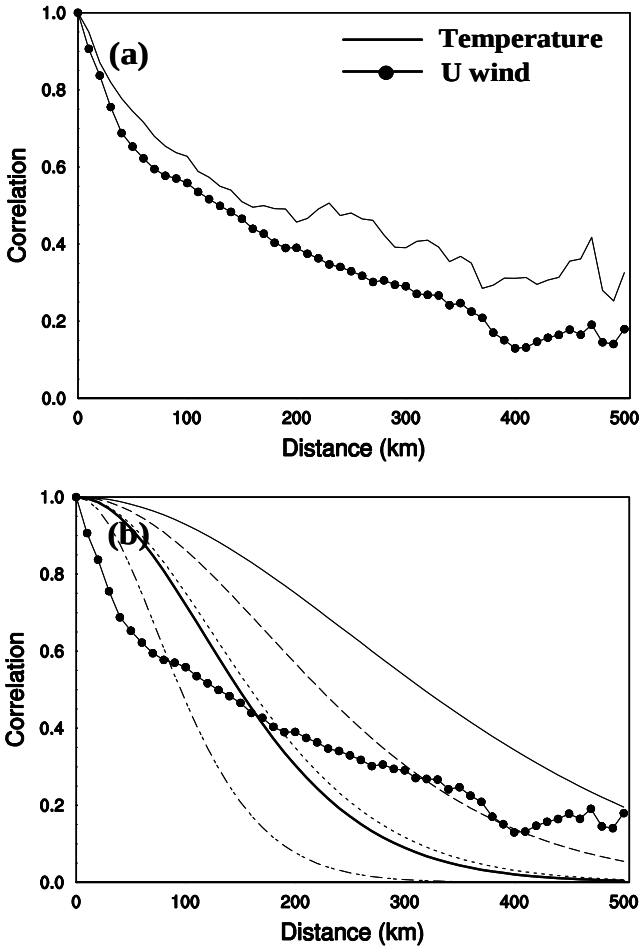

Figure 2.4 (a) Correlations from O-B for temperature (solid) and zonal wind (solid line with filled circles), and (b) correlation from O-B (solid line with filled circles) and from the recursive filter of the length scale with tuning factors of 1 (solid line), 0.7 (dashed line), 0.5 (dotted line), and 0.3 (dash-dotted line) for zonal wind. The thick solid line in (b) indicates the correlation from the recursive filter with a length scale calculated by the Gaussian function.

5) Advance the WRF forecast to the following analysis time to prepare the background for the next 3DVAR analysis

Steps 1–5 were repeated throughout the assimilation procedure. The idea here is to exploit the convergence on large scales during the first iteration of the minimization (Thépaut and Courtier 1991; Tanguay et al. 1995), and to allow convergence on the smaller scale during the next iteration. In this manner, both the large- and small-scale structures represented in the surface data were incorporated to the model field, and therefore the shape of the correlation seemed to approximate the features of the O-B correlation.

To select the length scales representing large and small length scales in the double iteration experiment, we calculated the root-mean-square errors (RMSEs) of correlations from the recursive filter of the length scales with tuning factors from 0.1 to 1 versus the O-B correlation. Considering the MCSs case and the density of the AWS data, the large scale was defined as 200–300 km, approximately corresponding to the meso-α scale, and the small scale was defined as 20–100 km, which was roughly consistent with the meso-β scale. The statistics indicated that for a heavy rainfall case associated with MCSs and the background error in the summer month of July 2006, the correlation from the recursive filter with a tuning factor of 0.7 was close to the O-B correlation in the large-scale, while in the small-scale, the correlation from the recursive filter with a tuning factor of 0.3 showed the best result. As a result, a length scale with a tuning factor of 0.7 was selected as the large length

scale in the first iteration of the assimilation, and a tuning factor of 0.3 was used for the small length scale in the second iteration.

Chapter 3 Data

The data used in this study were taken from Japan Meteorological Agency analysis charts, National Center for Environmental Prediction (NECP) global final (FNL) analyses with a 1°×1° horizontal resolution and 26 vertical layers from 1000 to 10 hPa, Korean Meteorological Administration (KMA) surface observations, multi-functional transport satellite (MTSAT) enhanced IR images, and rawinsonde data at Osan. The data from Osan weather station were analyzed by the complete Rawinsonde Observation Program, version 5.5 (RAOB55; ERS 2004). The RAOB55 allowed analysis of skew T and calculation of thermodynamic indices and wind data, which helped to infer convective potential, storm type, and movement.

As shown in Fig.1.1, there are a total of 18 radars comprising the radar network over the Korean Peninsula: eleven KMA radars, five Republic of Korean Air Force (KAF) radars, and two WSR-88D (USAF) radars. The KAF radars operate at a wavelength near 5.5 cm, while the KMA radars operate at wavelengths near 5.5 or 10 cm. These radars are approximately 120 km apart on average, with the observable range for each exceeding 100 km; hence, most areas over the southern Korean Peninsula are overlapped by the coverage of two or more radars. These characteristics of the Korean operational radar network allow the coverage of the entire southern Korean Peninsula. The radar data are observed at intervals of 6-10 min. In contrast to the networks of the

European countries and Japan, all of the operational radars throughout the Southern Korean Peninsula have Doppler capabilities and thus they can measure the radar reflectivity factor (Z), radial velocity (V_r), and spectrum width (SWD). The radial velocity contains information of vertical motion; the vertical velocity is sometimes important for the convective initiation and forecasting. The reflectivity is a measurement of the precipitation hydrometeors (rain, snow, etc). In this study, the preprocessing of radar data was based on a method proposed by Park and Lee (2009). The aliasing effect of the radial velocity measurements was corrected for under a spatial continuity constraint along the radial and azimuth directions. The radial velocities observed by radar were first processed along the radial direction at each azimuth ray, starting from the gate adjacent to the radar at which the observed velocity was assumed to be a true value unaffected by the aliasing effect. After terminating at all the azimuth rays, the correction procedure was then performed along the azimuthal direction. Then, winds derived from a simplified velocity-azimuth display (VAD) method were used as the reference values for determining whether the radial velocities were aliased or not. In addition, the speckling due to the dual–pulse repetition frequency (PRF) velocity error was removed. When the difference from the mean value within an area surrounded by 15 gates along the radial direction and seven rays along the azimuth direction centered at the given gate was greater than 10 m s^{-1}, the radial velocity at the given gate was considered to be noise and therefore removed. Then, the Doppler

radar data were then extracted by the Custom Editing and Display of Reduced Information in Cartesian space (CEDRIC) method and the Sorted Position Radar Interpolation (SPRINT) package provided by National Center for Atmospheric Research (NCAR). The vertical profiles of radar data used this study had a resolution of 0–12 km with 0.5 km vertical grid intervals. However, information below the 2 km height was limited due to complex topography over the Korean Peninsula. In order to investigate the characteristics of MCSs, the horizontal grid of 1 km radar data from 5 radar sites over the central Korean Peninsula was used, while, that of 5 km radar data from 18 radar sites was used for the radar data assimilation. The radial velocity and reflectivity observation errors used in the data assimilation were 2 m s^{-1} and 5 dBZ, respectively.

Approximately 550 AWS data points with an average horizontal resolution of 18 km were used to investigate the amount of hourly accumulated rainfall, and variations in surface pressure and temperature. For the surface data assimilation, AWS data were preprocessed using WRF 3DVAR observation preprocessor. Because observation errors can be introduced at all stages including measurement, reporting practices, transmission and decoding, it is important that quality control is performed to avoid the assimilation of erroneous observations (Gu et al. 2005). A number of quality control were performed including the removal of observations outside the time range and domain, re-ordering and merging the duplicate data reports in time location, retrieving the pressure or height based on the observation information with the

hydrostatic assumption, ensuring the vertical consistency, and estimating observational error based on the pre-specified errors. To make use of AWS data for the assimilation, when the elevation of an AWS station was above the height of the model's lowest level or below the model's lowest level but not more than 100 m, the observations were used for the surface data assimilation. However, the observation at the station was discarded if the elevation of an observed site was lower than the height of the model's lowest level by more than 100 m.

Chapter 4 Heavy rainfall case

The heavy rainfall case associated with the MCSs, which resulted in the total rainfall amount greater than 330 mm on 11-12 July 2006 in the Goyang area over the central Korean Peninsula with the band-shaped structure, was chosen to investigate the impact of not only the radar and surface data assimilation on heavy rainfall forecast but also the tuned length scale of recursive filter in background error on assimilating the surface data for a heavy rainfall case. One of the reasons for studying this case is that the operational forecasts failed to predict the amount of the precipitation. This was because the convective system responsible for the localized heavy rainfall developed over a period of just a few hours.

4.1 Synoptic background

Figure 4.1 shows the distribution of the accumulated 12 h and hourly rainfall amounts from 2100 UTC 11 July to 0900 UTC 12 July. The area with intense rainfall greater than 210 mm was localized within the band-shaped structure. The maximum rainfall amount over 12 h was 335.0 mm, with a maximum intensity of 77.5 mm h^{-1} at 2300 UTC 11 July. The heavy rainfall was mainly concentrated from 2200 UTC 11 July to 0100 UTC 12 July at Goyang, and then there was a 2 h break-off period from 0300 UTC to 0500 UTC 12 July. The heavy rainfall with 30 mm h^{-1} intensity occurred at 0600 UTC 12 July again, and then the rainfall was almost stopped. In

order to examine in detail the evolution of the heavy rainfall, we investigated the 15 min rainfall amount at Goyang station from 2200 UTC 11 July to 0100 UTC 12 July (not shown here). It showed the first peak at 2215 UTC 11 July with a rainfall intensity of 20 mm $(15\ \text{min})^{-1}$. After a temporary break for a few minutes, the maximum intensity reached 32 mm $(15\ \text{min})^{-1}$ at 2250 UTC 11 July. Then, the rainfall intensity over 25 mm $(15\ \text{min})^{-1}$ was maintained from 2315 UTC to 2330 UTC 11 July and then the rainfall was dissipated.

Similar to most heavy rainfall cases that have occurred in the Korean Peninsula, this case had a synoptic environment that was favorable for heavy rainfall (Lee et al., 1998), except for some discrepancies. Figure 4.2 shows the surface and 850 hPa weather charts at 1800 UTC 11 July, which is 3 h before the heavy rainfall begin. The *Changma* front over the central Korean Peninsula extended from a low-pressure center located in the northern East Sea and Typhoon Bilis (2006) was located over the South China Sea (Fig. 4.2a). The Korean Peninsula was on the edge of a high-pressure system extending westward over the Pacific Ocean between roughly 30° and 40°N. A belt of relatively strong southwesterly winds, or low-level jets (LLJs), developed along the high-pressure systems (Fig. 4.2b). In addition, the area of equivalent potential temperature over 345 K was extended to the heavy rainfall region. In this environment, the strong southwesterly flow continuously transported a large amount of warm and moist air to the central Korean Peninsula so that the relative humidity there was greater than 90 %. The LLJ was oriented almost

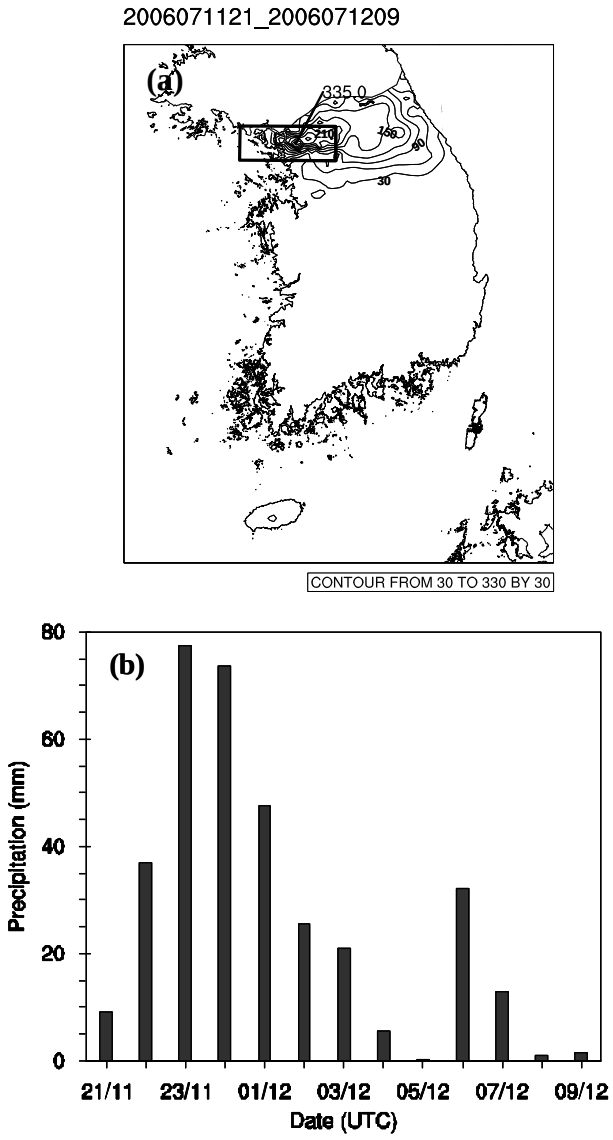

Figure 4.1 (a) Observed rainfall amount (mm) accumulated from 2100 UTC 11 July 2006 to 0900 UTC 12 July 2006. (b) Hourly precipitation (mm) at Goyang station. The black box in (a) denotes the analyzed area in Fig. 4.7.

perpendicular to a pronounced horizontal thermal gradient that increased baroclinicity. LLJs are known to slope upward when they overrun a lower layer of cooler air, which is usually present in large-scale MCS environments (Trier and Parsons, 1993; Augustine and Caracena, 1994; Rochette and Moore, 1996). At 0000 UTC 12 July, the LLJ was located in the vicinity of the heavy rainfall region because Typhoon Bilis pushed the high-pressure system further north to intensify the southwesterly flow. The upper-level synoptic pattern exhibited almost the same features as the low-level synoptic pattern. In the 500 hPa chart, the upper-level trough and the synoptic low pressure little influenced the Korean Peninsula, except for the low pressure system over the northern China (not shown here). Therefore, the cold air by the trough had little impact on the development and maintenance of the convective systems. In the 1000 hPa divergence and 850 hPa moisture advection fields at 1800 UTC 11 July, low-level convergence and moisture advection was dominant over the central Korean Peninsula, while divergence and dry advection was dominant over the northern Korean Peninsula (Fig. 4.3). This means that in this case, the MCSs strongly depended upon the development of a synoptic-scale system. The synoptic overview indicates that the heavy rainfall over Goyang could be attributed to the enhancement of convective activity over the *Changma* front.

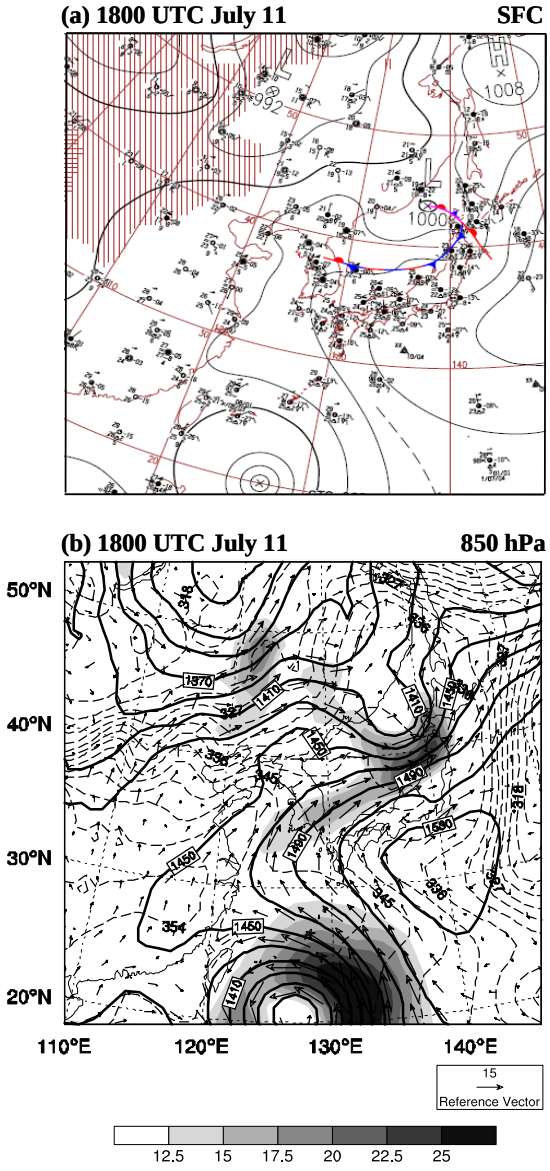

Figure 4.2 (a) Surface weather chart and (b) geopotential height (gpm, solid), equivalent potential temperature (K, dashed), wind speed (m s^{-1}, shaded) and wind vectors (m s^{-1}) for the 850 hPa weather chart at 1800 UTC 11 July.

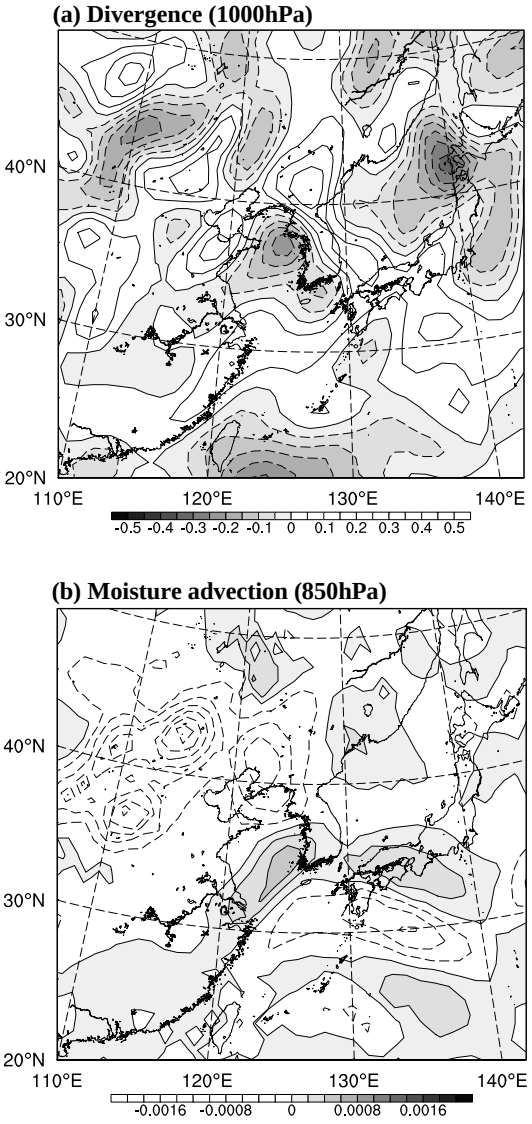

Figure 4.3 FNL analysis chart for (a) 1000 hPa divergence ($10^{-4}\,s^{-1}$, negative areas are shaded) and (b) moisture advection at 850 hPa (10^{-4} $kg\,kg^{-1}\,s^{-1}$, positive areas are shaded) at 1800 UTC 11 July.

Figure 4.4 shows MTSAT IR images from 2100-2300 UTC 11 July at 1 h intervals. MTSAT IR image clearly showed the development of MCS affecting the heavy rainfall at Goyang. An isolated storm appeared near the west coast of the central Korean Peninsula at 2100 UTC and moved eastward while developing quickly from 2200-2300 UTC. The size of the most intensive convection system at 2300 UTC 11 July was approximately 2000 km^2, which corresponded to the meso-β scale. This storm stagnated around Goyang for about 4 h from 2100 UTC 11 July to 0100 UTC 12 July. After 0100 UTC, the storm started to decay and began moving out of Goyang.

To investigate the storm environment in more detail, we analyzed sounding data from Osan, which is approximately 60 km south of Goyang, at 1800 UTC 11 July and 0000 UTC 12 July (Fig. 4.5). At 1800 UTC 11 July, moist layers were strengthened between 450 and 350 hPa and below 600 hPa. A dry layer was confined between the two moist layers. At 0000 UTC 12 July, the low level was almost saturated, while the dry layer became weaker. A lifting condensation level (LCL) appeared near the surface at 1800 UTC 11 July and 0000 UTC 12 July because the air there was almost saturated. The level of free convection (LFC) at 621 m at 1800 UTC 11 July was relatively low. Thus, only a small lifting force was needed to release its potential instability. A surface southeasterly of 2.5 m s^{-1} veered to the westerly of about 10 m s^{-1} at 500 hPa at 1800 UTC 11 July. At 0000 UTC 12 July, wind speed in the low level between 1000 and 700 hPa increased due to the increase of the LLJ; at the same time,

there was relatively weak midlevel flow. Consequently, a strong vertical wind shear occurred in the low level, and the vertical wind shear in the midlevel was relatively weak. The hodographs during the target period also showed that there was veering in the lower troposphere, that is, warm advection was prominent from the southern or southwestern region (not shown here). Weisman and Klemp (1982) used the bulk Richardson number (BRN), which is the ratio of the total energy available due to buoyancy to the total energy available from vertical wind shear, as nondimensional parameter and concluded that supercells form and are stably maintained when $5 \leq \text{BRN} \leq 50$; high values of BRN favor the development of multicell storms. In this case, the BRN was 7 at 1800 UTC 11 July and decreased to 5 at 0000 UTC 12 July. The estimated BRN value indicated that the environment in this case would support the formation of supercells. Therefore, the storm had the potential to develop into a supercell-type storm from 1800 UTC 11 July to 0000 UTC 12 July. However, the convective available potential energy (CAPE) maintained a value of 123 J kg^{-1} at 1800 UTC 11 July and 136 J kg^{-1} at 0000 UTC 12 July. This value was smaller than that of operational thresholds (NWS, 2004) and case studies (e.g. Stensrud et al., 1997) in the United States. Hong (2004) showed that the atmosphere over the Korean Peninsula during the summer season was identified as thermodynamically neutral when compared to the United States by investigating the summer season climatology of the synoptic environment during from 1979 to 2002. The environment in the United States during the summer season was

characterized by the large CAPE value. While, the atmosphere during the summer season over the Korean Peninsula was affected by baroclinicity.

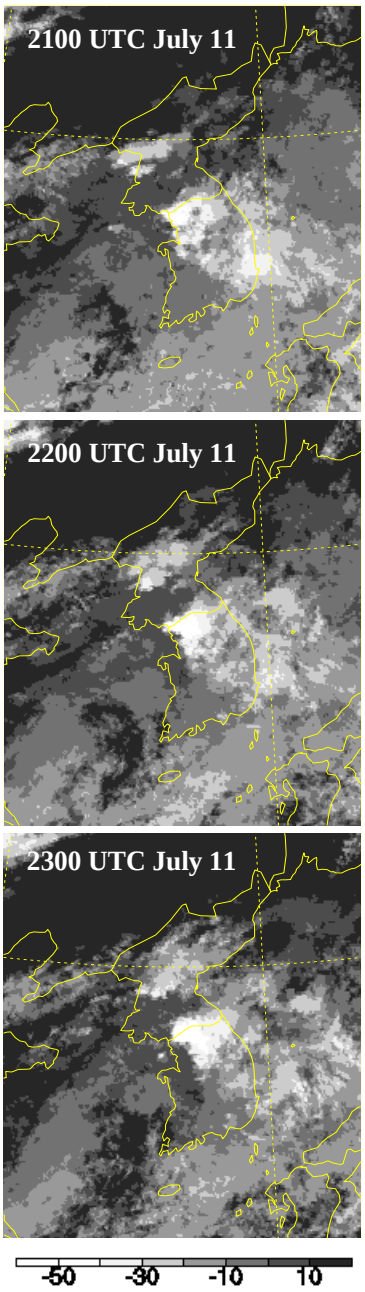

Figure 4.4 MTSAT IR satellite images from 2100 UTC 11 July to 2300 UTC 11 July at intervals of 1 h.

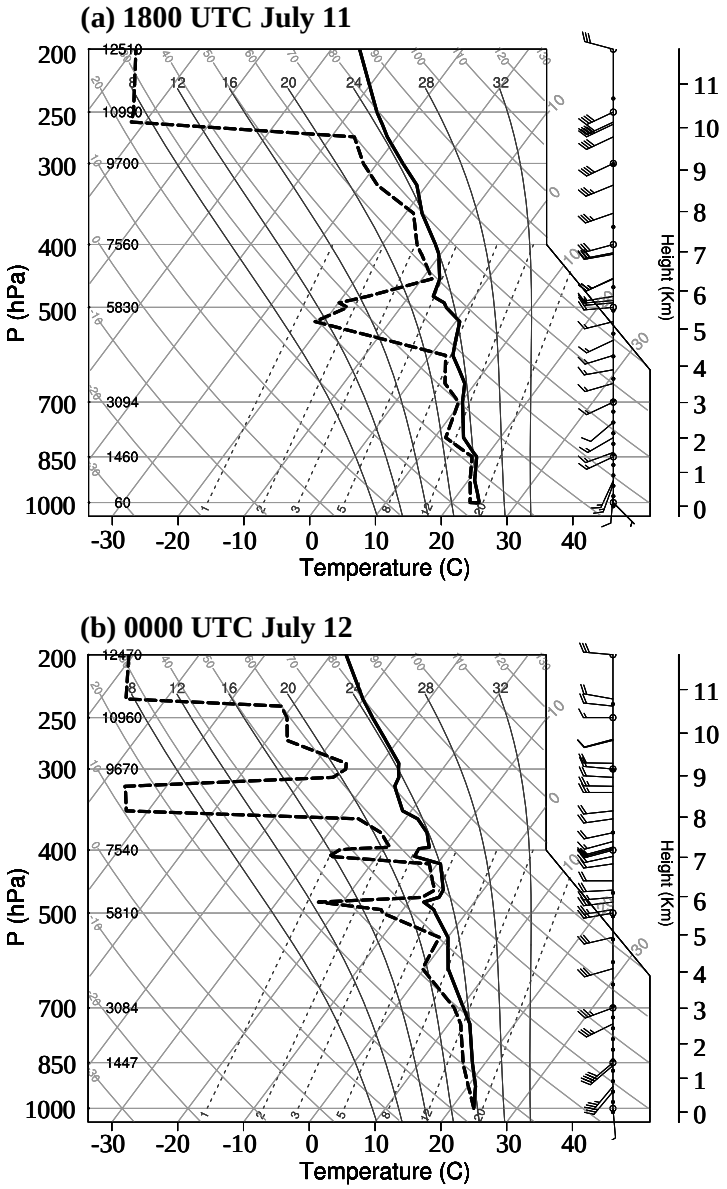

Figure 4.5 Skew T-log P diagram at Osan station at (a) 1800 UTC July 11 and (b) 0000 UTC July 12.

4.2 Mesoscale features

Figure 4.6 shows surface analyses at 2100 UTC 11 July and differences in the fields of surface analyses at 0000 UTC 12 July compared to 2100 UTC 11 July. The storm initiated at 2100 UTC 11 July and developed to its mature stage at 0000 UTC 12 July. At 2100 UTC 11 July, a surface mesoscale high pressure system was dominant over the central and northern Korean Peninsula, and a surface mesoscale low pressure system was located over the west coast of the central Korean Peninsula. A region of cold air collocated with the mesoscale high and mesoscale low pressure system located in a region of warm air. The region of cold air extended to the northwest of Goyang, and relatively warm air was dominant south of Goyang. A cold pool persisted in the central Korean Peninsula for many hours from 1200 UTC 11 July. Cold pools frequently play a role in triggering a storm in the initiation phase of successive convective storms (Doswell, 2001). Mesoscale boundaries such as the land-water contrast, rain-no rain areas, saturated-unsaturated soils, and steep horizontal pressure and thermal gradients can provide favorable conditions for storm initiation (Funk, 2004). In this case, thermal gradient between the mesoscale warm and cold air increased the convective instability over the *Changma* front accompanying a low-level convergence and thus provided favorable conditions for initiating storms. At 0000 UTC 12 July, the warm air region extended toward the heavy rainfall region due to the LLJ, and the temperature around Goyang dropped

slightly; however, there was no noticeable change in pressure over the central Korean Peninsula. The surface temperature fell less than 1°C as the storm passed over Goyang.

To investigate the evolution of the storm, the reflectivity, convergence, vertical vorticity and wind vectors of radar data were analyzed at a height of 4 km from 2150-2330 UTC 11 July at 10 min intervals (Fig. 4.7). The arrows plotted in Fig. 4.7 represent the movements of convective cells. The reason that the level of 4 km was selected is that the radar data could present the convergence and vorticity fields well. From 2100-2150 UTC, the storm initiated near the western coast of the central Korean Peninsula and moved northeastward together with the southwesterly flow. At 2150 UTC, a westerly flow occurred in the northern part of the storm, and a southwesterly flow dominated in the southern part of the storm. The speed of the westerly flow was weaker than that of the southwesterly flow. This environment facilitated the formation of a convergence line owing to the horizontal wind shear. Therefore, a linear storm along the convergence line developed that was aligned in the east-west direction. At the same time, a negative vorticity appeared in the middle of the vorticity field, and the strong echo region was separated into two parts at 2200 UTC. From 2150-2230 UTC, the antecedent storm M1 and rear storm M2 moved at speed of 16.7 and 8.3 m s^{-1}, respectively; the strong negative vorticity between storms M1 and M2 seemed to inhibit the propagation of storm M2. As a

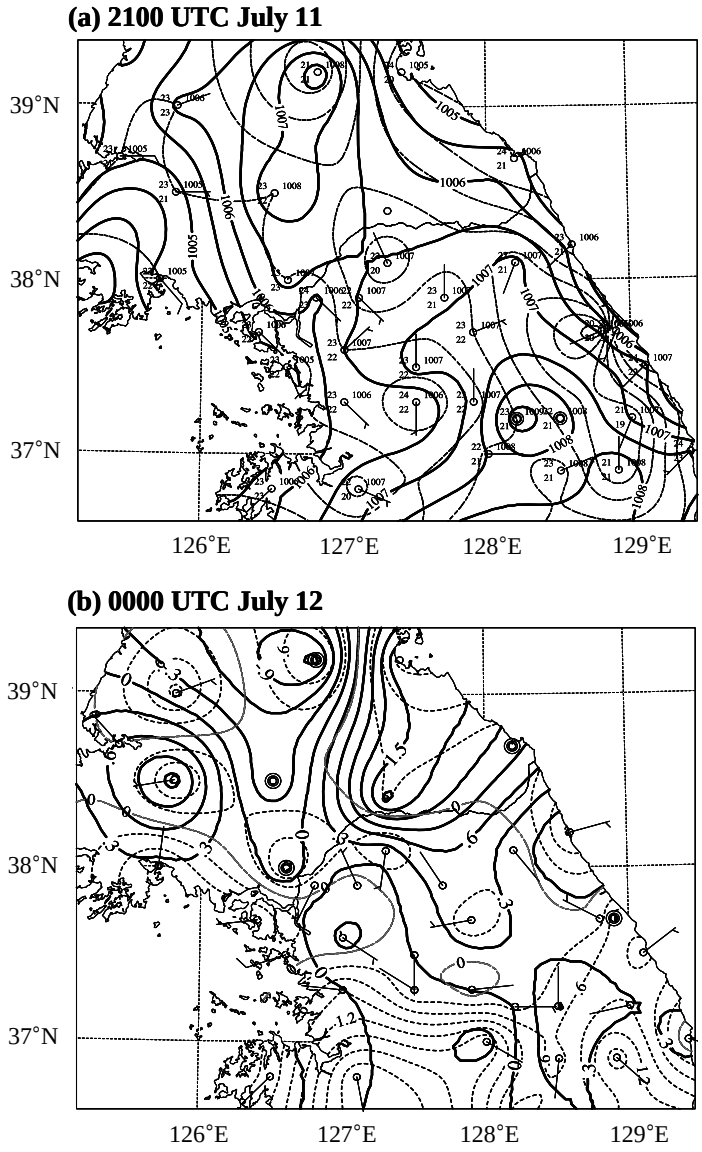

Figure 4.6 (a) Surface observation at 2100 UTC 11 July and (b) the difference in the mesoanalysis chart at 0000 UTC 12 July compared to 2100 UTC 11 July. Solid lines denote the pressure (hPa), and dotted lines denote temperature (°C).

result, storm M1 moved eastward faster than storm M2, and the two storms were then further divided. The small convective cell M3 developed near the western coast at 2210 UTC, moved eastward while developing, and merged into the leading storm M2. Therefore, at 2220 UTC, the reflectivity of storm M4, which caused heavy rainfall over Goyang, became stronger as it accompanied the intensified convergence and vorticity. Storm M4 stagnated around Goyang for 30 min from 2220-2250 UTC and moved out of Goyang. At 2300 UTC, the ambient flow of the storm changed from westerly to southwesterly. As a result, the convergence line was weak and moved downstream; the linear storm then began to spread. After 2300 UTC, storms repeatedly initiated on the coastal area, moved northeastward relatively rapidly, and developed continuously around Goyang due to repeated replacement of the storms. The propagation of the convective system showed clear evolution on the development of back-building MCSs, such as stagnation of the entire convective system oriented in the east-west direction. Bluestein and Jain (1985) noted that the back-building environment was very similar to that of supercells and that back-building occurred in environments of strong vertical shear, large CAPE, and small BRN. In this study, the storm development environment was consistent with that used by Bluestein and Jain, except for the relatively low CAPE values.

The maintenance and stagnation of the storms were closely related to the strength of convergence, which was associated with the upward motion, induced by the westerly flow in the northern part of the storm

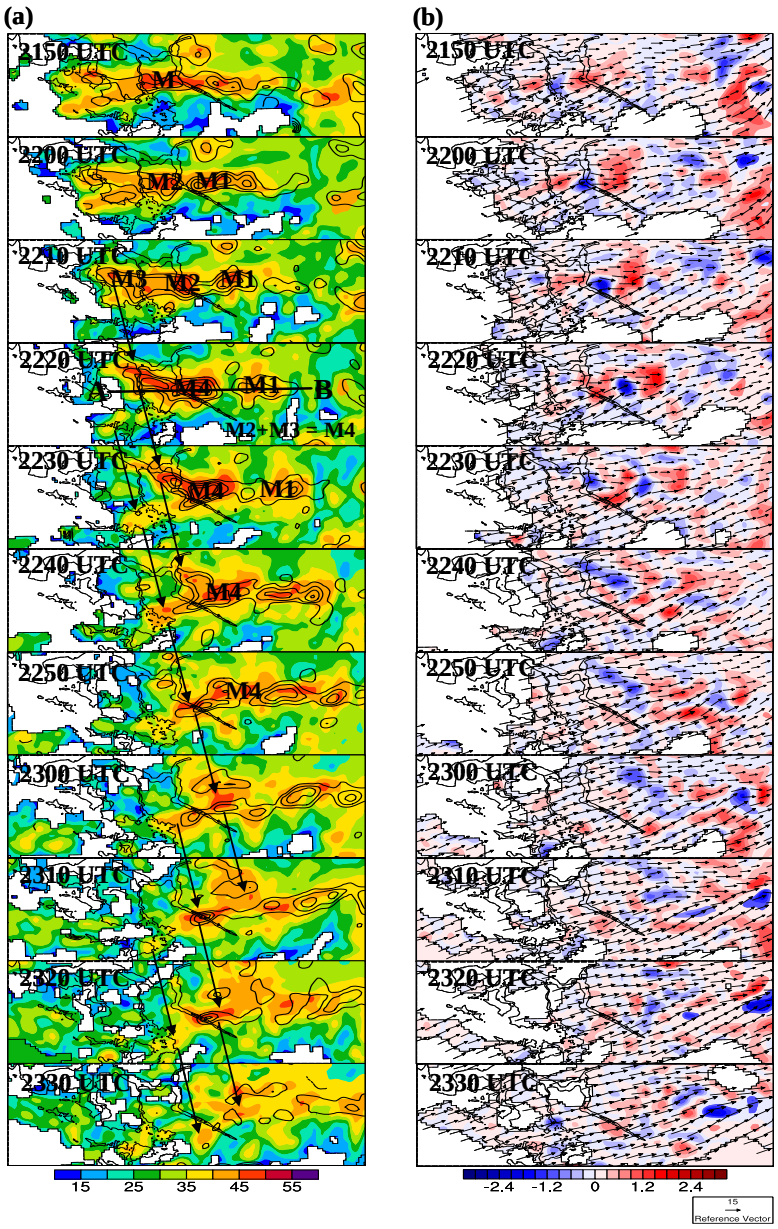

Figure 4.7 (a) Reflectivity (dBZ, shaded) and convergence (10^{-4} s^{-1}, lines) and (b) vertical vorticity (10^{-4} s^{-1}, shaded) and wind vectors (m s^{-1}) at the 4 km height from 2150–2330 UTC 11 July. Arrows indicate the movement of the convective cells.

and the strong southwesterly flow in the southern part of the storm. When the southwesterly flow dominated around and within the storm after 2300 UTC 11 July, the storms around Goyang were maintained by continuous development of new cells because the strong southwesterly flow provided warm and moist air to the heavy rainfall region. Therefore, the strong southwesterly flow, or LLJ, was an important factor in maintaining storm development.

Figure 4.8 shows the vertical cross section of reflectivity, wind vectors, vertical velocity and divergence along line A-B in Fig. 4.7 from 2150-2250 UTC at 10 min intervals. From 2150-2200 UTC, the storm developed and propagated eastward at a speed of 18.3 m s^{-1}, and the strong reflectivity corresponded well to the strong convergence areas. From 2200-2220 UTC, the storm M1 and M2 moved slowly eastward at 12.5 and 6.7 m s^{-1}, respectively, because these storms were blocked by the mountains. Along the east-west direction around Goyang, topography surrounds the eastern side with mountains whereas the western side is exposed. Topography played a role in preventing convective storms from moving eastward, which restrained the storms to within the area and this resulted in a heavy rainfall intensity of approximately 20 mm $(15\ min)^{-1}$ over Goyang by storm M1. A downdraft due to the heavy rainfall appeared at the center of storm M1, which corresponded well to the strong negative vorticity, as shown in Fig. 4.7. The downdraft provided the cold and dry air at the leading edge of storm M2. This environment not only facilitated the development of storm M2 due to enhancing

convective instability but also inhibited the propagation of storm M2 due to the downdraft. Consequently, a new storm (M3) merged into storm M2 to build a large convective storm (M4) with a 20 km width of strong reflectivity (>45 dBZ). From 2230-2250 UTC, storm M4 moved eastward and produced a heavy rainfall intensity of approximately 30 mm (15 min)$^{-1}$ over Goyang, which corresponded well to the peak of precipitation (Fig. 4.1). Analysis indicated that the topography seemed to play an important role in blocking the eastward movement of the convective storms, which helped merge storms and resulted in the heavy rainfall over Goyang. A convergence region was observed to the west of the strong reflectivity; this was consistent with the upstream convective development, especially the back-building system. The convergence field was tilted to the east, and the divergence region below the convergence region was collocated with the downdraft. The upper-level divergence maximum was directly above the core of the reflectivity, which was characterized by the back-building MCSs (Schumacher and Johnson, 2005) and this implied that the air was lifted from a low level through an upright pattern rather than a tilted pattern in the back-building MCSs. Another important feature in the storms was that the downdraft was approximately 2 m s^{-1}, corresponding to the weak cold pool shown in Fig. 4.6. It is known that a strong downdraft in a convective cell causes a strong outflow, which cuts the warm and moist air supply to an updraft of a convective cell (Weisman and Klemp, 1982). In this study, the weak downdraft did not cut off the southwesterly flow and therefore new

storms developed continuously because the southwesterly flow provided moist air from the regions of high equivalent potential temperature to the convective system, which was a favorable condition for heavy rainfall.

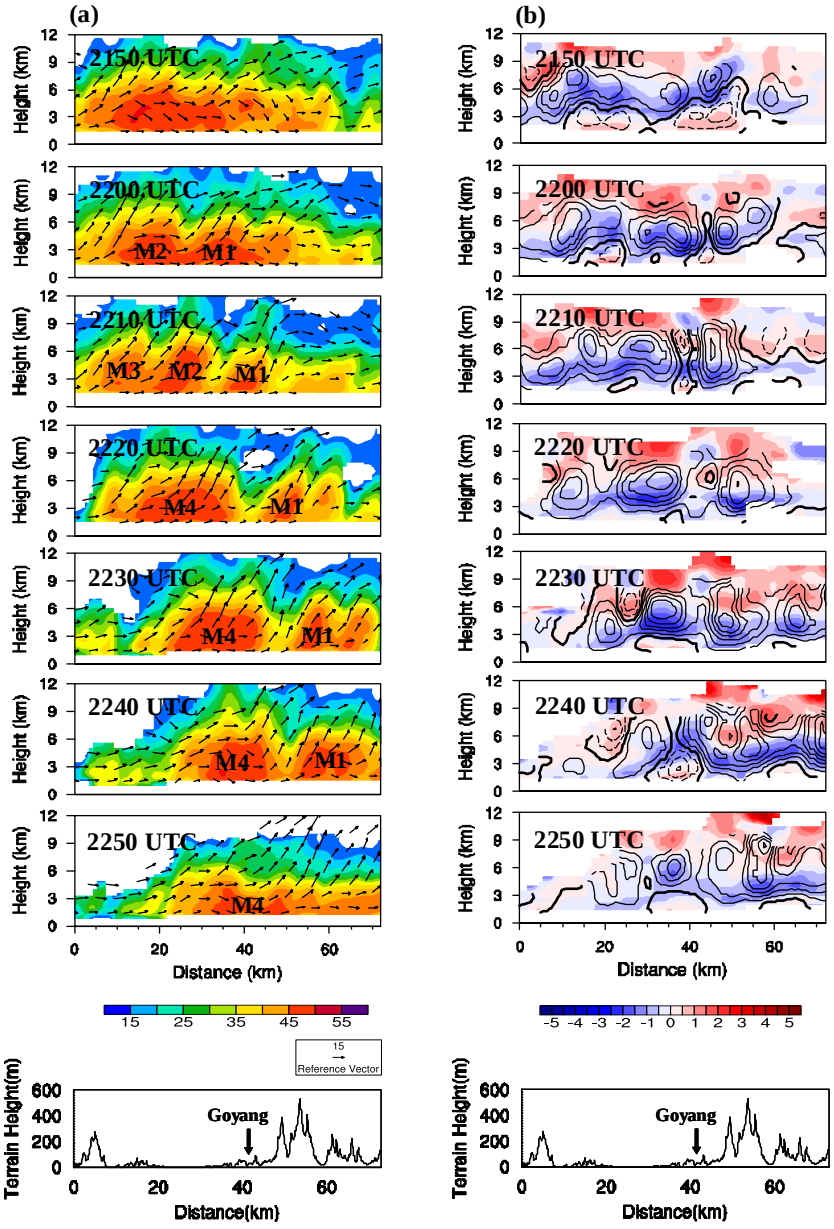

Figure 4.8 (a) Vertical cross section of reflectivity (dBZ, shaded) and wind vectors (m s^{-1}) and (b) divergence (10^{-4} s^{-1}, shaded) and vertical velocity (m s^{-1}, positive (negative) values solid (dashed)) from 2150–2250 UTC 11 July at 10 min intervals with topographical distribution.

Chapter 5 Numerical results

5.1 Configuration of numerical model

In order to investigate the impact of radar and surface data assimilation and the tuned length scale of recursive filter in the assimilation of surface data for the heavy rainfall forecast, we conducted numerical simulations using WRF and WRF 3DVAR version 3.1. The model domain consisted of 18 km (170 × 150 horizontal grid points) and 6 km (211 × 211 horizontal grid points) grid domains (Fig. 5.1); this was a one-way nested grid system with 31 η vertical layers and 50 hPa at the model top. A modified version of Kain-Fritsch cumulus parameterization scheme was used for subgrid scale convection (Kain and Fritsch, 1990, 1993), and the mixed-phase microphysics scheme for WRF single-moment 6-class (WSM6) (Hong et al., 2004) was used for moist processes at the grid-resolvable scale. The YSU scheme, NOAH land-surface model, and rapid radiative transfer model (RRTM) longwave/Dudhia shortwave schemes were used for the planetary boundary layer, land surface, and atmospheric radiation process, respectively. The vertical layers and all physics for the model were the same in all domains, except for cumulus parameterization, which was not used in the 6 km grid domain. Design of domain and model schemes used in this study is summarized in Table 1.

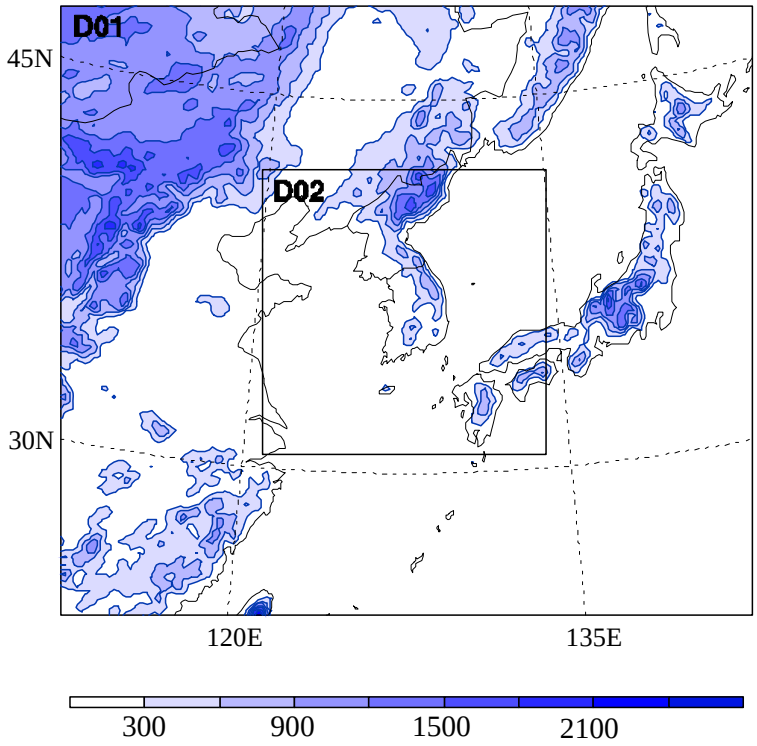

Figure 5.1 Domain configurations for WRF 3DVAR and WRF forecasting experiment and topography (m). Domain 1 (D01) has 170 × 150 grid points with a grid spacing 18 km, and domain 2 (D02) has 211 × 211 grid points with grid spacing of 6 km.

Table 1. Summary of model schemes and design of domain

description	Domain 1 (D01)	Domain 2 (D02)
Horizontal resolution	18 km	6 km
Horizontal grid number	170 × 150	211 × 211
Vertical layers / Model top	31 sigma layers / 50 hPa	
Cumulus parameterization	Kain-Fritsch scheme	No
Explicit moisture	WSM6 scheme	
Boundary layer	YSU scheme	
Longwave radiation	RRTM scheme	
Shortwave radiation	Dudhia scheme	
Surface physics	Thermal diffusion scheme	

5. 2 Radar and surface data assimilation

5.2.1 Experiment design

The initial conditions for the 18 km horizontal resolution coarse grid in the numerical simulations were obtained from FNL analysis data at 6 h intervals. In addition, we assimilated conventional data for the coarse-grid domain to produce more accurate initial and boundary conditions for the 6 km horizontal resolution fine grid using WRF 3DVAR. 3DVAR system has practical advantages; 1) observations can easily be assimilated directly without the need for prior retrieval. This results in a consistent treatment of all observations and, as the observation errors are less correlated, practical simplifications to the analysis algorithm. 2) The 3DVAR system solution is found using all observations simultaneously. 3) Balance constraints can be built into the preconditioning of the cost-function minimization. Even though 3DVAR system has practical advantages, there are still issues associated with spinup and initialization. Linear balance constraints in the variational system are often insufficient for preventing the development of spurious energy on fast timescales in numerical forecasts (Polavarapu et al., 2004). Therefore, a separate filtering procedure is required to remove spurious high-frequency gravity wave noise, which can have a detrimental effect on the first few hours of the forecast, and on the data assimilation cycle as a whole. A number of approaches have been developed to address the initialization

problem including nonlinear normal mode initialization (Daley 1979), damped time-differencing scheme (Baker et al. 1987), diabatic initialization (Puri 1985), digital filter initialization (DFI; Lynch and Huang 1992), nudging technique (Grell et al. 1995), and incremental analysis updates (IAU; Bloom et al. 1996). The U.K. Met Office (UKMO) has used the digital filter and IAU as an alternative initialization technique (Clayton 2003). In this study, we applied the IAU method for data assimilation of the WRF model to reduce the noise of spurious high-frequency gravity waves. By gradually incorporating analysis increments, IAU method removes high frequencies (Lee et al., 2006). In general, numerical weather prediction centers filter analysis increments rather than analysis in order to preserve the physical signal in the background state (Polavarapu et al., 2004). The increments are considered as additional forcing terms in the model equation

$$\frac{dX(t)}{dt} = F(X) + W(X_a - X_b), \tag{8}$$

where $X(t)$, $F(X)$, W, and $X_a - X_b$ are model variables, the original model forcing term, the predefined weighting coefficient, and the analysis increment, respectively. In Eq. (8), if the analysis increment is zero at a specific location, IAU forcing will be zero. This implies that IAU forcing is significant only in regions where observations induce an analysis increment. In this study, the IAU process progressed as follows:

1) WRF 3DVAR generated increments at analysis time. The

increments were transformed into tendencies of the model variables, i.e. winds, temperature, and water vapor mixing ratio, and the results were saved for use in IAU forcing.

2) An earlier model state was recovered for forecasting.
3) WRF advanced a forced forecast to the following analysis time to prepare the background for the next 3DVAR analysis.

Steps 1–3 above were repeated throughout the assimilation. More detailed information on IAU can be found in Bloom et al. (1996). In this study, the coarse-grid simulation was carried out for 21 h forecasts after two update cycles (1200 and 1500 UTC) from 1200 UTC 11 July, in which a 3 h time window from -1.5 h to + 1.5 h centered at the analysis time was given (Lee et al., 2006). The initial condition for the 6 km horizontal resolution was obtained at 1800 UTC 11 July. The 6 km horizontal resolution began at 2100 UTC after assimilating radar (reflectivity and radial velocity) and surface data by IAU from 1800 UTC 11 July and was integrated for 15 h until 1200 UTC 12 July. IAU in the 6 km horizontal resolution was applied at a 1 h time window from -30 to +30 min centered at the analysis time.

A number of assimilation and forecast experiments were conducted to examine the sensitivity of the forecast with different data sources. The sensitivity tests included the impacts of radar data, or surface data, or both combined. The CNTL experiment examined the results without data assimilation. The RADAR experiment examined the impact of assimilating only reflectivity and radial velocity data, and the AWS

experiment investigated the impact of assimilating only AWS surface data. Xiao et al. (2007) obtained sound analyses that had adjustments in both the dynamical and thermodynamical fields when both radial velocity and reflectivity were assimilated, and thus, we also assimilated both radial velocity and reflectivity data. RADAR+AWS experiment investigated the impact of radar and surface data assimilation.

5.2.2 Results of numerical simulations

Prior to the investigation of the impact of radar and surface data assimilation on the heavy rainfall forecast, the result of applying the IAU method to the initialization of the analysis is described. Figure 5.2 shows the evolution of the absolute tendency of surface pressure averaged over the 6 km horizontal resolution domain. The average absolute surface pressure and surface pressure tendency are indicators of the level of noise in the external gravity wave component to measure the degree of balance of the model solution. The initialization performance for the experiment with IAU at 1800 UTC was very similar to that for the experiment without IAU because both experiments used the same forecast at that time. The inclusion of data at 1900 and 2000 UTC caused a fluctuating curve in the experiment without IAU, indicating the presence of gravity waves for initialization in the experiment without IAU. It implies significant imbalance and inconsistency in the initial condition for the experiment without IAU method. However, using the IAU

method, the gravity wave fluctuation at 1900 and 2000 UTC was greatly reduced and the noise was effectively removed. This reduction in gravity wave by IAU will help to reduce aliasing in subsequent analyses. After 3 h, the noise stayed low with small oscillations for both experiments.

Figure 5.3 shows the 12 h accumulated rainfall amounts from 2100 UTC 11 July to 0900 UTC 12 July for CNTL, RADAR, AWS, and RADAR+AWS. The major difference between the results of these experiments was in the spatial distribution of the rainfall. The RADAR+AWS simulation showed better agreement with the observation results than the other simulations. CNTL, RADAR, and AWS were likely to simulate the spread of rainfall distribution. In CNTL, RADAR, and AWS, erroneous precipitation was simulated near the southwestern coast, with local maxima of 240.8, 220.2, and 220.8 mm, respectively. In contrast, for RADAR+AWS, the heavy rainfall was concentrated near Goyang. RADAR+AWS captured a 12 h rainfall amount that was greater than 220 mm, with a band-shaped structure over the central Korean Peninsula, even though the rainfall peak was delayed by approximately 2 h when compared with the observed rainfall peak (Fig. 5.4) and the maximum precipitation (287.9 mm) was underestimated compared with observation (335.0 mm). In the comparison between AWS and RADAR, the simulated rainfall over the central Korean Peninsula in AWS was comparable to that in RADAR+AWS, and the maximum rainfall amount over the central Korean Peninsula was 348.8 mm, representing an overestimation when compared with the observation. The rainfall distribution in RADAR was

more spread out with a relatively lower maximum rainfall amount (191.4 mm) over the central Korean Peninsula as compared to that in AWS. In the time series of the maximum rainfall point, the simulated rainfall in RADAR began and ended in the early hours of the forecast, but in AWS it began in the late hours of the forecast and continued up until the final hour as compared with the other experiments (Fig. 5.4). The evolution of hourly rainfall at the maximum rainfall point in RADAR+AWS was similar to that in CNTL.

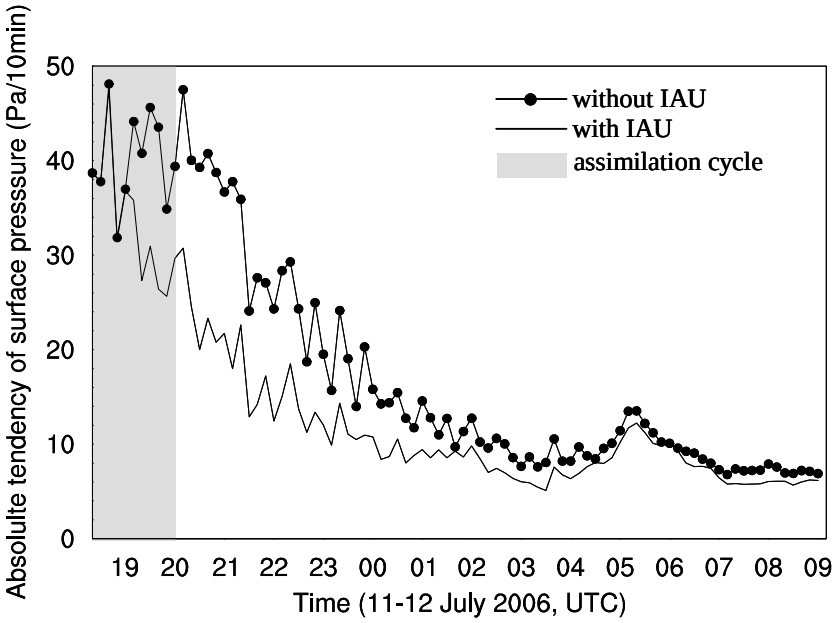

Figure 5.2 Averaged absolute tendency of surface pressure by 3DVAR without IAU (solid line with dots) and with IAU (solid).

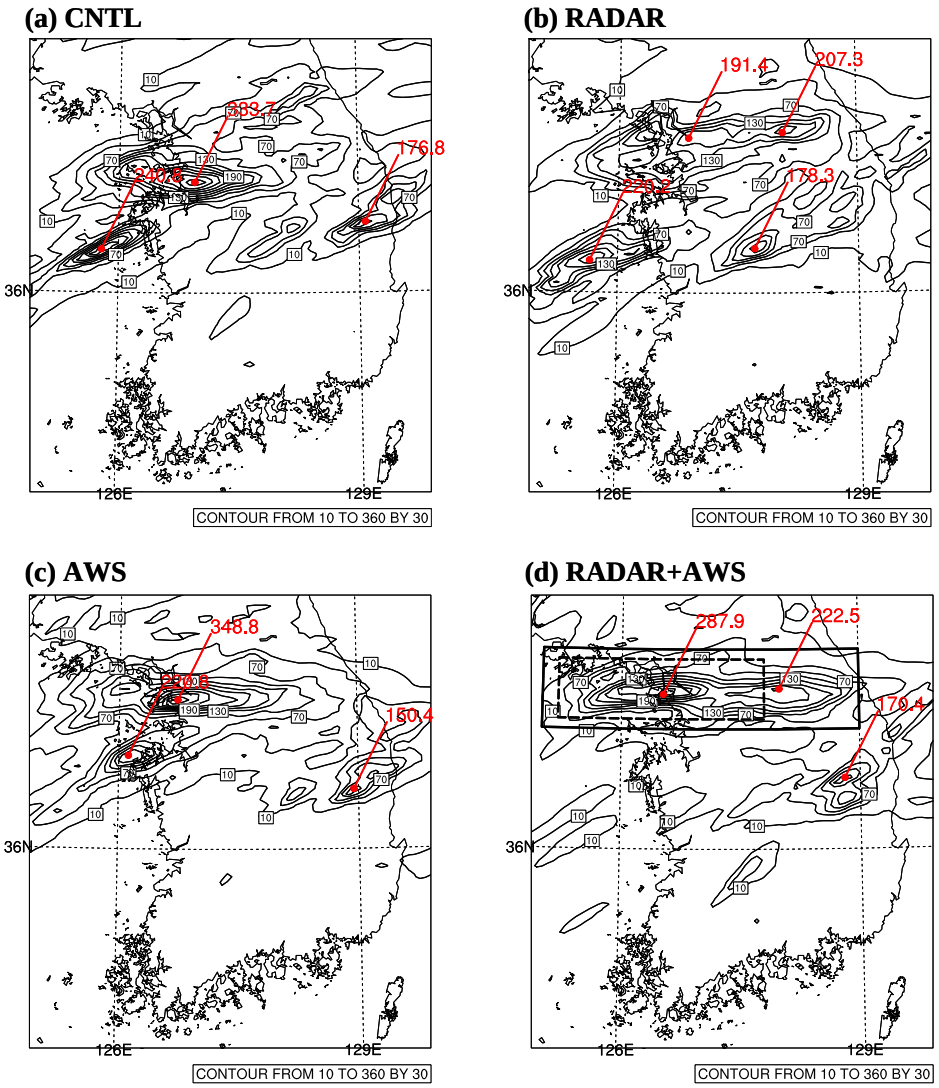

Figure 5.3 Total accumulated 12 h rainfall (mm, lines) from 2100 UTC 11 July to 0900 UTC 12 July for (a) CNTL, (b) RADAR, (c) AWS, and (d) RADAR+AWS. The dashed box in (d) denotes the analyzed area in Fig. 5.6, and the solid box in (d) shows the area of interest used for area-mean analysis in Fig. 5.11.

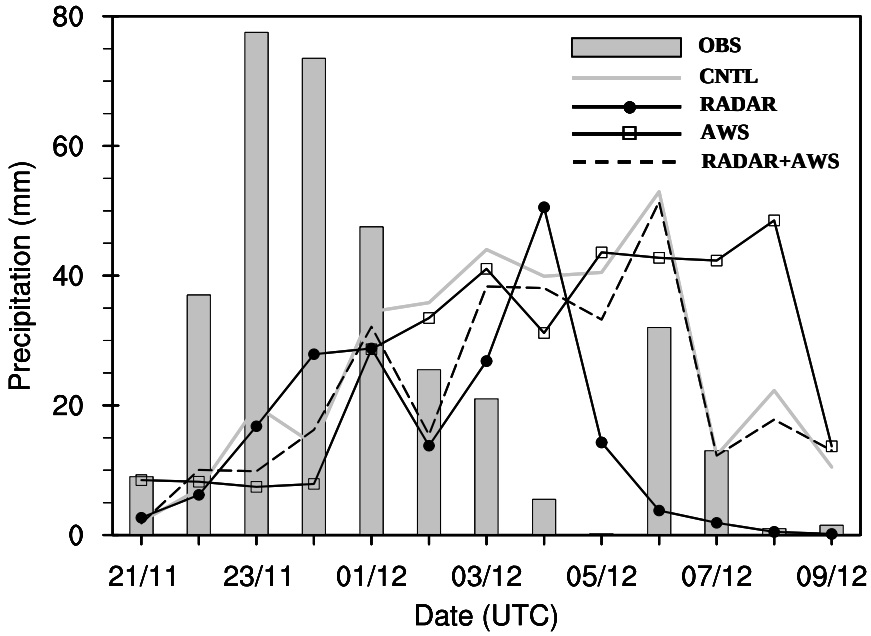

Figure 5.4 Time series of the observed and simulated hourly precipitation (mm) at the grid point of the maximum accumulated 12 h rainfall.

To evaluate the ability of the model at QPF skills for the designed experiments, the threat scores of precipitation forecast in each experiment, verified against the 12 h accumulated precipitation from 2100 UTC 11 to 0900 UTC 12 July are calculated. Figure 5.5 shows the threat scores of precipitation for CNTL, RADAR, AWS, and RADAR+AWS experiments. The threat scores of the threshold value with 10 mm showed the similar values among the experiments because the land area with the 10 mm rainfall was comparable in all experiments (Fig. 5.3). However, the experiments with data assimilation produced higher threat scores than CNTL for heavy rainfall (thresholds of 50 and 70 mm). It should be noted that RADAR+AWS resulted in the highest threat score among the experiments with the data assimilation for 50 and 70 mm threshold value, because RADAR+AWS improved the location of the heavy rainfall near Goyang as compared with the other experiments (Fig. 5.3). This means that assimilation of both multiple-Doppler radar and surface data had a more positive impact on QPF than the assimilation of either radar or surface data alone. In addition, the improvement in QPF with multiple-Doppler radar and surface data assimilation was more obvious for heavy rainfall than light rainfall.

To investigate the evolution of the simulated convective system, Figure 5.6 shows the simulated reflectivity with a 2 km height from 2300 UTC 11 July to 0100 UTC 12 July at 10 min intervals for RADAR+AWS, which produced the best precipitation forecast among the four experiments. The plotted arrows show the movements of convective cells.

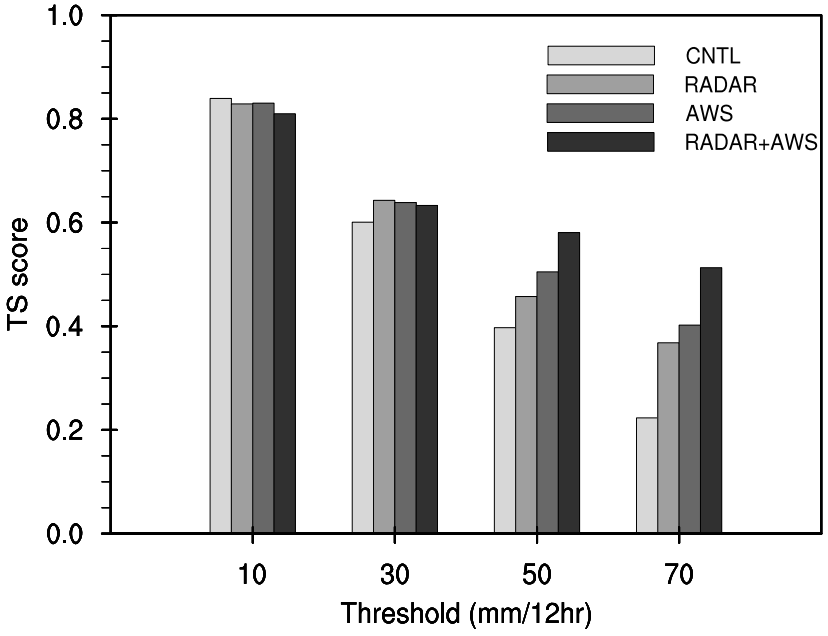

Figure 5.5 Threat scores for accumulated 12 h precipitation from 2100 UTC 11 July to 0900 UTC 12 July in the CNTL, RADAR, AWS, and RADAR+AWS experiments.

To the west of the convective system, new cells were successively generated in the Yellow Sea. The generated cells moved eastward and merged into the existing convective system. The evolution of the simulated convective systems agreed with observations (Fig. 4.7), albeit with some discrepancies. The general evolution of the simulated storms in RADAR+AWS indicated good performance of the model and data assimilation. We now discuss the impact of radar and surface data on simulating storm development.

Figure 5.7 shows the simulated reflectivity, wind vectors and LLJ, defined as regions with wind speeds greater than 12.5 m s^{-1}, at a 2 km height for RADAR+AWS, RADAR, and AWS in order to investigate the impact of each data type. At 2300 UTC 11 July, the simulated storm initiated, and at 0100 UTC 12 July the maximum rainfall intensity occurred in the targeted storm. At 2300 UTC 11 July, RADAR, AWS, and RADAR+AWS simulated strong reflectivity in the east–west direction over 40 dBZ near the west coast of the central Korean Peninsula. In RADAR+AWS and RADAR, the simulated reflectivity pattern over the central Korean Peninsula was comparable with observations, while that in AWS showed a wide-spreading reflectivity over the central western Korean Peninsula. At 0100 UTC 12 July, the simulated reflectivity was highly concentrated in the central Korean Peninsula by RADAR+AWS and RADAR, but much more spread out by AWS. The results indicated that the radar data, rather than the surface data, contributed to the development of convective cells in the models. It is noted that the strong

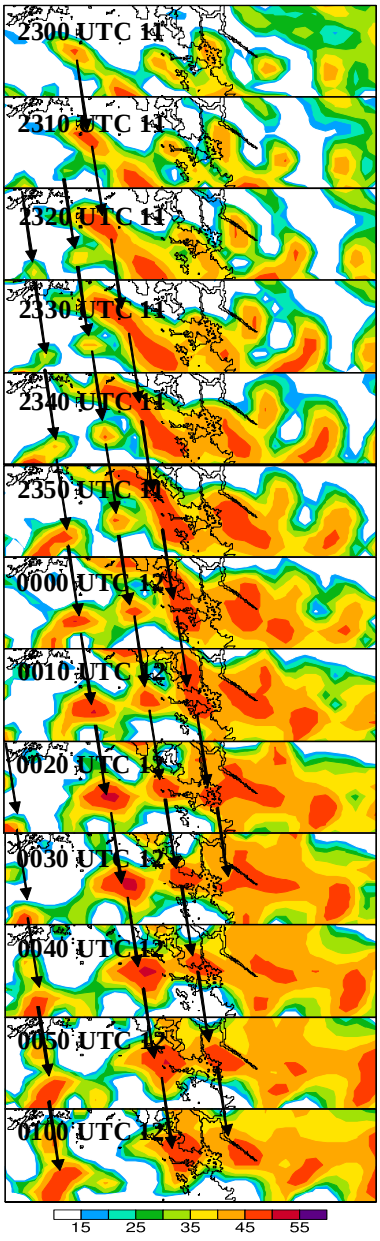

Figure 5.6 Simulated reflectivity at the 2 km height in the RADAR+AWS experiment from 2300 UTC 11 July to 0100 UTC 12 July. Arrows indicate the movement of convective cells.

reflectivity occurred along the northern edge of the LLJ in RADAR+AWS, RADAR and AWS. This suggests that interactions existed between the MCS and LLJ, which was consistent with observations (Fig. 4.2b and Fig. 4.7). At 0100 UTC 12 July, the strong reflectivity in RADAR+AWS moved northeastward in relation to at 2300 UTC, and it then developed around Goyang along the northern edge of the LLJ (Fig. 5.7b). RADAR simulated the strong reflectivity around Goyang; however, another strong reflectivity over the southern Korean Peninsula developed along the southward-shifted LLJ (Fig. 5.7d). This implies that the rainfall distribution in RADAR was somewhat more spread out relative to RADAR+AWS. AWS also simulated the strong reflectivity over the central Korean Peninsula along the LLJ, even though the LLJ did not extend over the eastern Korean Peninsula (Fig. 5.7f). On the basis of these results, the radar data contributed to better development of convective storms, and the surface data contributed to an increase in low-level wind over the central Korean Peninsula. Therefore, the assimilation of multiple-Doppler radar and surface data together resulted in better simulation of heavy rainfall in RADAR+AWS than in each individual source only.

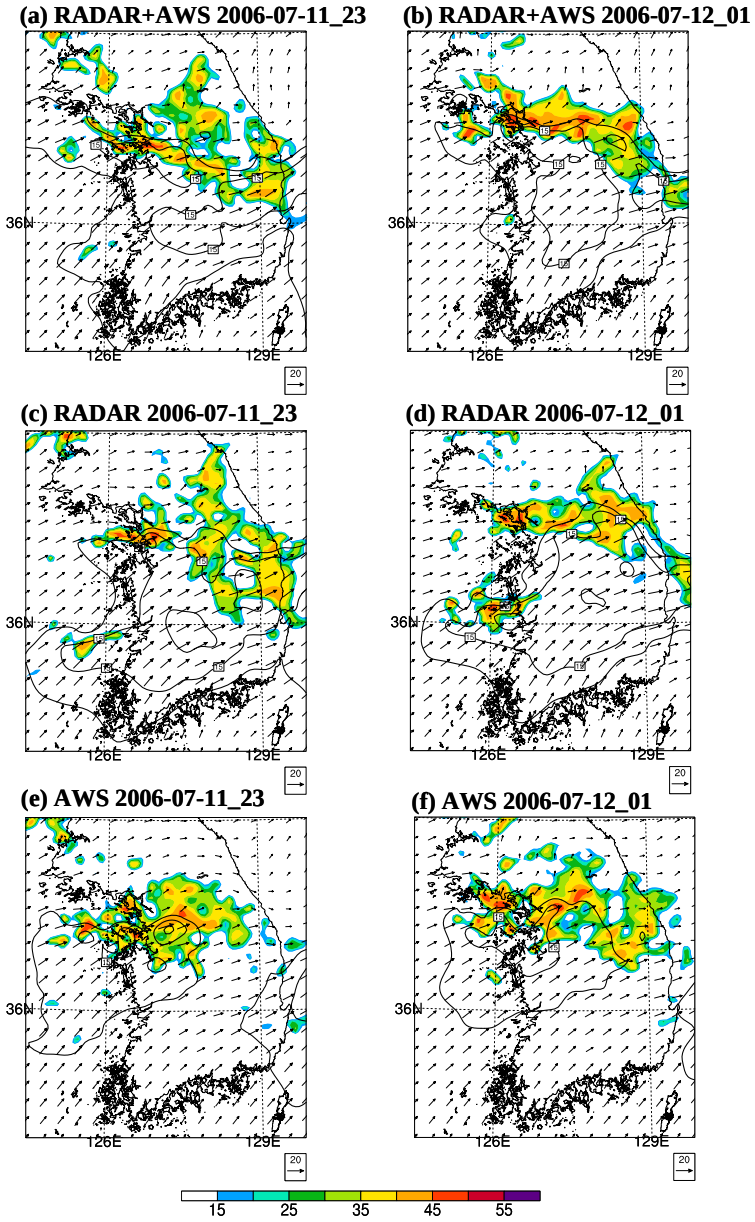

Figure 5.7 Simulated reflectivity (dBZ, shaded), wind speed (m s^{-1}, lines) and wind vectors (m s^{-1}) at 2 km height at 2300 UTC 11 July and 0100 UTC 12 July for (a, b) RADAR+AWS, (c, d) RADAR, and (e, f) AWS experiments.

To more investigate the effect of the radar data, Figure 5.8 shows the 850 hPa rain water mixing ratio difference from AWS experiment (Figs. 5.8a and 5.8b) for experiments RADAR+AWS and RADAR at 2300 UTC 11 July. The rainwater mixing ratio showed a positive difference over the west coast of the central Korean Peninsula, which was consistent with the area of strong reflectivity as shown in Fig. 5.7. The maximum increase of the rainwater mixing ratio was over 2 g kg^{-1}. These positive differences in rainwater mixing ratio seemed to cause highly concentrated convection over the west coast of the central Korean Peninsula as compared with AWS experiment, and contributed to the development of the convection. This is because the reflectivity of the radar data seemed to facilitate storm development in the early hours of the forecast, which was consistent with the time series of the maximum rainfall point, and more radar data points (~10,000) were available near the Goyang area compared with surface data. The north-south vertical cross sections of an intense convective cell in RADAR+AWS and RADAR experiment at 2300 UTC 11 July were analyzed in Fig. 5.9. The distinct character of the simulated convective cell could be seen in the pronounced vertical structure, and it was noted that the vertical structure of both experiments was similar. The strong updraft associated with the convergence at the lower troposphere was simulated, and the divergence existed in the upper troposphere (Figs. 5.9a and c). Comparing the equivalent potential temperature (Figs. 5.9b and d), both experiments showed that the low-level southerly or southwesterly flow supplied the warm and moist air to

the convective cell, and most of this inflow became the updraft. And a significant outflow was found at the top of the convection cell, and extended far to the north. This is typical characteristics of developed convective cell as other studies have shown (Sun and Lee, 2002; Kim et al. 2006). The results indicated that the radar data assimilation contributed to the establishment of the vertical structure of convective cell, even though RADAR+AWS experiment simulated the enhanced convergence and divergence in the low-level perhaps due to the assimilation of the surface data as compared with RADAR experiment.

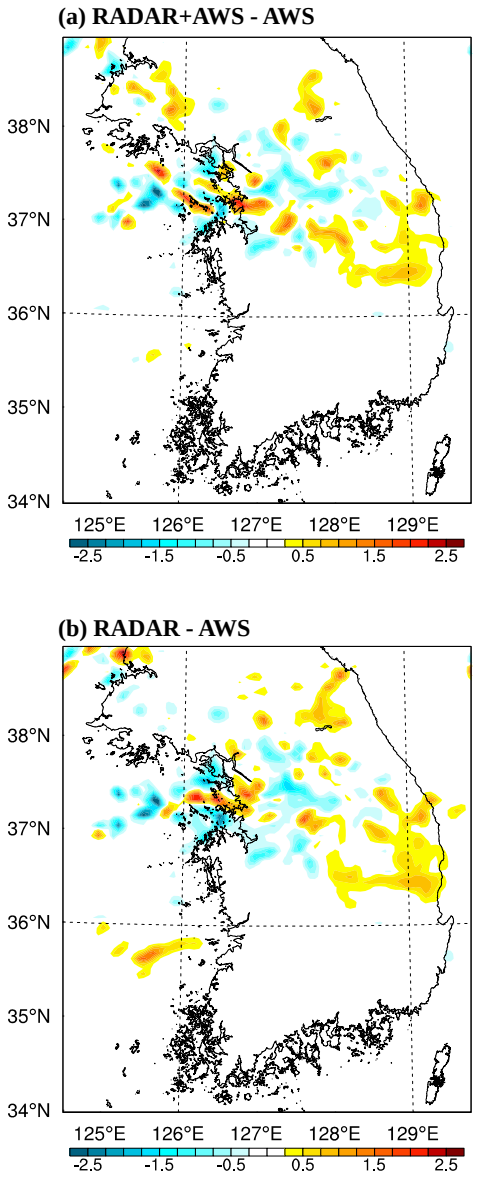

Figure 5.8 Analyzed 850 hPa water vapor mixing ratio difference from AWS by (a) RADAR+AWS and (b) RADAR at 2300 UTC 11 July 2006.

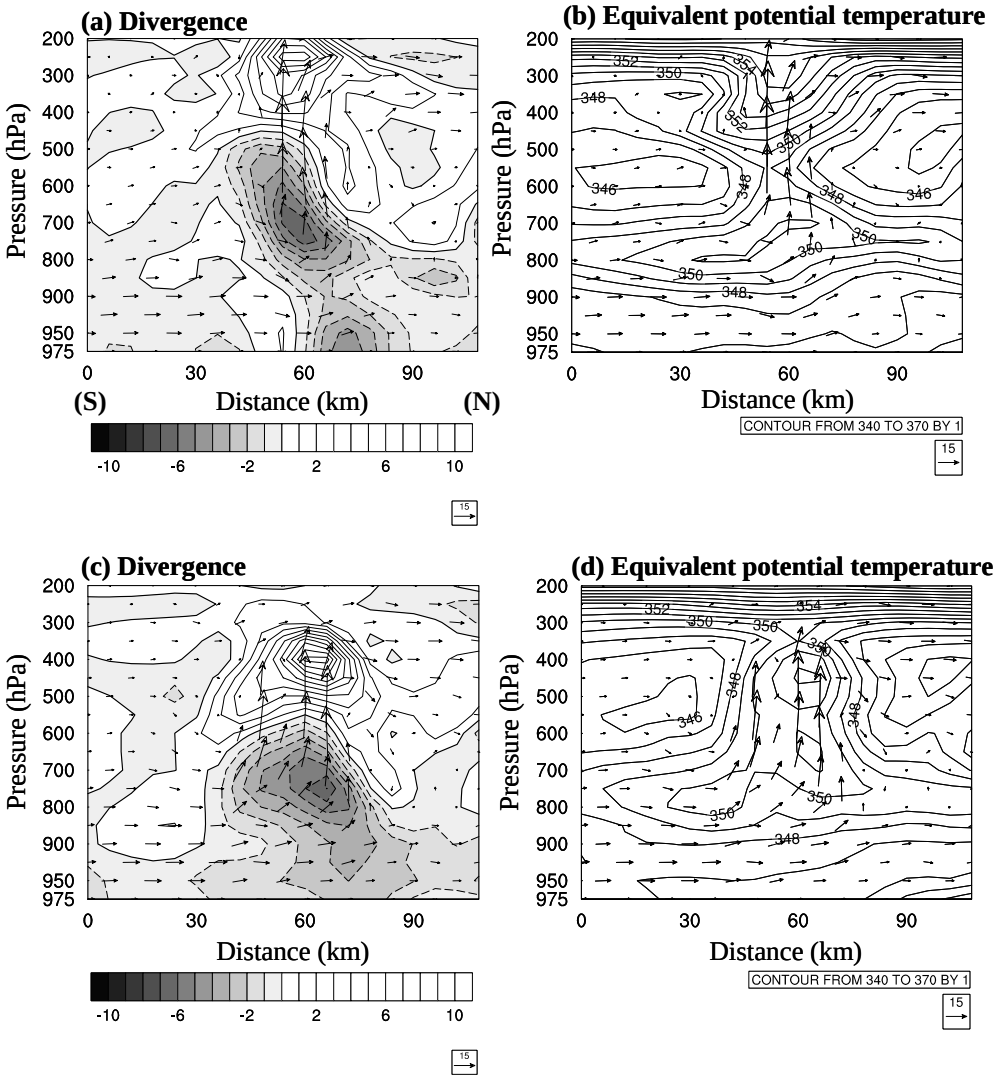

Figure 5.9 North-south vertical cross sections of the convective systems at 2300 UTC 11 July 2006 for (a, b) RADAR+AWS, and (c, d) RADAR experiment: (a, c) divergence (10^{-4} s^{-1}, positive (negative) values solid (dashed) and negative areas are shaded) and wind vectors (m s^{-1}), and (b, d) equivalent potential temperature (K, lines) and wind vectors (m s^{-1}).

We analyzed the temperature field at a height of 2 m (air temperature) at 2300 UTC 11 July and 0100 UTC 12 July (Fig. 5.10) to further investigate the effect of surface data. The distribution of the simulated temperature in RADAR+AWS was similar to that observed: cold air dominated over the northern and central Korean Peninsula, while warm air dominated over the southern Korean Peninsula at 2300 UTC 11 July. After 2 h, the warm air extended to the heavy rainfall region. In the difference fields (Figs. 5.10c and d), the negative difference between RADAR+AWS and RADAR over the central Korean Peninsula was due to cold advection and evaporative cooling caused by precipitation, while the positive difference over the southern Korean Peninsula seems to have been mainly due to warm advection by the southwesterly flow. Thus, the thermal gradient in the lower troposphere was enhanced when surface data were added to the data assimilation. Tao and Soong (1991) showed that surface forcing played an important role in the development of squall lines. Therefore, in this study, we investigated surface forcing when surface data assimilation was added. Figure 5.11 shows the time series for the ground temperature, friction velocity, height of the planetary boundary layer (PBL), and precipitation averaged over the heavy rainfall region (125–129°E and 37.2–38°N) in RADAR+AWS and RADAR, which is defined by the solid box in Fig. 5.3d. RADAR+AWS and RADAR had similar ground temperature distribution, as shown in Fig. 5.11a. On the other hand, the friction velocity was larger for RADAR than for RADAR+AWS (Fig. 5.11b). The

larger friction velocity induced a higher PBL height, which resulted in a decrease of the specific humidity at lower levels due to entrainment of relatively dry air from the above PBL; consequently, there was a smaller potential for convection in RADAR (Fig. 5.11c). That was why greater rainfall occurred in RADAR+AWS experiment with the surface data assimilation than in RADAR experiment, as shown in Fig. 5.11d. This result indicates that the surface data played an important role in inducing the lower tropospheric environment to be favorable for convection development in RADAR+AWS.

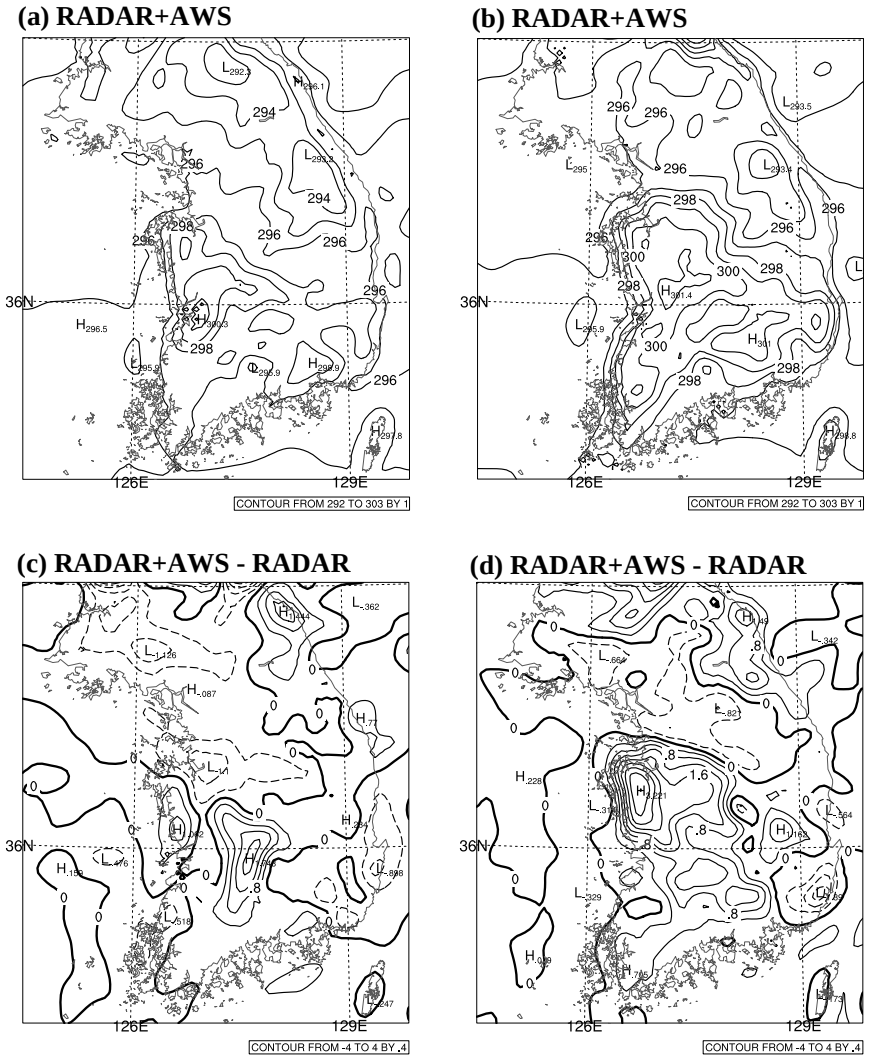

Figure 5.10 Simulated temperature (K, lines) at 2 m height in the RADAR+AWS experiment at (a) 2300 UTC 11 July and (b) 0000 UTC 12 July; (c) and (d) indicate the difference in temperature (K, positive (negative) values solid (dashed)) at 2 m height between the RADAR+AWS and RADAR experiments at 2300 UTC 11 July and 0100 UTC 12 July, respectively.

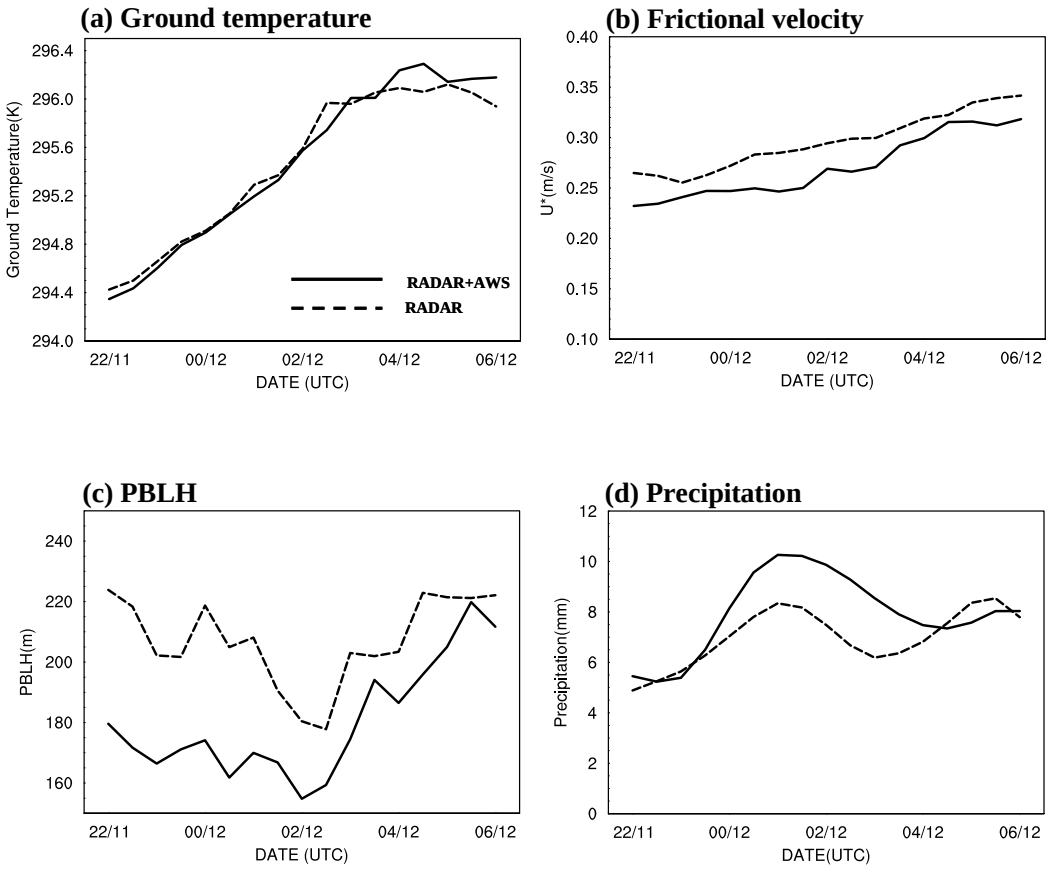

Figure 5.11 Temporal variation of (a) ground temperature (K), (b) friction velocity (m s^{-1}), (c) height of PBL (m), and (d) precipitation (mm) of the RADAR+AWS and RADAR experiments averaged over the black box shown in Fig. 5.3d.

5. 3 Tuning of the background error correlation length scale for surface data

5.3.1 Experiment design

The WRF and WRF 3DVAR systems are used to investigate the impact of the tuned length scale of background error statistics for the surface data assimilation. The model configuration was identical with the one in numerical experiments as mentioned in section 5.1. The initial and lateral boundary conditions for the coarse grid with 18 km horizontal resolution were provided by the NCEP FNL analyses with a 1° × 1° horizontal resolution at 6 h intervals. In addition, we assimilated conventional data for the coarse-grid domain to produce more accurate initial and boundary conditions for the 6 km horizontal resolution fine grid. The coarse grid simulation was carried out for 21 h forecasts after two update cycles (1200 and 1500 UTC) from 1200 UTC 11 July. The simulation on the 6 km horizontal resolution began at 2100 UTC 11 July after assimilating the surface data from 1800 UTC 11 July at 1 h time window and was integrated until 0900 UTC 12 July. For the data assimilation, a Rapid Update Cycle (RUC) 3DVAR system was applied to incorporate more frequent observations. Update cycling can supplement a disadvantage of non-varying with time in 3DVAR, and reduce the spinup time at the analysis state.

A number of assimilation and forecast experiments were conducted

to examine the forecast with respect to various changes of the length scale in the surface data assimilation for the heavy rainfall case. In the assimilation experiments, CNTL examined the results without data assimilation, and LS examined the impact of the original length scale from the NMC background error statistics to investigate the effect of the NMC background error statistics in the assimilation of the surface data. LS0.7, 0.5, and 0.3 examined the impact of the tuning factors of 0.7, 0.5, and 0.3 applied to the original length scale, respectively. We investigated the impact of the assimilated surface data with the large length scale of the recursive filter in LS0.7, the half of the original length scale in LS0.5, and the small length scale in LS0.3. In addition, we investigated the impact of the double iteration of the tuned length scales with tuning factors of 0.7 and 0.3 in DILS.

Figure 5.12 shows the vertical zonal wind increments using a single surface observation that is performed in section 2.2 to examine the vertical response with respect to the length scale. The assimilation with the tuned length scale (i.e., LS0.7, LS0.5 and LS0.3) showed a comparatively smaller correlated area than that obtained using the original length scale from the NMC background error statistics. It was noted that the experiments using smaller length scales tended to generate the shallow propagation depth with the enhanced vertical response in the boundary layer. The propagation depth in LS0.3 was slightly shallow as compared to the other experiments around an observation point. However, the vertical response in DILS was comparable with that in LS,

and DILS was characterized by a strongly correlated vertical pattern, particularly in the boundary layer.

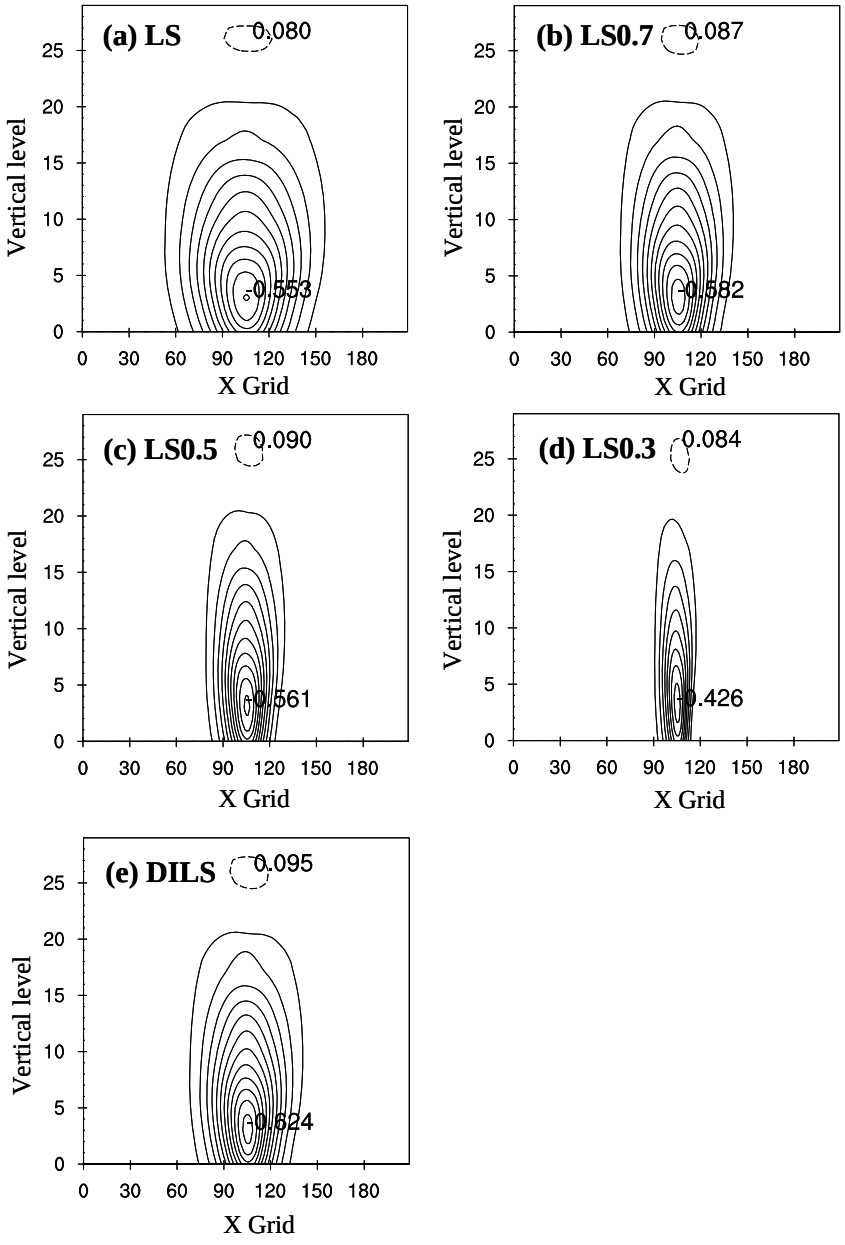

Figure 5.12 Vertical zonal wind responses for single observation test for (a) LS, (b) LS0.7, (c) LS0.5, (d) LS0.3, and (e) DILS.

5.3.2 Idealized experiment

Prior to investigating the impact of length scale on the heavy rainfall forecast, we conducted an ideal experiment that could illustrate the response of the length scale tuning. Even though the estimates derived in ideal experiments are unrealistically good, such experiments are valuable because we can gain information on the sensitivity of the estimation procedure to various inputs. Regarding the true state of the atmosphere as a composition of sinusoidal waves ($\sum_{n=1}^{13}\sin^2 nx$), a first guess was made using the filter of the initial condition for the 6 km horizontal grid mesh, with the frequency of wave number 2 set to approximately 300 km. The values extracted at the AWS locations shown in Fig. 1.1 from the true state were used as the observation data for the ideal experiment. Using the first guess and the observation, we conducted a number of assimilations with the tuned length scales. The results of these experiments were analyzed in the area defined by the dashed box in Fig. 1.1, because the influence of AWS data was limited around the Korean Peninsula.

Figure 5.13 shows the distribution of the analysis after the surface data assimilation for the various length scales. The relatively large length scale used in LS and LS0.7 contributed to the simulation of the large-scale features of the truth state, and the small-scale features tended to filter out (Figs. 5.13b and c). However, the assimilation using tuning factors of 0.5 and 0.3 enhanced the simulation of the small-scale features related to the

mesoscale phenomena due to the reduced length scale (Figs. 5.13d and e). LS0.3 simulated the large- and small-scale features of the truth state compared with the experiments using the tuned length scale; however, in the southwest part, the distribution and intensity in LS0.3 was different from that in the truth state. It was noted that DILS showed the best agreement with the truth state in terms of distribution and intensity (Figs. 5.13a and f). To further analyze the effect of the tuned length scale, Figure 5.14 presents the wave number distribution produced by a two-dimensional Fast Fourier Transform (FFT). The analysis of the wave number spectra clearly showed that the assimilation with the tuned length scale was able to contain small-scale features when compared to the LS simulation, which was consistent with Fig. 5.13. It should be noted that LS0.3 overestimated the simulation of small-scale features in the y-direction compared with the truth state, which resulted in contamination of the subsequent forecast (Fig. 5.14e). However, DILS better identified both large- and small-scale features of the truth state in the model fields, albeit with some discrepancies (Figs. 5.14a and f). These results indicate that the length scale tuning of background error statistics can be a critical process in the 3DVAR system, and that assimilation using the double iteration with two different length scales is able to reflect the observed information more effectively in the model fields by designing a suitable strategy for assimilating the surface data. Another effect of the double iteration method was that the assimilation system was able to utilize more observations than when using only a single iteration with the tuned

length scale. With the two different length scales approach, over 92% of the AWS observations were incorporated in the assimilation procedure of DILS, while about 73% of AWS observations were incorporated in LS, LS0.7, LS0.5, and LS0.3.

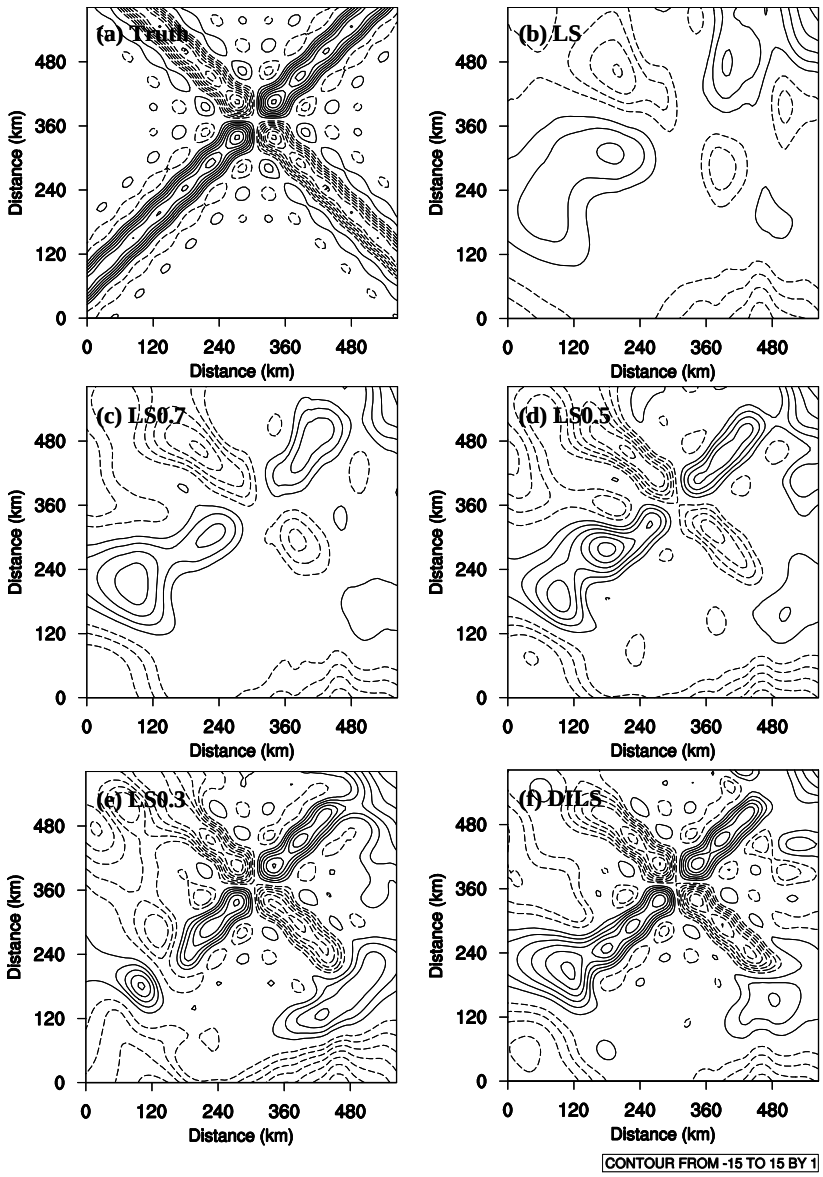

Figure 5.13 The distribution of (a) truth state, and (b), (c), (d), (e), and (f) assimilated fields for LS, LS0.7, LS0.5, LS0.3, and DILS, respectively, in the dashed box shown in Fig. 1.1 [Positive (negative) isoline is solid (dashed), and the zero isoline is omitted].

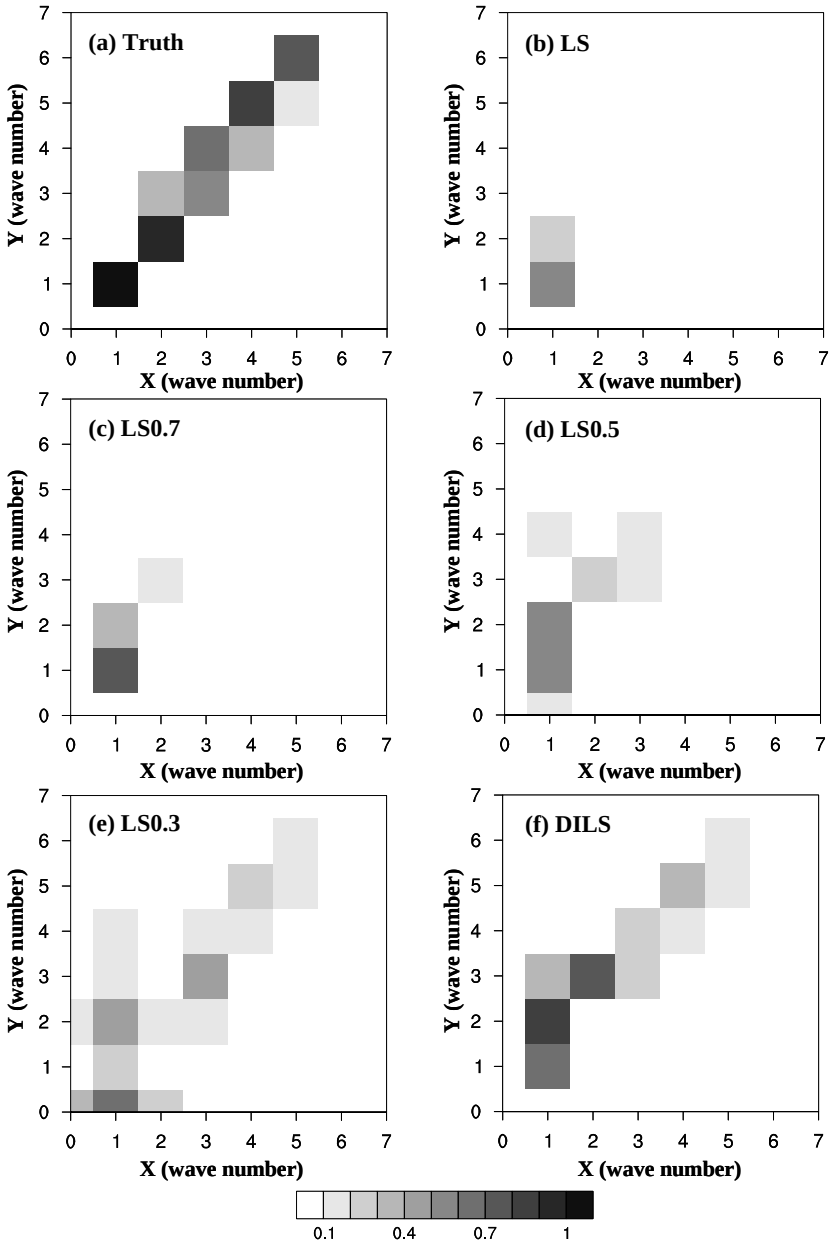

Figure 5.14 Two-dimensional wave number distribution calculated by two-dimensional Fast Fourier Transform (FFT) for ideal experiments in (a) truth, (b) LS, (c) LS0.7, (d) LS0.5, (e) LS0.3, and (f) DILS.

5.3.3 Results of numerical simulations for heavy rainfall

Figure 5.15 shows the bias of domain-averaged zonal wind (Fig. 5.15a) and meridional wind (Fig. 5.15b) to compare the analysis fields at 1800 UTC 11 July. The biases were calculated over the southern Korean Peninsula (34°-38°N, 126°-130°E) by using the NCEP FNL data with its horizontal resolution of 1.0°, approximately 100 km. The NCEP FNL data were used as an observational reference for the comparison among the experiments because there was no availability of atmospheric mesoscale wind observations in comparison for this study. The results showed that the surface data assimilation affected large difference of the bias in the layer between 950 and 800 hPa. In addition, the bias differences were more significant for meridional wind than zonal wind. The experiments with the surface data assimilation resulted in smaller biases at low levels than CNTL. It should be noted that, even with the smallest length scale in this study, LS0.3 resulted in the largest bias among the experiments with the surface data assimilation, which was consistent with the result in Fig. 5.14. Among all experiments, the assimilation using DILS was in better agreement with the observation reference. This meant that data assimilation using a double iteration of two different length scales contributed positive impact to the analysis fields than the other experiments.

The 12 h accumulated rainfall predictions from 2100 UTC 11 July to 0900 UTC 12 July for CNTL, LS, LS0.7, LS0.5, LS0.3, and DILS are shown

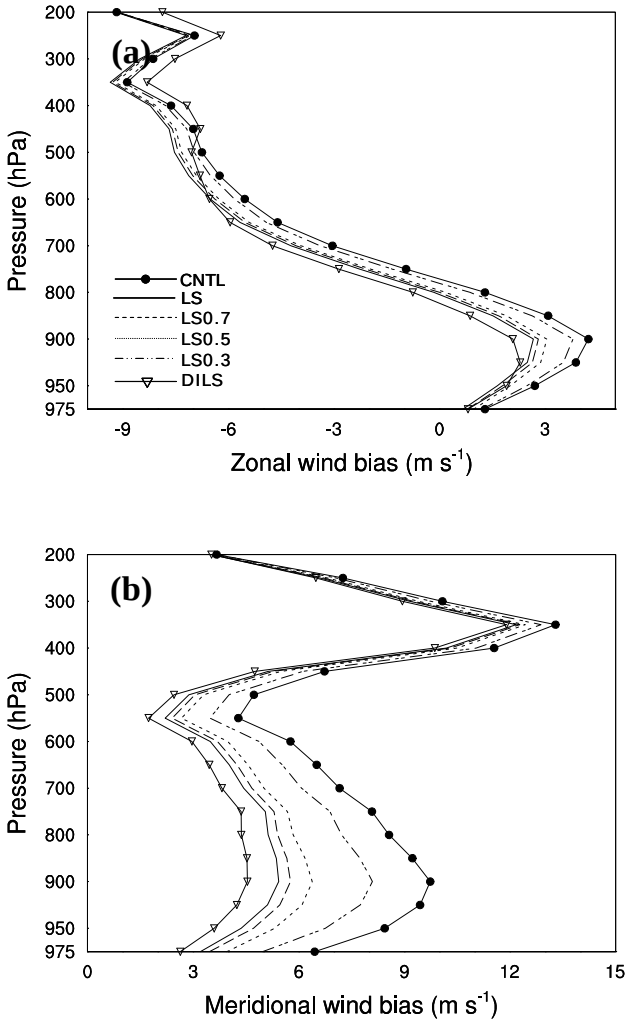

Figure 5.15 Vertical distribution of (a) zonal and (b) meridional wind bias in the analysis field at 1800 UTC 11 July 2006; CNTL (solid line with dots), LS (solid line), LS0.7 (dashed line), LS0.5 (dotted line), LS0.3 (dash-dotted line), and DILS (solid line with down-triangles).

in Fig. 5.16. The main difference between the results was in the spatial distribution and amount of rainfall. The experiments using smaller length scales tended to generate a comparatively narrow rainfall distribution with an increased rainfall amount. Among the experiments, the DILS simulation showed better agreement with the observational results than the other simulations. It captured a 12 h accumulated rainfall amount of over 300 mm, with a band-shaped structure over the central Korean Peninsula, even though some erroneous precipitation was simulated near the southwestern coast. In contrast, the CNTL simulation produced a rainband that shifted southward, while the location of the maximum rainfall in LS was shifted eastward as compared to the observation data. The simulations with the tuned length scale (i.e., LS0.7, LS0.5, and LS0.3) were in better agreement with observations than LS; however, their maximum precipitation underestimated the observed amount. Between LS0.7, LS0.5, and LS0.3, we could see that the location of the maximum rainfall in LS0.5 agreed well with the observation data, which was consistent with the results in Fig. 2.4, but gave a relatively low maximum rainfall amount (233.6 mm).

To investigate the structure of the simulated convective system, Fig. 5.17 shows the north–south vertical cross sections of an intensive convective cell in DILS, which produced the best precipitation forecast among the six experiments, at 2230 UTC 11 July 2006. The convective core of strong reflectivity (>45 dBZ) reached 500 hPa altitude with a width of approximately 30 km (Fig. 5.17a). The strong updraft was

located in the mid-troposphere (500 hPa), and a downdraft of less than 1 m s^{-1} existed at the low-level (Fig. 5.17b), which was consistent with the observation. The intense convective cells simulated strong convergence below 600 hPa and divergence in the lower troposphere, as for general convective cells in their mature stage (Fig. 5.17c). Behind the convection (northern side of the convection), a downdraft with a maximum speed of 0.27 m s^{-1} occurred at approximately 900 hPa, in the area between 45–55 km where the strong reflectivity was found. This downdraft enhanced the convergence, and consequently the updraft, at 30–50 km. However, this did not seem to be the most crucial factor in the development of strong updrafts when we considered its strength. Indeed, it seemed that the strong convergence along the northern edge of the LLJ corresponded to the strong updraft. Lee et al. (1998) suggested that strong convergence ahead of the LLJ region played a key role in the production of heavy rain associated with an eastward propagation disturbance. In addition, the strong low-level southerly or southwesterly flow supplied warm and moist air to the convective system. Thus, the equivalent potential temperature and convective instability $(\partial\theta_e / \partial p)$ showed the presence of unstable stratification between 700–800 hPa by a high θ_e inflow (Fig. 5.17d). The present convective system showed both similarities and dissimilarities to other convective systems. It showed some similarities in its structure, such as the low-level inflow ahead of the band, and the downdraft behind the band. However, unlike the squall lines, the updrafts tilted northward below about 700 hPa and stood upright above

about 700 hPa, which induced the continued development of convection and significant precipitation. This was consistent with the analysis of observed radar data.

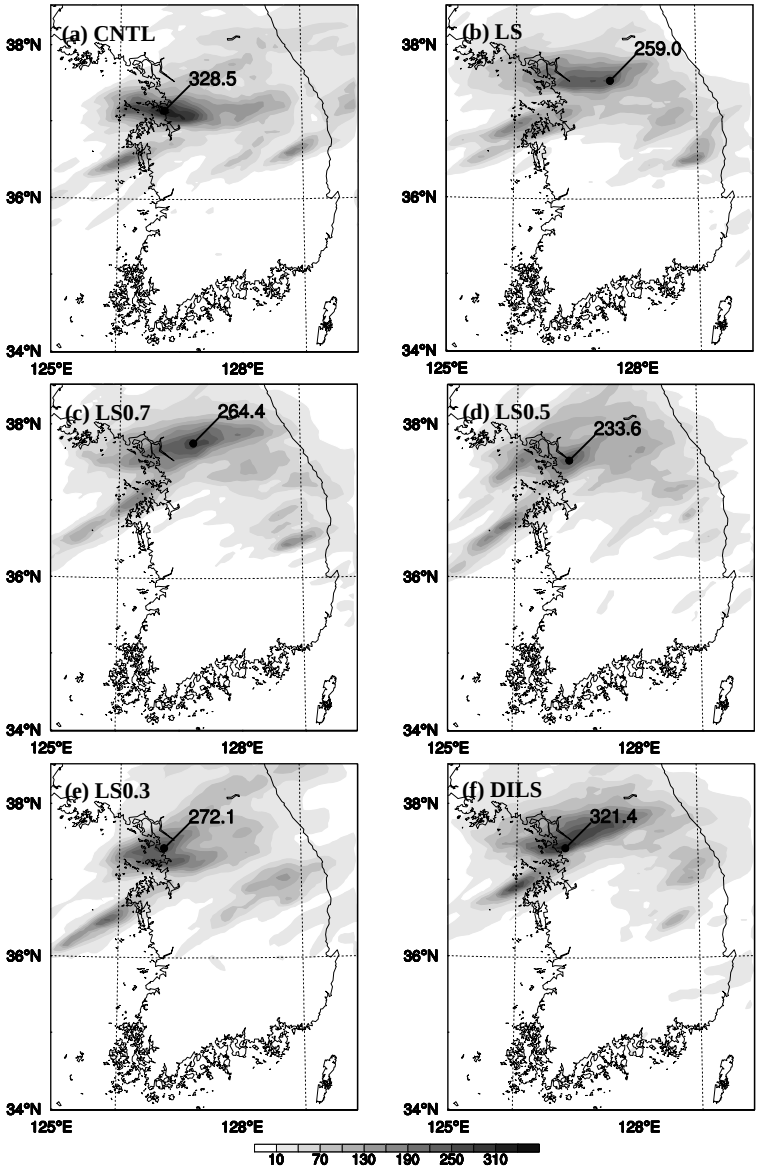

Figure 5.16 Total accumulated 12 h rainfall (mm) from 2100 UTC 11 July to 0900 UTC 12 July for (a) CNTL, (b) LS, (c) LS0.7, (d) LS0.5, (e) LS0.3 and (f) DILS.

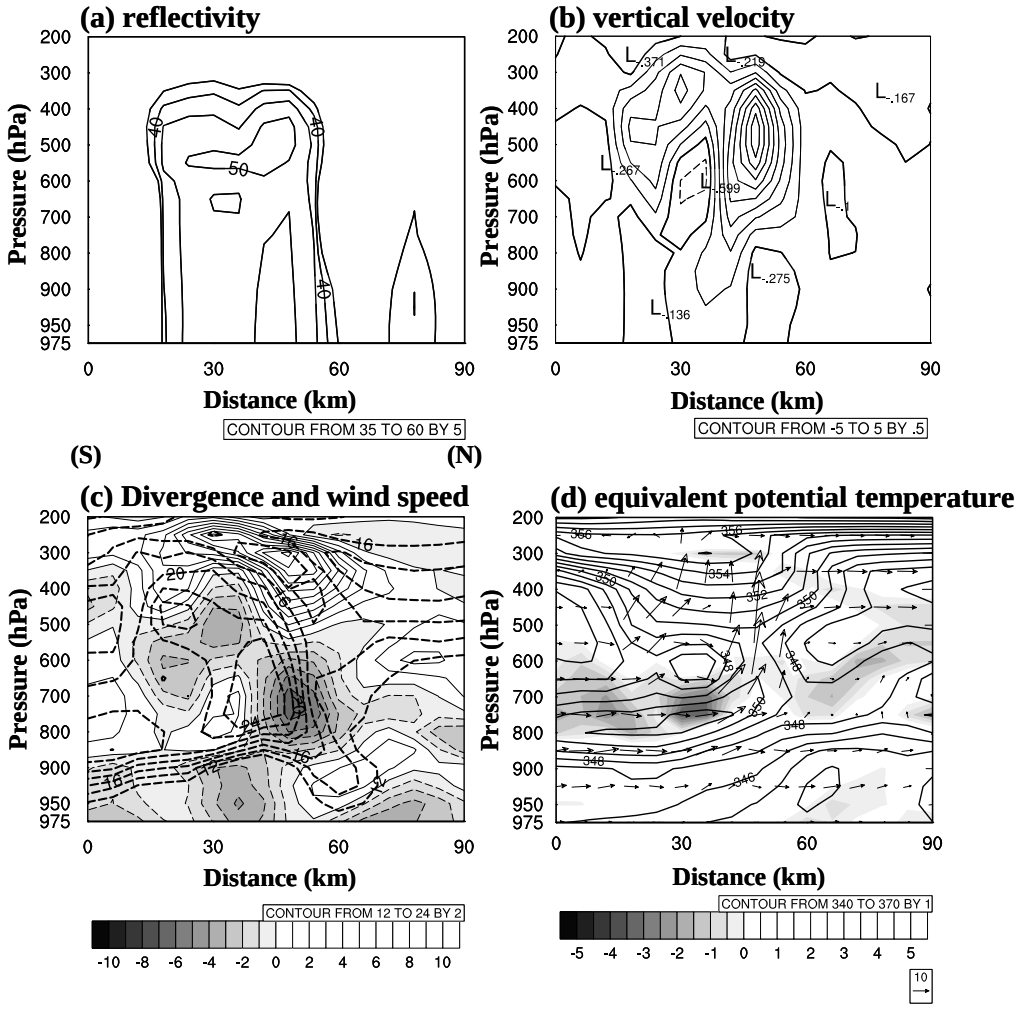

Figure 5.17 North–south vertical cross sections of the intense convective system in DILS for 2230 UTC 11 July 2006: (a) reflectivity (dBZ, lines), (b) vertical velocity (m s^{-1}, positive (negative) values solid (dashed)), (c) divergence (10^{-4} s^{-1}, positive (negative) values solid (dashed) and negative areas are shaded) and wind speed greater than 12 m s^{-1} (thick dashed), and (d) equivalent potential temperature (K, lines), convective instability (K hPa^{-1}, negative areas are shaded), and wind vectors (m s^{-1}).

Figure 5.18 shows the observed and simulated reflectivity, wind speed (≥ 12.5 m s^{-1}), divergence, equivalent potential temperature, and wind vectors at a 1.5 km height for DILS, LS0.5, and LS at 2200 UTC 11 July. DILS simulated strong reflectivity over 40 dBZ in the east–west direction near the west coast of the central Korean Peninsula, which agreed with the radar observations, albeit with some discrepancies (Figs. 5.18a and b). However, LS0.5 generated a convective band that was shifted northwestward with a wide-spreading reflectivity (Fig. 5.18c), and the simulated reflectivity in LS was shifted southward and much more spread out (Fig. 5.18d). The results indicated that the simulation using the double iteration with two different length scales, rather than the simulation using the single iteration with the tuned length scale, contributed to the better development of convective cells in the early hours of the forecast. After length scale tuning in DILS and LS0.5, the low-level wind over the central Korean Peninsula was intensified compared to the LS simulation (Figs. 5.18b, c, and d), because the length scale tuning in the surface data assimilation had significantly affected the extent to which the information was spread and the intensity of wind increment (Fig. 5.19). Figure 5.19 shows the lowest level analysis increments of zonal wind and temperature at 1800 UTC 11 July. The wind increments were propagated over small spatial distance and enhanced locally as smaller length scale of recursive filter was used. However, the distribution of temperature increments was similar with respect to the length scale tuning, except for reducing the correlated area

as the small length scale was used. From the results, we noted that the length scale tuning in the surface data assimilation more affected wind increments. Therefore, the length scale tuning in the surface data assimilation resulted in reduced spatial correlation of the winds and gave rise to a strengthened low-level wind increment. The wind strengthening over the central Korean Peninsula was consistent with the skew T log P chart at Osan, which is approximately 60 km south of Goyang. The sounding data showed that the wind speed at the 850 hPa level increased to 20 m s^{-1} between 1800 UTC 11 July and 0000 UTC 12 July (Fig. 4.5). In DILS and LS0.5, an area of equivalent potential temperature over 352 K extended almost to the observed heavy rainfall region and the strong convergence corresponded well to the area of strong reflectivity (Figs. 5.18e and f). In this environment, a strong southwesterly could continuously transport a large amount of warm and moist air to the central Korean Peninsula, and so the strengthened low-level wind due to the surface data assimilation with the tuned length scale facilitated the development of the convective system near the observed heavy rainfall region. Based on these results, the length scale tuning in the surface data assimilation contributed to an increase in the low-level wind, which played an important role in the development of this convective system. In addition, the simulation using the double iteration with two different length scales contributed to the better development of convective storms, which resulted in better simulation of the heavy rainfall.

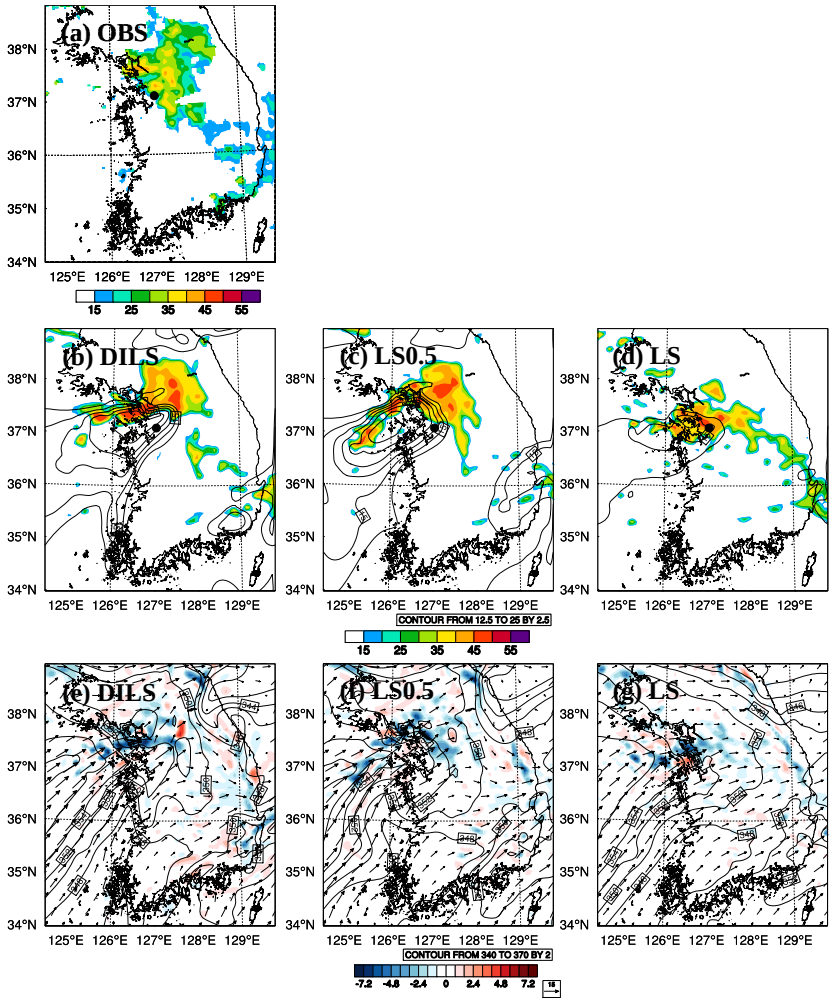

Figure 5.18 (a) The radar-observed reflectivity, and (b), (c), and (d) show the simulated reflectivity (dBZ, shaded) and wind speed (m s^{-1}, contours) for DILS, LS0.5, and LS, respectively, at 1.5 km height at 2200 UTC 11 July. (e), (f), and (g) represent divergence (10^{-4} s^{-1}, shaded), equivalent potential temperature (K, contours), and wind vectors (m s^{-1}) for DILS, LS0.5, and LS, respectively. The black dot in (a), (b), (c), and (d) denotes the location of Osan.

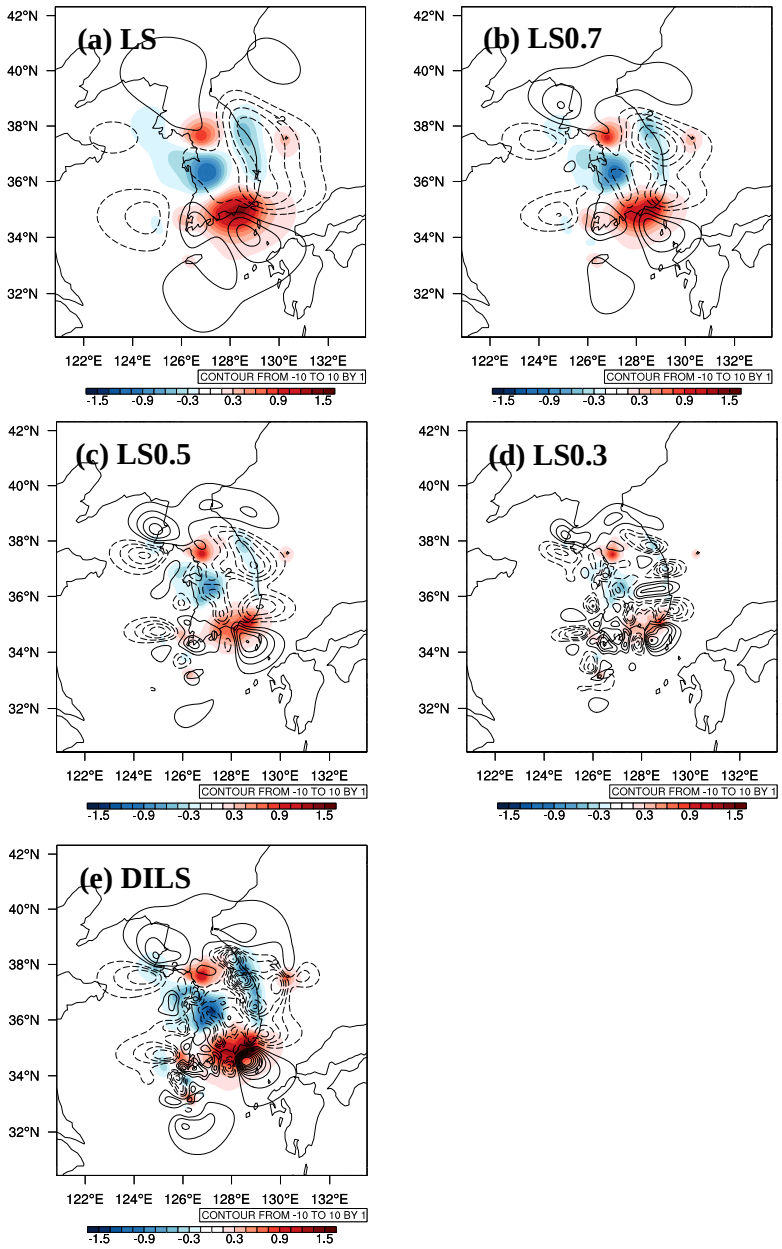

Figure 5.19 Analysis increment of zonal wind (positive (negative) isoline is solid (dashed)) and temperature (shaded) at 1800 UTC 11 July 2006.

In order to assess the performance of the assimilation of surface data with the tuned length scale in the WRF 3DVAR, the surface data were assimilated in 3DVAR during one-month period (July 2006). Figure 5.20 shows the RMSEs of lowest level zonal wind (Fig. 5.20a), meridional wind (Fig. 5.20b), and temperature (Fig. 5.20c) after the 3DVAR assimilation verified against the AWS surface observations for CNTL, LS and DILS experiments. For most of cases, the analysis errors for winds and temperature were improved in the experiments with the surface data assimilation as compared with CNTL. The analysis error statistics showed that the 3DVAR assimilation using the double iteration with two different length scales yielded smaller analysis error than the other experiments. This result suggests that the reflection of the observed information effectively in the model field for DILS can eliminate improper information caused by the longer length scale of the background error in the NMC method. The RMSEs of zonal wind was significantly improved in the DILS as compared with LS, while, the RMSEs of temperature was improved slightly.

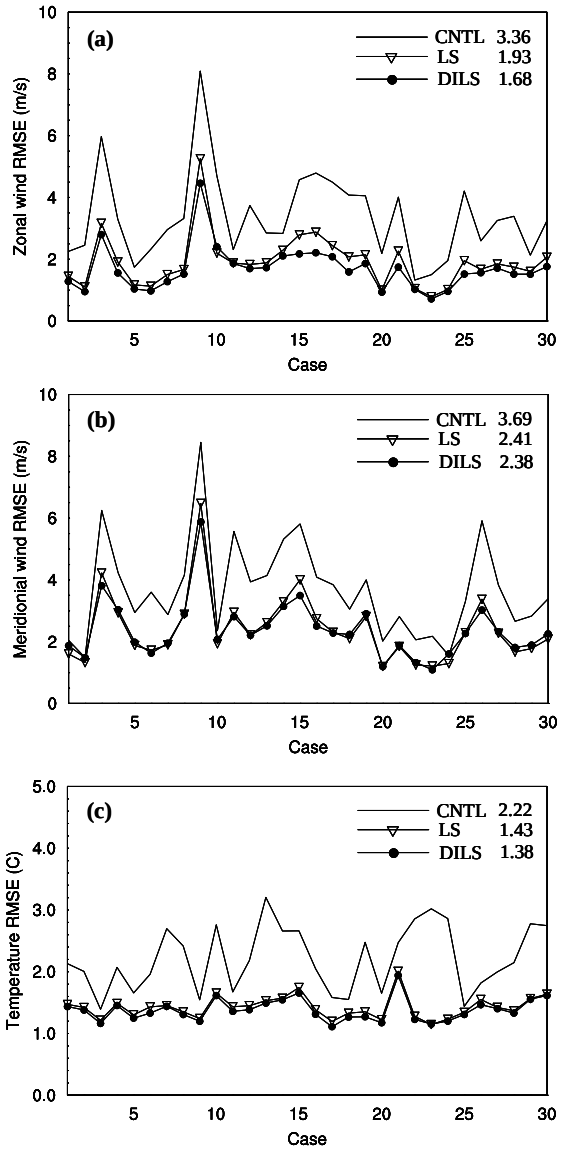

Figure 5.20 (a) RMSE errors of the 3DVAR analysis for (a) zonal wind, (b) meridional wind and (c) temperature against AWS surface observations for a one-month period (July 2006); CNTL (solid line), LS (solid line with down-triangles), and DILS (solid line with dots).

As mentioned above, the AWS surface data provide the various variables. Therefore, in this study, the sensitivities of variables for AWS observations on this heavy rainfall case are also investigated. The sensitivity tests included the impact of winds (hereafter referred to as UV) versus temperature and relativity humidity (hereafter referred to as TRH) from the AWS measurement. Figure 5.21 shows the 12 h accumulated rainfall amounts from 2100 UTC 11 July to 0900 UTC 12 July 2006 for UV and TRH experiment. Comparing the rainfall distribution of different experiments as shown in Fig. 5.3 leads to two characteristics in the heavy rainfall forecast. First, assimilation of wind variables alone (UV) resulted in similar rainfall distribution to assimilation of all variables (AWS in Fig.5.3). As shown in Fig. 2.2, background errors for winds were larger than observation errors. Thus, the analysis from the surface data assimilation inclined toward the observation more than the background ($\sigma^o < \sigma^b$) and hence produced large wind analysis increments. While observation errors for temperature and relative humidity were larger than background errors, and the analysis produced only small analysis increments. That seemed why rainfall distribution in UV was similar to that in AWS experiment. Second, assimilation of the wind variables produced better agreement with the observational results than assimilation of thermodynamic variables. When the wind variables were added to the data assimilation, the southwesterly flow over the western coast of the central Korean Peninsula was enhanced as compared with the other experiments, which was an important factor influencing

continuous storm development (not shown here). While, the rainfall distribution by temperature and relative humidity assimilation was shifted southward, which was similar to that in CNTL. Thus, it was noted that in this case, the wind variables in the surface data assimilation assisted in better positioning and amount of the rainband.

In order to further investigate the impact of surface observations, the adjoint-based observation impact is evaluated using WRF and WRF 3DVAR system. For optimal use of surface observations and forecast skill improvement, it is important to evaluate and monitor the impact of the surface observations in the numerical weather prediction system. Baker and Daley (2000) showed that the sensitivity of the analysis to observations could be computed using the adjoint of analysis system. And Langland and Baker (2004) showed that the adjoint of a data assimilation system could be used to measure the impact of observation on forecast skill. Thus, in this study, we investigated the impact of surface observations on the forecast skill using the WRF adjoint model. While nonlinear forecast model is integrated with full physics packages including moist subgrid processes, the adjoint model is integrated with only dry physics. The surface observations, wind speed and direction, temperature, and relative humidity, were assimilated in the WRF 3DVAR. To calculate the observation impact, a forecast error (e) from an analysis x_a and background x_b was calculated based on total dry energy.

$$e = (x^f - x^t)^T C (x^f - x^t) \tag{9}$$

where x^t is analysis state and C is a symmetric, usually diagonal matrix of weights. x^f is the forecast from the analysis and background field, respectively. The difference $\delta e = e_a^f - e_b^f$ measures the combined impact of all observations assimilated at the given time and $\delta e < 0$ indicates that the observations improve the forecast skill or reduce the forecast error. In this study, the observation impact in reducing 6 hour forecast error was evaluated from 0000 UTC 11 to 1200 UTC 14 July 2006, with variant formula of third-order approximation of forecast error variation. Figure 5.22 shows the time-averaged surface observation impact in the 6 km horizontal resolution grid. In the results, the water vapor mixing ratio showed the negative value, however, the observation impact of water vapor mixing ratio was very small compared with the other variables. This seems because the relative humidity is less reported than the other variables. Among the observation variables, wind observations showed larger impact in improving the 6 hour forecast than the thermodynamic observations in this modeling system framework, which was consistent with the results of Fig. 5.21. And it was noted that the meridional wind showed the largest impact in reducing 6 hour forecast error.

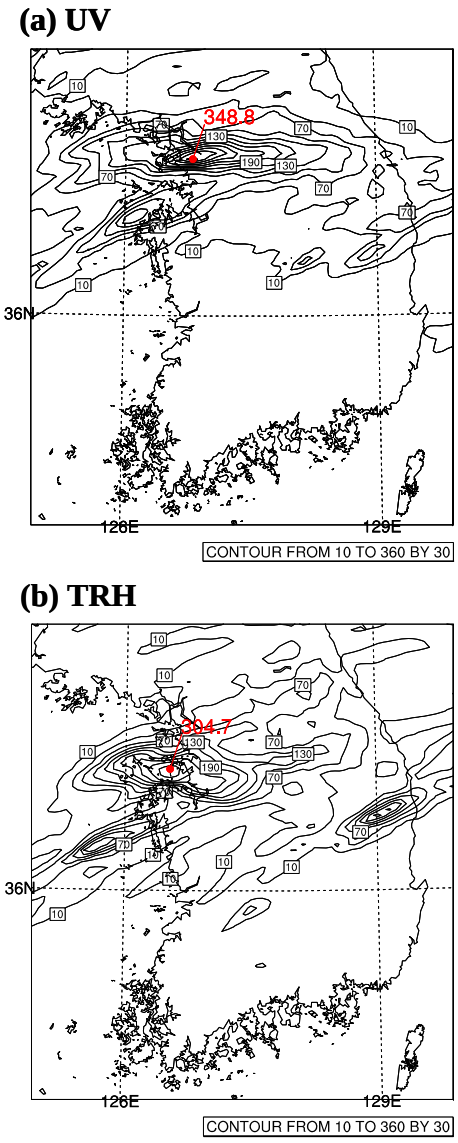

Figure 5.21 Total accumulated 12 h rainfall (mm, lines) from 2100 UTC 11 July to 0900 UTC 12 July for (a) UV and (b) TRH experiment.

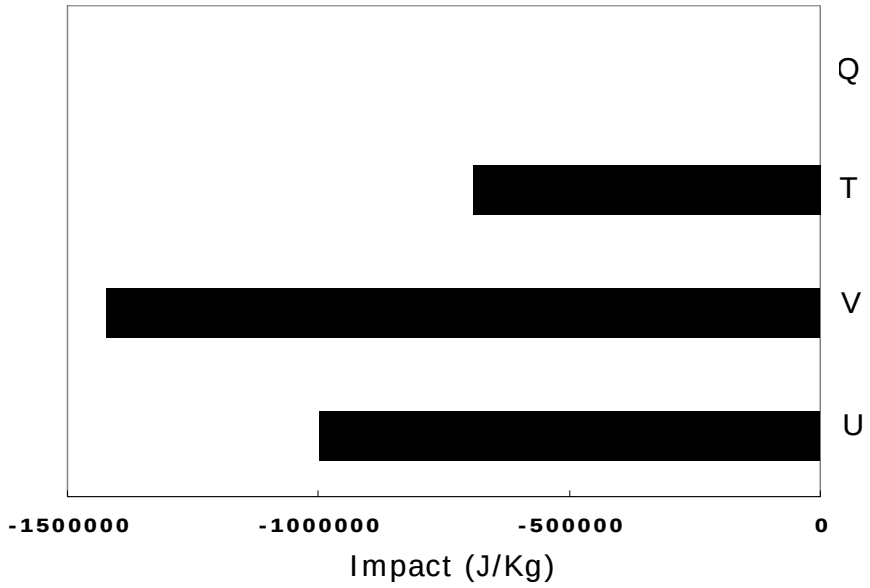

Figure 5.22 Time-averaged observation impacts for surface observation variables from 0000 UTC 11 to 1200 UTC 14 July 2006.

6. Summary and conclusions

This study investigated the impact of the multiple-Doppler radar and AWS surface data assimilation on the heavy rainfall forecast, and tuning of the length scale of the background error statistics with the surface data assimilation. To evaluate the impact of data assimilation, we selected the heavy rainfall case that the operational forecasts failed to predict the amount of the precipitation. The selected heavy rainfall case was a band-shaped torrential rainfall associated with MCSs, which resulted in the total rainfall amount greater than 330 mm on 11-12 July 2006 in the Goyang area of the central Korean Peninsula, located approximately 13 km northwest of Seoul. In the analysis of Doppler radar data, MCSs responsible for the heavy rainfall in Goyang were characterized by the back-building formation of convective storms, which were successively formed upstream of the rainband.

Numerical simulations were carried out using a three-dimensional non-hydrostatic mesoscale model and its three-dimensional variationial data assimilation with the radar and surface data. In addition, to reduce the noise of spurious high-frequency gravity waves, IAU method was applied for the data assimilation. Using the IAU method, the gravity wave fluctuation was greatly reduced and the noise was effectively removed, which help to reduce aliasing in subsequent analyses. The simulation without data assimilation failed to capture the precipitation location, while the simulations with assimilation using radar and surface

data showed better agreement with the observations. The assimilation of both radar and surface data showed the best agreement with the observations in terms of location and amount of rainfall, and had a more positive impact on the QPF than the assimilation of either radar data or surface data only. In addition, the improvement in the QPF with the assimilation of both radar and surface data was observed more clearly for heavy rainfall than for light rainfall. Based on the data assimilation experiments, we found that the simulation using radar data contributed to the development of convective storms in the early hours of forecast, and the simulation using surface data led to the establishment of the low-level environment that was favorable for convection by enhancing the thermal gradient, and modulating the frictional velocity and height of the PBL over the central Korean Peninsula. These results demonstrate that the prediction of heavy rainfall can be potentially improved through the data assimilation of radar and surface observations.

In order to improve the assimilation of surface data, the length scale of the recursive filter in the background error was tuned by observation minus background (O-B) statistics. This was necessary because the NMC method tended to exaggerate the length scale that determined the shape and extent to which observed information spread out for the assimilation of high-resolution surface data. A comparison revealed that the length scale of the NMC method should be halved in order to better assimilate the surface data with that of O-B. However, even using half of the original length scale, the correlation between the NMC method and O-B

statistics was still poor. To solve this problem, we applied a double iteration method with two different scales representing large and small lengths.

The results of the data assimilation using an ideal field clearly showed that the assimilation using a length scale with tuning factors was able to contain small-scale features effectively, as the horizontal influence extension of the analysis increments was reduced. The two-dimensional wave number distribution was calculated by a two-dimensional Fast Fourier Transform (FFT). This showed that the assimilation using the double iteration with two different length scales improved the reflection of the large- and small-scales features from the observed information to the model fields, allowing the 3DVAR system to extract high-resolution observed information more effectively. In the results of the heavy rainfall forecasts, the results of the data assimilation showed that length scale tuning of the recursive filter improved the low-level wind circulation, and subsequent precipitation forecasts, compared to the simulation with the original length scale of the NMC method. The strengthened low-level winds forecast over the west coast of the central Korean Peninsula enhanced the transport of warm and moist air towards the heavy rainfall region. This consequently enhanced low-level convergence, so that the simulations with tuned length scales could identify the location of the observed rainfall, although the maximum rainfall amount was underestimated compared to observational data. The simulation using the double iteration with two different length scales not only reproduced

the band-shaped rainfall area over the central Korean Peninsula, but also the maximum rainfall amount (321.4 mm). By using a suitable strategy to assimilate surface data that allowed the representation of large- and small-scale features in the model fields, the assimilation using the double iteration with two different length scales contributed to the better development of convective cells near the observed heavy rainfall region. We assessed the impact of the background error tuning on the improvement of the analysis through the assimilation of surface data during a one-month period, in comparison with the background error estimated by the NMC method. The statistics results for the one-month period indicated that the analysis using the double iteration with two different length scales improved the RMSEs verified against the surface observations, particularly in zonal wind analysis.

Regarding the surface observations, the adjoint-based surface observation impact was evaluated for optimal use of surface observation and forecast skill improvement. The observation impact in improving 6 hour forecast was evaluated with variant formula of third-order approximation of forecast error variation. The results indicated that wind observations showed larger impact in improving the 6 hour forecast than thermodynamic observations. In addition, the meridional wind showed the largest impact in reducing 6 hour forecast error.

References

Augustine, J. A., and F. Caracena, 1994: Lower-tropospheric precursors to nocturnal MCS development over the Central United State. *Wea. Forecasting*, **9**, 116-135.

Baker, N. L., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.*, **126**, 1431-1454.

Baker, W. E., S. Bloom, H. Woollen, M. Nestler, E. Brin, T. Schlatter, and T. Branstrator, 1987: Experiments with a three-dimensional statistical objective analysis scheme using FGGE data. *Mon. Wea. Rev.*, **115**, 272-296.

Barker, D. M., W. Huang, Y.-R. Guo and Al Bourgeois, 2003: A three-dimensional variational (3DVAR) data assimilation system for use with MM5. NCAR Tech. Note NCAR/TN-453+STR. 68pp.

_____, and Q. Xiao, 2004: A three-dimensional variational (3DVAR) data assimilation system for use with MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897-914.

Bloom, S. C, L. L Takacs, A. M. da Silva, and D. Ledvina, 1996: Data Assimilation Using Incremental Analysis Updates. *Mon. Wea. Rev.*, **124**, 1256-1271

Bluestein, H. B., and M. H. Jain, 1985: Formation of mesoscale lines of Precipitation: Severe squall lines in Oklahoma during the spring. *J. Atmos. Sci.*, **42**, 1711-1732.

Burgess, D., and P. S. Ray, 1986: Principle of radar. *Mesoscale Meteorology and Forecasting*, P. S. Ray, Ed., Amer. Meteor. Soc., 85-117.

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O., 2004: Properties and first application of an error statistics tuning method in variational assimilation. *Q. J. R. Meteorol. Soc.*, **130**, 2253-2275.

Chun, J. K., and M. L. Oh, 1991: Presumption of the horizontal wind and divergence field using a single Doppler radar data. *J. Kor. Meteor. Soc.*, **27**, 55-66. (in Korean)

Clayton, A., 2003: Incremental analysis Update (IAU) scheme. Unified Model Documentation Pater 31, UK Met Office, 20pp.

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

Cressman, G. P., 1959: An Operational Analysis Scheme. *Mon. Wea. Rev.*, **87**, 367-374.

Crook. N. A., J. Sun, 2002: Assimilating Radar, Surface and Profiler Data for the Sydney 2000 Forecast Demonstration Project, *J. Atmos. Oceanic Technol.* **19**., 888-898.

Daley, R., 1979: The application of nonlinear normal mode initialization to an operational forecast model. Atmosphere. *Atmos.-Ocean*, **17**, 97-124.

Daley R, 1991. Atmospheric Data Analysis. Cambridge University Press.

Desroziers, G. and Ivanov, S. 2001: Diagnosis and adaptive tuning of information error parameters in a variational assimilation. *Q. J. R.*

Meteorol. Soc., **127**, 1422-1452.

Doswell, C. A., III, Ed., 2001: *Severe Convective Storms. Meteor. Monogr*., No. 50, Amer. Meteor, Soc., 570 pp.

ERS, 2004: RAOB version 5.5 user's guide and technical manual. Environmental Research Services, 102 pp.

Funk, T., cited 2004: Heavy convective rainfall forecasting: Parameters, progresses, patterns, and rule of thumb. Science and Technology, NWS, Louisville, KY, 38 pp. [Available online at http://www.crh.noaa.gov/lmk/soo/presentations/heavy_rainfall_forcasting.pdf]

Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variation method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev*., **127**, 2128-2142.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Tech. Note NCAR/TN-398 +STR, 122pp.

Gu, J., Q. Xiao, Y.–H. Kuo, D. M. Barker, J. Xue, and X. Ma, 2005: Assimilation and simulation of typhoon Rusa (2002) using the WRF system. *Adv. Atmos. Sci*., **22**, 415-427.

Guo, Y. R., H.-C. Lin, X. X. Ma, X.-Y. Huang, C. T. Terng, and Y.-H. Kuo, 2006: Impact of WRF-Var (3DVar) Background Error Statistics on Typhoon analysis and Forecast. *WRF user's workshop*, Boulder, Colorado, 19-22pp.

Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis

of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3-15.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111-136.

Hong, S.-Y., 2004: Note and correspondence, comparison of heavy rainfall mechanisms in Korea and the central US. *J. Meteor. Soc. Japan*, **82**, 1469-1479

Hong, S.–Y., J. Dudhia, and S.–H. Chen, 2004: A Revised Approach to Ice Microphysical Processer for the Bulk Parameterization of Clouds and Precipitation. *Mon. Wea. Rev.*, **132**, 103-120.

Ingleby, N. B., 2001: The statistical structure of forecast errors and its representation in the Met Office Global 3-D Variational Data Assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **127**, 209-231.

Jarvinen, H., 2001: Temporal evolution of innovation and residual statistics in ECMWF variational data assimilation systems. *Tellus*, **53A**, 333-347.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/ detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2748-2802.

____, and ____, 1993: Convective parameterization for mesoscale models: The Kain-Fritcsh scheme. *The representation of cumulus convection in numerical models*, K. A.

Kato, T., 1998: Numerical simulation of a band-shaped torrential rain

observed over southern Kyushu, Japan on 1 August 1993. *J. Meteor. Soc. Japan*, **76**, 97-128.

Kim, H. W., and D. K, Lee, 2006: An Observational Study of Mesoscale Convective Systems with Heavy Rainfall over the Korean Peninsula. *Wea. Forecasting*, **21**, 125-148.

KMA (Korea Meteorological Administration), 2001: Climatological standard normals of Korea (1971-2000).

Langland, R. H., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-203.

Lee. D. K., H. R. Kim, and S. Y. Hong, 1998: Heavy rainfall over Korea during 1980-1990. *Kor. J. Atmos. Soc*., **1**, 32-50.

Lee, J. H., H. H. Lee, Y. C. Choi, H. W. Kim, and D. K. Lee, 2010: Radar data assimilation for the simulation of mesoscale convective systems, *Adv. Atmos. Sci.*, **27**, 1025-1042.

Lee, M. S., and D. K. Lee, 2003: An application of a weakly constrained 4DVAR to satellite data assimilation and heavy rainfall simulation. *Mon. Wea. Rev*., **131**, 2151-2176.

Lee, M. S., D. M. Barker, and Y.-H. Kuo, 2006: Background Error Statistics using WRF Ensemble Generated by Randomized Control Variables. *J. Kor. Meteor. Soc.*, **42(3)**, 153-167.

Lee, M. S., Y.–H. Kuo, D. M. Barker, and Eunha Lim, 2006: Incremental analysis updates initialization technique applied to 10-km MM5 and MM5 3DVAR. *Mon. Wea. Rev*., **134**, 1389-1404.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177-1194.

Lorenc, A. C., and Coauthours, 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **126**, 2991-3012.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019-1034.

National Emergency Management Agency, 2009: Disaster yearbook, Korea, 755pp.

NWS, cited 2004: Convective season environmental parameters and indices. Science and Technology, NWS, Louisville, KY. [Available online at http://www.crh.noaa.gov/lmk/soo/docu/indices.php.]

Park, S. G., and D. K. Lee, 2009: Retrieval of High-Resolution Wind Fields over the Southern Korean Peninsula Using the Doppler Weather Radar Network. *Wea. Forecasting*, **24**, 87-103.

Parker, M. D., 2007: Simulated Convective Lines with Parallel Stratiform Precipitation. Part I: An Archetype for Convection in Along-Line Shear. *J. Atmos. Sci.*, **64**, 267-288.

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

Polavarapu, S., S. Ren, A. M. Clayton, D. Sankey, and Y. Rochon, 2004: On the relationship between incremental analysis updating and

incremental digital filtering. *Mon. Wea. Rev.*, **132**, 2495-2502.

Puri, K., 1985: Sensitivity of low-latitude velocity potential fields in a numerical weather-prediction model to initial conditions, initialization, and physical processes. *Mon. Wea. Rev.*, **113**, 449-466.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian convariacnes. *Mon. Wea. Rev.*, **131**, 1524-1535.

Rinehart, R. E., 1997: *Radar of Meteorologists*. Rinehart Publications, 3rd ed., 166-167.

Rochette, S. M., and J. T. Moore, 1996: Initiation of an elevated mesoscale convective system associated with heavy rainfall. *Wea. Forecasting*, **11**, 443-457.

Schumacher, R. S., and R. H. Johnson, 2005: Organization and Environmental Properties of Extreme-Rain-Producing Mesoscale Convective Systems. *Mon. Wea. Rev.*, **133**, 961-976.

Serafin, R., and J. W. Wilson, 2000: Operational weather radar in the United States: Progress and opportunity. *Bull. Amer. Meteor. Soc.*, **81**, 501-518.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Grill, D. M. Barker, W. Wand, and J. G. Powers, 2008: A description of the advanced research WRF version 3. NCAR Tech. Mote NCAR/TN-475 +

STR, 113 pp.

Stensrud, D. J., R. L. Grall, and M. J. Nordquist, 1997: Surges over the Gulf of California during the Mexican monsoon. *Mon. Wea. Rev.*, **125**, 417-437.

Sun J., and T.–Y. Lee, 2002: A numerical Study of an Intense Quasi-stationary convection Band over the Korean Peninsula. *J. Meteor. Soc. Japan*, **80**, 1221-1245.

Tanguay, M., Bartello, P. and Gauthier, P., 1995: Four-dimensional data assimilation with a wide range of scales. *Tellus*, **47A**, 974-997.

Tao, W.–K. and S.–T. Soong, 1991: Numerical simulation of a subtropical squall line over Taiwan Strait. *Mon. Wea. Rev.*, **119**, 2699-2723.

Thépaut, J.-N. and Courtier, P., 1991: Four-dimensional data assimilation using the adjoint of a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.*, **117**, 1225-1254.

Trier, S. B., and E. B. Parsons, 1993: Evolution of environmental conditions preceding the development of a nocturnal mesoscale convective complex. *Mon. Wea. Rev.*, **121**, 1078-1098.

Wee, T. K., 1999: A cumulus parameterization based on cloud simulations of mesoscale convective systems (in Korean). Ph. D. thesis, Seoul National University, Seoul, South Korea, 308 pp.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504-520.

Wu, W.-S., J. Purser, and D. E. Parrish, 2002: Three-dimensional

variational analysis with spatial inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905-2916.

Xiao, Q., Y.–H. Kuo, J. Sun, W.–C. Lee, Eunha Lim, Y.–R. Guo, and D. M. Barker, 2005: Assimilation of Doppler radar observations with a regional 3D-VAR system: Impact of Doppler velocities on forecasts of a heavy rainfall case. *J. Appl. Meteor.*, **44**, 768-788.

Xiao, Q., and J. Sun, 2007: Multiple-Radar Data Assimilation and Shor-Range Quantitative Precipitation Forecasting of a Squall Line Observed during IHOP_2002. *Mon. Wea. Rev.*, **135**, 3381-3404.

Zipser, E. J., 1982: Use of a conceptual model of the life-cycle of mesoscale convective systems to improve very short-range forecasts. *Nowcasting*, K. Browning, Ed., Academic Press, 191-204.

Zou, X., F. Vandenberghe, M. Pondeca, and Y.-H. Kuo, 1997: Introduction to adjoint techniques and the MM5 adjoint modeling system. NCAR Tech. Note NCAR/TN-435 + STR, 110 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO, 80307.]

초록

본 연구에서는 집중호우의 예보 향상을 위해 레이더 및 AWS 지표 관측 자료의 동화 영향을 살펴보았으며, 이를 위해 WRF와 WRF 3DVAR 시스템을 이용하였다. 또한 자료 동화 시 발생하는 고주파의 중력파를 제거하기 위하여 IAU 방법을 적용하였으며, 백빌딩 (back-building) 중규모 대류계에 의해 발생한 2006년 7월 11-12일의 집중호우를 대상으로 각 관측 자료 동화가 강수 예보에 미치는 영향을 살펴보았다. IAU 방법은 중력파 변동을 현저하게 감소시키고 노이즈를 효과적으로 제거함으로 분석장을 향상시키는데 도움을 주었다. 먼저 레이더와 지표 관측 자료 동화의 효과를 살펴보기에 앞서, 레이더와 지표 관측 자료를 동화한 실험이 집중호우의 예측을 향상시키는지를 살펴보았다. 레이더와 지표 관측 자료를 동시에 동화한 실험은 강수 강도와 위치에 있어 관측과 유사한 결과를 모의하였으며, 정량적인 검증에 있어서도 각 관측 자료를 동화한 실험에 비해 긍정적인 효과를 나타내었다. 또한 백빌딩 중규모 대류계의 특징을 잘 모의하였다. 자료 동화 실험 결과를 바탕으로, 레이더 관측 자료의 동화는 모형의 초기 시각에 집중 호우를 유발하는 중규모 대류계 발달에 긍정적인 영향을 주고, 지표 관측 자료 동화는 강화된 하층 바람의 생성에 영향을 준다는 결과를 얻어내었다. 또한 지표 관측 자료는 하층의 온도 경도를 강화시키고 행성경계층을 변화시켜 하층에 대류가 발생하기 좋은 조건을 형성하는데 중요한 역할을 하였다. 이러한 결과들은 레이더 및 AWS 지표 관측 자료 동화가 모형의 집중호우 예보 능력 향상에 기여할 수 있다는 가능성을 제시한다.

고해상도의 지표 관측 자료를 동화함에 있어, NMC 방법으로 계산된 배경오차는 관측 정보의 전파와 형태를 결정하는 길이규모를 과장되게 표현하는 경향이 있다. 따라서 효과적인 지표 관측 자료 사용으로 집중호우의 예측을 향상시키기 위해, NMC 방법으로 계산된 배경오차의 상관도와 관측과 배경장의 차이인 O-B 상관도를 비교하여 NMC 방법으로 계산된 배경오차의 길이규모를 조절하였다. 비교를 통해, 보다 효과적으로 지표 관측 자료를 동화하기 위해서는 NMC 방법으로 계산된 길이규모를 반으로 줄여야 한다는 결과를 얻어내었다. 하지만 그럼에도 불구하고 O-B 와 NMC 방법의 상관도 형태에서는 차이가 나타났으며, 이를 해결하기 위하여 본 연구에서는 큰 규모와 작은 규모로 표현되는 두 개의 길이규모를 적용하여 자료 동화하는 방법 (DILS) 에 대해 살펴보았다.

이상화 실험에서, 길이규모를 조절한 실험은 NMC 방법으로 계산된 배경오차를 이용한 실험에 비해 관측 정보의 작은 규모를 효과적으로 나타내었다. 또한, DILS 는 자료 동화 시스템이 고해상도 지표 관측 정보를 효과적으로 얻어내는 것을 가능하게 함으로써 NMC 방법보다 관측 정보를 모형에 보다 효과적으로 반영하는 결과를 보였다. 집중 호우 사례에 대해, DILS 실험은 길이규모를 조절한 후 한 번의 자료 동화를 수행한 실험에 비해 강수 분포와 양에 있어 관측과 유사한 결과를 나타내었다. 이는 강화된 하층 바람과 이와 연관된 수렴에 의해 집중 호우가 발생했던 지역에 중규모 대류계를 잘 모의하였기 때문이다. 이러한 결과들은 자료 동화 시스템에서 적절하게 계산된 배경오차를 이용함으로써 고해상도 AWS 지표 관측 자료의 활용성을 극대화하고 이를 통해 중규모 수치 모형에서

집중호우의 예측 성능을 향상시킬 수 있음을 제시한다. DILS 방법을 적용한 배경오차가 중규모 예보 모형에 미치는 영향을 평가하기 위하여 1 개월의 기간에 대하여 지표 관측 자료 동화 실험을 수행하고, 초기 시각의 분석장을 지표 관측 자료에 대하여 평균제곱근 오차를 계산하여 검증하였다. 두 번의 길이규모를 적용한 배경오차를 이용하여 지표 관측 자료의 동화 실험에 적용한 결과, 온도장의 분석결과는 NMC 방법을 적용한 분석장과 유사하게 나타났지만, 동서 바람장의 분석오차는 현저하게 감소하는 결과를 보였다.

지표 관측 자료 동화에 더해 여러 변수를 제공하는 AWS 지표 관측 자료가 예보장에 미치는 영향을 살펴보기 위해 WRF 수반모형을 이용하여 6시간 예보장에 대한 지표 관측 자료의 영향에 대해 살펴보았다. 그 결과, 온도와 상대 습도의 열역학 변수보다는 바람 자료가 예보장을 향상시킨다는 결과를 얻어냈으며, 바람 변수 중에서도 남북 바람이 예보장의 오차를 감소시키는데 큰 역할을 한다는 결과를 얻어내었다. 이러한 결과들은 지표 관측 자료의 변수를 효과적으로 사용함으로써 중규모 수치 모형의 예측 성능을 향상시킬 수 있음을 제시한다.

주요어: 레이더와 지표 관측 자료, 자료 동화, 3차원 변분 자료 동화, 중규모 대류계, 집중 호우, 조절

학 번: 2005-30149

감사의 글

길었던 학위 과정을 마무리하고 이제 새로운 출발점 위에 서게 되었습니다. 논문의 마무리를 글로 적으려니 그 동안 많은 분들의 도움으로 논문을 마무리 할 수 있었다는 것을 새삼 느끼게 됩니다. 학위과정 동안 과학자로서 연구자의 자세를 일깨워 주셨으며, 이 논문이 완성될 수 있도록 지도해주신 이동규 교수님께 깊이 감사 드립니다. 학위 논문 완성을 위해 격려와 조언을 아끼지 않으시고 논문 지도를 해주신 임규호 교수님께도 깊은 감사의 뜻을 전합니다. 논문 심사위원장으로서 연구 내용에 귀중한 조언들을 해주신 백종진 교수님과 논문의 부족한 점을 채워주신 전종갑 교수님께 감사 드립니다. 연구 결과를 꼼꼼하게 살펴봐주시고 연구 내용의 향상을 위해 조언을 해주신 장동언 박사님께도 깊이 감사 드립니다. 학문적 견해를 넓혀주시고 논문 심사를 위해 시간을 내주신 임은하 박사님께 깊은 감사의 말씀을 드리고 싶습니다. 석사, 박사 과정 동안 대기과학이라는 학문에 대한 이해를 넓혀 주신 박순웅, 윤순창, 손병주, 강인식, 최우갑, 허창회 교수님께도 감사 드립니다. 또한 학부 때 처음으로 접한 대기과학 학문에 흥미를 갖게 해주시고 늘 관심을 가지고 지켜봐 주신 변희룡 교수님께도 감사의 말씀을 드립니다.

석사, 박사 과정의 긴 시간을 보낸 수치예보 실험실의 선·후배님들께 진심으로 감사하다는 말을 전하고 싶습니다. 이승우 선배님, 강현석 선배님, 김형우 선배님, 박정균 선배님, 이조한 선배님, 하종철 선배님, 조정훈 선배님, 김춘지 선배님, 차동현 선배님, 노준우 선배님, 김주완 선배님, 이현하 선배님 감사합니다. 선배님들의 도움과

격려로 논문을 마무리 할 수 있었습니다. 옆에서 힘이 되어준 용한이와 천실, 민정, 현영, 대용, 경나, 경동, 재복에게도 감사의 마음을 전합니다. 그리고 1년 동안 힘이 되어준 종관규모 기상학 실험실의 신우, 남규, 의현, 은호, 상원씨에게도 고마운 마음 잊지 않겠습니다. 많은 시간을 보내면서 격려와 위로를 주었던 03학번 대학원 동기들에게도 고마운 마음을 전합니다.

끝까지 학위를 마칠 수 있도록 지켜봐 주신 가족, 여러 친지 분들의 사랑과 관심에 감사 드립니다. 특히 힘든 일이 있을 때마다 버팀목이 되어주시고 부족한 자식을 믿어주신 부모님께 진심으로 감사 드립니다. 그리고 사랑합니다. 학위 과정 동안 옆에서 많은 응원의 말씀을 해주신 시부모님께도 감사의 마음을 전합니다. 항상 저를 응원해주고 일상의 고민을 들어주었던 형부와 언니, 동생에게도 감사의 마음을 전합니다.

이 논문을 완성하는 데에는 무엇보다도 멋진 남편 최석진 박사의 응원과 격려가 큰 힘이 되었습니다. 진심으로 사랑하고 감사하다는 말을 전하고 싶습니다. 마지막으로 앞으로 태어나게 될 똘똘이와 함께 기쁨을 나누고 싶습니다.