Retrieval of Moisture from Slant-path Water Vapor Observations of a Hypothetical GPS Network Using a Three-Dimensional Variational Scheme with Anisotropic Background Error

Haixia Liu and Ming Xue[*]

School of Meteorology and Center for Analysis and Prediction of Storms
University of Oklahoma, Norman OK 73019

[*]Corresponding Author Address:
Dr. Ming Xue
School of Meteorology
University of Oklahoma
100 E. Boyd
Norman OK 73019
mxue@ou.edu

Abstract

A three-dimensional variational (3DVAR) scheme is developed for retrieving three-dimensional moisture in the atmosphere from slant-path measurements of a hypothetical ground-based Global Positioning System (GPS) observation network. We assume that the observed data are in the form of slant-path water vapor (*SWV*), which is the integrated water vapor along the slant path between the ground receiver and GPS satellite. The inclusion of a background in our analysis overcomes the under-determinedness problem. An explicit Gaussian-type spatial filter is used to model the background error covariances which can be anisotropic. As a unique aspect of this study, an anisotropic spatial filter based on flow-dependent background error structures is implemented and tested and the filter coefficients are derived from either true background error field or from the increment of an intermediate analysis that is obtained using an isotropic filter. In the latter case, an iterative procedure is involved.

A set of experiments is conducted to test the new scheme with hypothetical GPS observations for a dryline case that occurred during the 2002 International H_2O Project (IHOP_2002) field experiment. Results illustrate that this system is robust and can properly recover three-dimensional mesoscale moisture structures from GPS *SWV* data and surface moisture observations. The analysis captures major features in water vapor associated with the dryline even when an isotropic spatial filter is used. The analysis is further improved significantly by the use of flow-dependent background error covariances modeled by an anisotropic spatial filter.

Sensitivity tests show that surface moisture observations are important for the analysis near ground, and more so when flow-dependent background error covariances are not used. Vertical filtering is necessary for obtaining accurate analysis increments. The retrieved moisture

field remains reasonably accurate when the surface moisture observations and GPS *SWV* data contain errors of typical magnitudes. The positive impact of flow-dependent background error covariances increases when the density of ground-based GPS receiver stations decreases.

1. Introduction

Accurately characterizing the three-dimensional (3D) distribution of water vapor in the atmosphere is very important for the understanding and prediction of mesoscale and storm-scale weather, especially with regard to quantitative precipitation forecast (QPF, Emanuel et al. 1995). QPF skills have been improving rather slowly owing to the high spatial and temporal variability of water vapor. Thus, high-resolution observations of 3D water vapor have the potential to significantly improve the prediction of precipitation and severe weather. In recent years, space- and ground-based Global Positioning System (GPS) receivers have seen significant development and can potentially provide water vapor measurements with high resolution under virtually all weather conditions (Businger et al. 1996; Ware et al. 2000; Wolfe and Gutman 2000; Bengtsson et al. 2003).

For ground-based GPS receivers, the raw measurement is the total amount of delay to the microwave radio signals caused by the atmosphere as the signals travel from the GPS satellite to the receiver. The total delay along the slant path between the receiver and satellite is composed of ionospheric delay, hydrostatic delay and wet delay. Upper atmosphere ionospheric delay, when observed by a dual-frequency GPS receiver, can be calculated to millimeter accuracy. The hydrostatic delay can be estimated with known knowledge of pressure and temperature. The slant-path wet delay (*SWD*), the residual part due to water vapor, can therefore be obtained by subtracting the ionospheric and hydrostatic delays from the total delay. Further, *SWD* can be linearly related to slant water vapor (*SWV*), defined as the integral of water vapor along the entire slant path (Davis et al. 1985). The accuracy of *SWV* thus derived is usually within a few millimeters (Ware et al. 1997; Braun et al. 2001).

Most of the past data assimilation or data impact studies related to ground-based GPS

water vapor observations deal with the precipitable water (*PW*) or zenith wet delay (*ZWD*) data. (e.g., Kuo et al. 1993; Kuo et al. 1996; Guo et al. 2000; Falvey and Beavan 2002). *ZWD* is a derived product that is obtained by projecting *SWD* observations onto the zenith then averaging all of them over a certain time period under the assumption of azimuthal isotropy and horizontal homogeneity (Rocken et al. 1993). Further, *PW* is derived from *ZWD* in a similar way as *SWV* is from *SWD*, i.e., it is linearly related to *ZWD*. A more recent study of Cucurull et al. (2004), using a 3DVAR scheme, directly assimilates zenith total delay (*ZTD*) which is also a derived product similar to *ZWD*, except it is for total delay. Simulated and real *ZTD* data were assimilated via a four-dimensional variational (4DVAR) scheme into a mesoscale model by De Pondeca and Zou (2001a; 2001b), respectively. These studies have all found a positive impact of assimilating the GPS data on precipitation forecast, though in some cases, the impact is small.

There exists a significant loss of information in *ZWD* or *ZTD* data, however, compared to the original slant-path data due to the spatial and temporal averaging involved in their derivation. It should, therefore, be beneficial to use the slant path total or wet delay or slant-path water vapor data directly. We do point out here that since the slant-path measurement represents an integrated quantity of water vapor along each ray path, it, as in the case of *PW* and *ZWD*, does not provide information on the vertical distribution of water vapor. It is hypothesized here that in the case of slant-path data (in contrast to *PW* and *ZWD* data), the multiple overlapping ray paths, forming a 'net' covering the atmosphere, are helpful in recovering the 3D structure of moisture. In the case of variational analysis, the inclusion of the analysis background, a good knowledge of the background error structure as well as the effective utilization of such knowledge, should all be helpful. It is the goal of this study to show that these hypotheses are true.

Currently, over the United States, there are approximately 125 surface GPS sites for

which precipitable water observations are available in near real time. These sites include those from the NOAA Forecast Systems Laboratory (FSL) network and the SuomiNet (Ware et al. 2000). Slant-path GPS data are available with several day latency from more than 20 sites but the real time availability from over 100 sites is planned. The potential availability of much more slant path data and their decent accuracy have prompted interests in analyzing and assimilating such data directly into numerical models. Limited number of existing studies include MacDonald et al. (2002) and Ha et al. (2003), both of which utilize simulated data from a hypothetical GPS network. This is the case partly because real slant-path data are very limited and their spatial coverage and resolution are still poor.

In MacDonald et al. (2002), a 3D variational method is used to analyze *SWV* data collected by a hypothetical high-resolution network of ground-based GPS receivers. It is shown that the 3D moisture field can be recovered from the *SWV* data in combination with the surface moisture observations taken at the same sites as the receivers. Water vapor soundings from a low-density network are also used to help their analysis. Ha et al. (2003) showed, using a 4DVAR system, that the direct assimilation of simulated *SWD* is superior to assimilating the derived *PW* data in terms of both recovering water vapor information and short-range precipitation forecasting.

In this study, we focus on the analysis of slant-path water vapor (*SWV*) data (instead of *PW* data) and choose to develop and use for the analysis a more complete 3DVAR system that employs an anisotropic spatial filter for modeling the flow- or field-dependent background error covariances. As an initial study, we perform our analysis of water vapor in a univariate 3DVAR system where we focus most of our attention on the impact of flow-dependent background error covariances on the quality of analysis. The flow-dependent background error is modeled using an

anisotropic spatial filter. The use of anisotropic spatial filter is a unique aspect of this study. To our knowledge, it has never been applied specifically to the analysis of GPS data. Since the GPS data are not point measurements of the analysis variables themselves, but rather integrated quantities, the variational approach is a natural choice.

As in MacDonald et al. (2002) and Ha et al. (2003), we also use simulated data to test our analysis system. The use of simulated data from a hypothetical receiver network gives us flexibility and complete control over the network design, data resolution and quality. Further, the knowledge of the truth allows us to unambiguously assess the quality of analysis. The roles of observation system simulation experiments (OSSE) in the design, development and evaluation of future observing systems are discussed in detail by Lord et al. (1997) and also by Atlas (1997). Another reason for using simulated data is that a high-density GPS receiver network with large spatial coverage is not yet available, over the United States at least. Ultimately, however, the goal of improving moisture analysis is to improve numerical weather prediction (NWP), especially the prediction of precipitation. The impact of simulated and real GPS slant-path data analyzed through our 3DVAR procedure on short-range precipitation forecast will be the subject of future studies.

This paper is organized as follows: section 2 introduces our 3DVAR analysis system and section 3 describes the generation of a simulated GPS *SWV* data set from a mesoscale model forecast. Results from analysis experiments are presented in section 4. Further discussions on the effectiveness of our scheme are given in section 5 through sensitivity experiments. Summary and conclusions are given in section 6, together with a plan for future work.

2. 3DVAR analysis system

In this work, we follow the standard practice of 3DVAR data assimilation for NWP

(Lorenc 1981; Daley 1991) by including the analysis background. Thus, the cost function includes both background and observation terms. For moisture analysis, a weak non-negative-moisture constraint is also included in this cost function. The use of a background makes the problem over-determined and the analysis feasible for realistic numbers of GPS satellites and ground-based receivers, because the number of control variables, which is the number of grid point values of moisture in our case, is much greater than the number of (*SWV* plus surface) observations at any instance. Because the GPS network does not directly observe point values of water vapor, the analysis involves 'retrieving' or 'recovering' the 3D distribution from the integral observations. We therefore often refer to the analysis process as 'retrieval'.

The previous 3DVAR work of MacDonald *et al*. (2002) differs from this study by lumping together all GPS data in the analysis domain during a one hour interval and treating them as observations at the analysis time so that the number of observations is significantly increased. Because their 3DVAR system does not include a background constraint, noise problems were encountered during their initial analyses. They obtained reasonable analyses by employing a multi-grid procedure, in which the analyses were performed on two grids of different resolutions repeatedly, while the solutions were transferred back and forth between the grids many times. While such a procedure appears to work, the amount of effective smoothing imposed upon the analysis by such a procedure is difficult to assess, especially in terms of the physically meaningful background error correlations.

In this work, the spread of observation increments in space is controlled by the background error covariances, which in some experiments are spatially inhomogeneous and flow-dependent. In addition, this 3DVAR system is formulated in a general terrain-following coordinate system; therefore, its analysis can be directly used to initialize a model formulated in

the same coordinate system.

The initial cost function of our 3DVAR system is defined as,

$$J(x) = J_b(x) + J_{swv}(x) + J_{sfc}(x) + J_c(x),$$ (1)

where,

$$J_b(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b),$$ (2a)

$$J_{swv}(x) = \frac{1}{2}\left[H_{swv}(x) - SWV\right]^T \mathbf{R}_{swv}^{-1}\left[H_{swv}(x) - SWV\right],$$ (2b)

$$J_{sfc}(x) = \frac{1}{2}\left[H_{sfc}(x) - q_{v_{sfc}}\right]^T \mathbf{R}_{sfc}^{-1}\left[H_{sfc}(x) - q_{v_{sfc}}\right],$$ (2c)

$$J_c(x) = \frac{1}{2}\left(\frac{|x| - x}{2}\right)^2.$$ (2d)

In Eq. (1), cost function J is composed of four terms: background constraint term, J_b, GPS SWV observation term, J_{swv}, the term for conventional surface moisture observations, J_{sfc}, and the weak non-negative-water vapor constraint, J_c. The vector x is the control variable which in our case contains the specific humidity q_v at every grid point. The corresponding background state vector is x_b. The background term, J_b, represents the departure of the control variable from the background. $\mathbf{B}$ is the background error covariance matrix, which determines how the observational information is spread in space as well as weighted (in combination with the observation errors) in the analysis.

The GPS observation term, J_{swv}, represents the departure of the analysis, calculated from the control variable q_v through the observation operator H_{swv}, from the observations of SWV measured by the ground-based GPS receivers. The matrix $\mathbf{R}_{swv}$ is the observation error covariance matrix for SWV, which is usually assumed to be diagonal under the assumption that

observation errors are not correlated (some of the correlated errors can usually be effectively removed through bias correction procedures, see e.g., Harris and Kelly 2001). The magnitude of error variances, the diagonal elements of matrix $\mathbf{R}_{swv}$, compared to the background error variances, determines the relative weight of observation and background for the analysis. In our paper, observation error variances for *SWV* and surface observations are specified.

Since the ground-based GPS receiver sites are commonly equipped with regular meteorological sensors, regular surface water vapor observations can be made together with *SWV* observations. Therefore, J_{sfc} is included in the cost function to better analyze moisture structure near the surface. Finally, in order to avoid creating a significant amount (in relative sense) of negative water vapor, especially at high levels, a weak non-negative-moisture constraint term, J_c, is also included in the cost function.

As pointed out previously, the inclusion of a background term is significant for our 3DVAR analysis. It not only eliminates the under-determinedness problem, but also allows for more accurate analysis through the background error covariances, contained in matrix $\mathbf{B}$. But because the dimensionality of $\mathbf{B}$ is very large for typical meteorological problems, the direct inversion of $\mathbf{B}$ as required in Eq. (2a) is never attempted. Huang (2000) presents a method named variational analysis using a filter (VAF) that avoids the need for the inversion. In this method, the control variable is redefined as,

$$v = \mathbf{B^{-1}}\,(x\text{-}x_b), \tag{3}$$

which is the increment field relative to the background multiplied by the inverse of $\mathbf{B}$. Using this new control variable, the cost function is redefined as,

$$J(v) = \frac{1}{2}v^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}v + \frac{1}{2}\left[H_{swv}\left(\mathbf{B}v + x^b\right) - SWV \right]^{\mathrm{T}} \mathbf{R}_{swv}^{-1} \left[H_{swv}\left(\mathbf{B}v + x^b\right) - SWV \right]$$

$$+ \frac{1}{2}\left[H_{sfc}\left(\mathbf{B}v + x^b\right) - q_{v_{sfc}} \right]^{T} \mathbf{R}_{sfc}^{-1} \left[H_{sfc}\left(\mathbf{B}v + x^b\right) - q_{v_{sfc}} \right] \tag{4}$$

$$+ \frac{1}{2}\left(\frac{\left|\mathbf{B}v + x^b\right| - \left(\mathbf{B}v + x^b\right)}{2} \right)^2 .$$

This new form of cost function contains no inversion of **B**. Moreover, the VAF method uses a spatial filter to model the effect of the **B** matrix instead of calculating and storing the matrix directly. The new variational analysis scheme is simpler and more flexible in practcal implementations.

The choice of spatial filter coefficients should be based on *a priori* knowledge of the covariance matrix **B**. For instance, the following Gaussian filter function can be used to represent **B** for a homogeneous and isotropic background error field for a three-dimensional univariate problem (Daley 1991),

$$b_{ij} = \sigma_b^2 \exp\left[-\left(\frac{r_{ij}}{L_r}\right)^2 \right], \tag{5}$$

where b_{ij} are the elements of **B,** σ_b^2 is the variance of background error, r_{ij} is the spatial distance between grid points i and j, and L_r is the length scale or the background error de-correlation length and is, in practical use, sometimes tied to the observation station density. In all our experiments, we use a constant weight for the background term so effectively the background error is assumed to be homogeneous. This model represents isotropic background error covariances. In practice, a truncated filter is used, as in this paper, to save memory and computation, as suggested by Huang (2000). The truncation, however, destroys the positive definiteness of the modified **B**. The problem is alleviated to some extent by applying the Lanczos

window (see, e.g., Duchon 1979) instead of a sharp cutoff window, as is also done in this paper. Previous work (Hayden and Purser 1995) has demonstrated that an implicit recursive filter that guarantees positive definiteness can asymptotically approach a Gaussian filter. The explicit Gaussian filter is simpler and much easier to implement, however, especially for modeling anisotropic covariances, although it more expensive. As a first implementation, this 3DVAR analysis will use an explicit filter, following Huang (2000).

The use of isotropic background error covariances is based on the assumption that background errors at nearby points are similar (Riishojgaard 1998). But real background errors are usually flow-dependent and spatially anisotropic and the use of flow-dependent covariances in the analysis should improve the results, especially when data are sparse. Therefore, an anisotropic filter is considered for modeling the flow-dependent **B** matrix. The following expression can be used to model the anisotropic **B** matrix (Riishojgaard 1998),

$$b_{ij} = \sigma_b^2 \exp\left[-\left(\frac{r_{ij}}{L_r}\right)^2\right] \exp\left[-\left(\frac{f_i - f_j}{L_f}\right)^2\right] \tag{6}$$

where f is a field whose pattern represents that of the background error which we will call the error field. In this study, f is either the true error of the background or an estimate of it. L_f is the length scale in error field space, in contrast to the length scale L_r in physical space, and is determined by the de-correlation scale of background error in terms of the spatial gradient of f. The new background error covariance between any two points defined by Eq. (6) will follow the shape of the error field and fall off rapidly in the direction of strongest gradient, while the isotropic component of covariance will dominate in directions where the error changes slowly. Eq. (6) shows that, as L_f goes to infinity, the anisotropic covariance reduces to the isotropic form in Eq. (5).

The isotropic and anisotropic filters will be used separately to model the behavior of the background error covariances, and their results will be compared. These will be done in sections 4 and 5.

3. Hypothetical GPS network and the generation of *SWV* data

For the reasons stated in the introduction, simulated data are used to conduct retrieval experiments with our analysis system. The model used to produce the 'truth' field is the Advanced Regional Prediction System (ARPS, Xue et al. 2000; 2001; 2003) which is a nonhydrostatic atmospheric model formulated in a generalized terrain-following coordinate. High-resolution observations from hypothetical GPS networks are created from a forecast field for a dryline case that occurred on June 19, 2002 over the Southern Great Plains of the United States during the CAPS real-time forecast period (Xue et al. 2002) for IHOP_2002 (Weckwerth et al. 2004). The ARPS model is initialized using analysis of the ARPS Data Analysis System (ADAS, Brewster 1996) at 1200 UTC June 19, 2002, and is integrated for 8 hours. The computational domain has a horizontal grid spacing of 9 km and 40 layers in the vertical. The vertical grid is stretched from a minimum grid spacing of 100 meters near the surface.

Considering that in the near future, the mean spacing of ground receivers of GPS networks will probably not be much less than a hundred kilometers, the scale of water vapor distribution we can obtain will probably be no smaller than the mesoscale. The 9-km 8-hour forecast field is therefore thinned by sampling or picking specific humidity values every 4 grid points, yielding a resolution of 36 km and a horizontal grid size of 46×41. This gridded field is defined as the 'nature' or 'truth' and is sampled, using Eq. (7) given in the following, to generate the hypothetical GPS *SWV* observations. The specific humidity field from the 'nature', on the grid of 36 km resolution, is presented in Fig. 1. A roughly north-south zone of strong horizontal

moisture gradient is located to the west of Kansas, Oklahoma and Texas, corresponding to a dryline located in this region (Fig. 1a). The east-west vertical slice through $y = 234$ km (Fig. 1b) shows that a vertically oriented boundary between dry and moist air is found in the lowest 1.5 km at about $x = 360$ km and becomes nearly horizontal to the east. The upward bulging moisture tongue near $x = 576$ km reflects upward motion there. To the west of the dryline, the atmosphere is well-mixed up to 500 hPa. Such a strong gradient as well as the variations in the strong gradient of water vapor may not be properly captured by ordinary moisture observation networks, especially at levels away from the ground surface.

From the data set, the slant-path water vapor is simulated by the hypothetical GPS network using the formula

$$ SWV_{ij} = \int_{i^{th} \ receiver}^{j^{th} \ satellite} q_v \, ds, \tag{7} $$

where SWV_{ij} is the integrated water vapor along the slant path between the ith ground-based receiver and the jth GPS satellite, and q_v is the specific humidity along the path elements. The value of q_v is given by a tri-linear interpolation from the eight grid points surrounding the center of the path element. No error is added to the collected SWV observations except for one sensitivity experiment. The hypothetical GPS network is composed of nine irregularly distributed satellites simultaneously in view, and of 132 ground-based receivers evenly distributed in the analysis domain. The horizontal spacing of GPS receivers is 144 km. Both the simulation of observational data and the data analysis are performed on the 36-km grid in ARPS terrain-following coordinate. A schematic is given in Fig. 2 to illustrate the GPS observation network. Surface moisture observations are available at the GPS receiver sites.

4. Retrieval experiments and results

a. Single surface observation tests

In order to validate our newly developed system and more importantly to understand the behavior of the isotropic and anisotropic spatial filters, we first perform two idealized experiments in which only one single surface moisture observation is analyzed in the whole analysis domain. The isotropic and anisotropic filters, based on Eqs. (5) and (6), respectively, are used to model the background error covariances.

The single specific humidity observation with a value of 8.29 g kg^{-1} is located at the grid point marked by a black dot in Fig. 1, which also shows the true field at the surface used in this set of experiments. No *SWV* observation is involved so that the second term in the cost function in Eq. (4) vanishes. For simplicity, only a two-dimensional horizontal filter is used so that the analyses at different levels are decoupled and the overall analysis is essentially two dimensional. In the case of an isotropic filter, a horizontal length scale (L_r) of 4 grid intervals is used. In the anisotropic filter case, L_r has a length of 6 grid intervals and the length scale in error field space, L_f, is 2.0 g kg^{-1}. The larger value of L_r is used in the anisotropic filter case so that the combined effective de-correlation length scale is of sufficient length in the direction of error-field contours.

For the single observation tests, the background value is assumed constant on each model level and is equal to 12.71 g kg^{-1} at the surface. In this case, the true error field, equaling to the true field minus the background, has the same pattern as the true field itself at each level. The true error field specifies *f* in Eq. (6) for this experiment. Here we assume that the observation is much more accurate than the background, and the relative weights, proportional to the inverse of error variances, of 1, 500 and 50 are given to the background, observation and the non-negative-constraint terms of the cost function, respectively. The resultant analysis should therefore be

much closer to the observation than to the background. The analysis increment fields are shown in Fig. 3.

The observation increment at the point of observation is spread in space through the background error covariances. Consequently, the analysis with isotropic covariances gives an analysis increment of circular shape while that with anisotropic covariances shows an increment that is related to the error field (Fig. 3). Since the error field in this case has the same pattern as the true specific humidity field shown in Fig.1a, the analysis increment in the anisotropic case (Fig. 3a) is oriented in the north-northeast to south-southwest direction and is narrower in the east-west and broader in the north-south directions compared to the circular increment of isotropic analysis. The spatial scales of the increment fields roughly match the effective de-correlation scales used in the filters. The analysis increments at the observation location are about - 4.4 g kg^{-1} in both cases, giving a total analysis of 8.30 g kg^{-1} that is, as expected, very close to the observed value.

These experiments confirm that our 3DVAR analysis system, using isotropic and anisotropic filters, performs as expected. In the following and in section 5, we apply this system to the analysis of 3D moisture filed using GPS slant water vapor data as well as surface observations.

b. SWV retrieval experiments

A list of retrieval experiments analyzing the simulated GPS *SWV* and surface moisture observations is given in Table 1. The overall correlation coefficients between analysis increment for these experiments and the 'truth' increment (truth minus background) are also given in the table.

First, a control experiment (CNTL) is performed. In this experiment, both *SWV* and

regular surface observations at receiver sites are used. The analysis background is created by smoothing the 'nature' field 50 times, using a 2-D 9-point filter (with 1-2-1 weightings in each direction) in the horizontal. It can be seen from Fig. 4 that this background shows a general pattern of higher moisture to the east and lower values to the west but detailed dryline structure is lost. Since both the truth and background are known, the background error can be calculated. It is therefore possible to model the background error covariances by taking the known background error field as f in Eq. (6) and this is done for CNTL.

The length scale L_r used is equal to 4 grid intervals in both horizontal and vertical directions. L_f is given as 2 g kg^{-1}. Owing to the insignificant effect of filter on the far distance, cutoff radii are used and chosen to be 10 grid intervals in the horizontal and 6 layers in the vertical, respectively. The selection of filter scale depends, for one thing, on the density of ground-based GPS receivers. The relatively small filter scales and cutoff radii are chosen here so that gaps between receiver stations are filled without excessive smoothing to the analysis.

In all except for one sensitivity experiment, no error is added to the simulated observations, so the observations are in a sense infinitely accurate compared to the background. Because it is the relative errors that matter, we choose to specify the inverse of error variances in terms of weights for each term of the cost function, and the relative weights of the background, GPS *SWV* observation and regular surface observation terms and of the non-negative constraint are specified as 1, 100, 500 and 50, respectively. The much higher weights given to the observation terms reflect the high accuracy of observations as compared to the background, and we are interested in finding out how well the 3DVAR scheme can do in recovering the 3D moisture structure under an ideal condition. The cost function defined by Eq. (4) is minimized with respect to the increment of specific humidity, using a conjugate gradient algorithm.

With the above parameter settings, the control experiment is conducted. Figure 5 presents a vertical cross-section of retrieved moisture at $y = 234$ km from CNTL, as compared to the 'truth'. Only the domain below 6 km altitude is shown here since water vapor has very low values above 6 km. It is obvious that this retrieved moisture field matches 'nature' very well. The dryline near $x = 360$ km is accurately captured. There is a strong east-west moisture gradient at the low levels and the moisture isohumes are almost perpendicular to the ground near the dryline. Meanwhile, due to presumably upward motion near the dryline at near $x = 576$ km, there exists a moisture tongue at this location which is surrounded by two troughs to its east and west due to return flows. Figure 6 shows the q_v increments at the surface. The retrieved increment matches almost exactly the increment of 'truth' (the difference between truth and background); their shapes match very well and extrema locations coincide. The correlation coefficient between the two increment fields is 0.926 on the entire grid (Table 1).

To see the performance when an isotropic filter is used instead, another experiment, named SNF is conducted, in which the background is based on the Smoothed 'truth' with No Flow-dependent background error (Table 1). This experiment is otherwise the same as CNTL. The length scale for an isotropic filter should be smaller, so L_r is given a length of 3 grid intervals. Figure 7 shows the vertical cross-section at $y = 234$ km and the analysis increment at the surface from SNF. The retrieved moisture field also exhibits a dryline around $x=360$ km, a moisture tongue due to updraft together with troughs on its sides due to downdraft (Fig. 7a). The strength of the updraft and downdraft as reflected by the isohume shapes is weaker than that in CNTL and 'nature'. Their locations near the ground are shifted eastwards relative to the 'truth'. The isohumes have shapes different from the 'truth' and are smoother than 'truth' (Fig. 7b). Overall, this analysis does not match the 'truth' as well as the analysis of CNTL. The overall

correlation coefficient of the increment fields is reduced to 0.83 from the 0.926 of CNTL (Table 1).

Experiment CNTL has a flow-dependent background error covariances based on known background error while experiment SNF assumes isotropic covariances. Their comparison illustrates the importance of background error covariances. The problem is, however, that the background error covariances are never known exactly. In order to improve actual analysis in NWP, it is necessary to seek feasible methods for representing the background error covariances as accurately as possible. Using an isotropic filter, we can obtain an analysis (the result of SNF) that is much closer to the 'truth' than the initial background field. As a result, the background error may be estimated by subtracting the background from the isotropic analysis, which we call the updated (from that based on initial background) background error. Based on this consideration, experiment SUF (with the Smoothed 'truth' as the background and the Updated Flow-dependent background error covariances) is performed (see Table 1), which does a second analysis starting from the same background but using an anisotropic filter based on the error field calculated as the difference between the output of SNF and the background. This analysis matches the 'truth' much better than that of SNF as shown in Fig. 8, and the improvement is also, though to a lesser extent, reflected in the overall correlation coefficient (0.832 versus 0.830, Table 1). In the vertical cross-section (Fig. 8a), the isohumes for specific humidity values of 4, 6 and 8 g kg^{-1} follow the 'truth' much better than those in Fig. 7a. The fine-scale moisture bulge near x = 360 km is also well recovered. Meanwhile, the maximum value of 18.65 g kg^{-1} from SUF is closer to the true value of 18.64 g kg^{-1} than the 18.79 g kg^{-1} of SNF. The surface increment field (Fig. 8b) contains a lot more finer-scale structures that are consistent with the pattern of 'truth' increment field.

These three experiments clearly demonstrate that the 3DVAR system with flow-dependent background error covariances realized through an anisotropic spatial filer provides better analysis than that with isotropic covariances. This is true even when the background error is estimated using a first-pass analysis that utilizes isotropic error covariances.

c. Retrievals with a vertically logarithmic background

The background in the above three experiments was constructed by smoothing the 'truth'. This background, shown in Fig. 4, still contains structures of moisture at the larger scales and some information on the vertical moisture distribution. To understand how much the analysis depends on the structure information present in the background, we conduct another two experiments in which the background is specified using a logarithmic vertical profile. This profile decreases from a value of 12 g kg^{-1} at the surface to zero at 17 km, the top of the analysis domain. This profile is used to specify the background q_v values on each terrain-following grid level, so that the background is uniform along the model levels. Such a background is artificial and supplies no realistic information on the structure of moisture, therefore a successful analysis has to extract all structure information from the observations with the help of background error covariances. For these tests, the weight of background term is reduced from 1.0 to 0.2 because of the reduced accuracy of the background.

Two experiments are LNF (Logarithmic background with No Flow-dependent **B**) using the isotropic filter, and experiment LTF (Logarithmic background and Truth-based Flow-dependent **B**) using the anisotropic filter that is based on the true error field (Table 1). The vertical cross-sections of retrieved moisture field for these two experiments are presented in Fig. 9 and Fig. 10, respectively. For the isotropic filter case (Fig. 9), the dryline is weaker and the boundary separating the moist and dry air shows a significant slope at the low levels. The

structure is, incorrectly, more symmetric in the east-west direction in accord with the isotropy and the moist bulge near $x = 288$ km is completely missed in the analysis (Fig. 9). However, the vertical structure of the analysis using flow-dependent background error is much better than that using isotropic error as expected. For the anisotropic case, the isohumes generally follow the 'true' isohumes except near the boundaries. The dryline is reflected by the almost vertically oriented boundary between the dry and moist air in the lowest 1.5 km. But there are more errors near the east and west boundaries in Fig. 10 (for LTF) than in Fig. 5 (for CNTL). This can be explained by the fact that, with the logarithmic background, the recovery of 3D moisture structure depends more on the GPS slant-path water vapor observations but there are fewer slant paths near the boundaries because there is no path coming in from outside the boundary (this problem will be gone for global analyses).

As we have explained previously, the inclusion of the background term in the 3DVAR analysis eliminates the under-determinedness problem. We have found that even with a background that is worse than the logarithmic one tested above, such as the case of a constant value applied to 3D, a convergent 3DVAR analysis is still feasible although the analysis is poorer (results not shown). On the contrary, the minimization fails (an unphysical analysis was produced before any convergence could be reached) when the background term is excluded and when no additional smoothing constraint is applied.

The above experiments show that our 3DVAR system is capable of recovering the 3D moisture structure from ground-based GPS slant-path water vapor and surface moisture observations even when an artificial analysis background is used. When flow-dependent background error information is known and properly used, the analysis is improved.

5. Sensitivity Experiments

In this section, sensitivity experiments are performed to test several factors that can affect the quality of moisture analysis.

a. Impact of surface moisture observations

Sensitivity experiment STFNSFC (STF, Smoothed background and Truth-based Flow-dependent error covariances, as in CNTL, plus No SurFaCe observations) is conducted which is same as CNTL except that surface observations are excluded. This is to test the effect of surface moisture observations on the retrieval. The overall correlation coefficient between the increment fields of this retrieval and that of 'truth' is now 0.894, a quite significant reduction from the 0.926 of CNTL (Table 1).

The surface increment field from this experiment is presented in the Fig. 11a. The pattern of the increment field is good, but the extrema are only half as large as those from CNTL (Fig.6). Due to the integral nature of the *SWV* observations, the deterioration in accuracy of surface analysis will also worsen the analysis at the upper levels. This is shown by Fig. 12, which shows the correlation coefficients between the retrieval and 'truth' increment fields plotted against the vertical model layers[*]. It is seen that, below the 5th model level (about 500 meters AGL), the correlation coefficients from experiment STFNSFC are always less than those from CNTL. At the upper levels, the correlation coefficients of STFNSFC are mostly smaller than those of CNTL. Clearly, the surface moisture observations improve the overall analysis by directly adjusting near-surface q_v field and by more accurately distributing water vapor in the vertical.

Another experiment, SNFNSFC (SNF plus No SurFaCe observations), is the same as

---

[*] The mean heights of the center of the model levels are 0.80, 0.89, 0.98, 1.12, 1.29, 1.50, 1.73, 2.00, 2.28, 2.58, 2.90, 3.24, 3.58, 3.94, 4.30, 4.66, 5.04, 5.41, 5.79, 6.16, 6.54, 6.92, 7.30, 7.67, 8.05, 8.43, 8.81, 9.20, 9.59, 9.98, 10.38, 10.79, 11.21, 11.65, 12.10, 12.57, 13.07, 13.59, 14.14, 14.72, 15.34, 16.00, and 16.66 km.

SNF (Table 1) except that surface observations are excluded. The results show that the analyzed dryline is much weaker, compared to the corresponding retrieval with surface observations, i.e., that of SNF. The increment field at the surface (Fig. 11b) does not match well the 'true' increment field (Fig. 6a) and the extrema are even weaker than in STFNSFC (Fig. 11a).

The overall correlation coefficient is 0.67 for SNFNSFC, about 0.26 and 0.22 less than those of CNTL and STFNSFC, respectively. The large drop in accuracy is mostly due to the differences at the lowest levels where the surface observations have the greatest impact. This is so partly because, near the surface, the model meshes are intersected by very few or no slant paths. Limited by the relatively small analysis domain, the lowest elevation angle of usable slant paths in our experiments is about 15 degrees. The inaccuracy in the surface moisture analysis influences the analysis at upper levels because of the integral nature of GPS observations.

The above comparisons of analyses tell us that the surface observations play an important role, especially when realistic flow-dependent background error covariances are not available. Flow-dependent background error covariances, consequently, have more impact on the analysis in the absence of surface observations. The best retrieval is obtained when both accurate background error covariance information and surface observations are included while the worst is obtained when neither is.

b. Impact of vertical filtering

To isolate the effect of vertical filtering, the horizontal filter only is used in experiment STFNVF (STF or CNTL plus No Vertical Filtering, Table 1). All other parameter settings are the same as in CNTL. Figure 13 shows the vertical profiles of correlation coefficients of analysis increment from CNTL and STFNVF with the 'truth' increment. It is clear that CNTL gives a better analysis than STFNVF. There is almost no difference right at the surface owing to the

dominant contributions of surface observations there but the correlation coefficient of STFNVF is significantly lower between the 2nd and 5th levels, with the difference being larger than 0.3 at the 2nd level. Still, the correlation coefficients from the 3rd through 13th level, are larger than or equal to 0.8, so the pattern of analysis increment remains reasonably good at those levels. The gradient of the low-level analysis increment is, however, clearly weaker than that of the 'truth' increment (not shown). Figure 13 also shows a general improvement in the analysis at the upper levels when vertical filtering is included. Therefore, the vertical filtering is very important for accurate analysis in the boundary layer and beneficial at the upper levels too. It is so because, in the absence of vertical filtering, surface observation information cannot be spread upward into the boundary-layer where information from GPS data is also lacking.

c. Sensitivity to observation error

One of the advantages of using simulated data is that observation data can be error free, but the sensitivity to observation errors should be examined for practical use. This is done in experiment STF_ER ('ER' for error), in which normally distributed errors with 5% and 7% standard deviations are added to the simulated surface and *SWV* observations, respectively. The experiment is otherwise the same as CNTL. The errors are consistent with the estimate of Braun et al. (2001) for real data. Compared to CNTL, the relative weights for *SWV* and regular surface observation terms are decreased to 80 and 400, respectively, because of the added observational errors.

The analysis of STF_ER also matches the 'truth' reasonably well, as shown in Fig. 14. In the vertical cross section, only the 2 g kg^{-1} isohume is prominently different from the CNTL result. The analyzed maximum is 19.49 g kg^{-1}, 0.75 g kg^{-1} larger than the 'truth' maximum of 18.64 g kg^{-1}. The horizontal structure (not shown) also matches 'truth' well below 7 km where

95% of water vapor concentrates. Therefore, even in the presence of realistic errors in the *SWV* and surface observations, our 3DVAR system is still able to produce good analysis, although the overall correlation coefficient of the increments decreases from 0.926 to 0.79 (Table 1). Still, all major structures of the dryline are recovered well.

d. Observation density test

Finally, the sensitivity of 3D moisture analysis to the density of ground-based GPS receivers is examined. The receiver density is halved in experiments SNF_LR and STF_LR (LR for Low Resolution), with one receiver station every 8 grid intervals and a station spacing of 288 km. The horizontal de-correlation length scale in physical space, L_r, and the corresponding cutoff radius are enlarged since their choices should be partly related to receiver network density. The length scale should be large enough to fill the gaps between receiver stations. Experiment SNF_LR uses an isotropic filter with L_r given a length of 5 grid intervals, while experiment STF_LR uses an anisotropic filter and a 6-grid interval physical length scale. L_f is still 2.0 g kg^{-1}. SNF_LR and STF_LR should be compared with the high-resolution counterparts SNF and CNTL, respectively.

Figure 15 presents the retrieval result from STF_LR. Comparing the east-west cross-section at $y = 234$ km (Fig. 15a) with that of CNTL (Fig. 5), we can see that the difference in the quality of analysis is relatively small, indicating that the 3D moisture retrieval is not very sensitive to the observation density, in this case when reliable statistics of the background error are available and used. This conclusion is also supported by the surface analysis increment field in Fig. 15b. The overall correlation coefficient is about 0.87 for STF_LR, 0.06 less than that of CNTL (Table 1). When the background error covariances are given an isotropic form in SNF_LR, the overall correlation coefficient decreases to 0.68 from 0.83, that of the

corresponding high density case (SNF). Clearly, this reduction is much larger than the flow-dependent error cases (CNTL and STF_LR). This implies that the retrieval quality is more sensitive to receiver station density when no good information on the background error structure is available or used.

6. Summary and conclusions

A new 3DVAR analysis system is developed for retrieving the 3D water vapor structure of the atmosphere from a GPS observation network. This network provides integrated water vapor along slant paths between GPS satellites and ground-based receivers, as well as direct moisture measurements at the ground receiver sites. The ARPS mesoscale model is used to create a 'true' atmospheric moisture field for a dryline case that occurred during the IHOP_2002 field experiment, and this 'true' atmosphere is sampled by the GPS observation network to produce simulated slant-path water vapor data and surface moisture observations.

Our analysis system is formulated in the same generalized terrain-following coordinate system used by ARPS. It includes a background term in the 3DVAR cost function, which for one thing overcomes the under-determinedness problem with GPS data retrieval. Three dimensional Gaussian-type explicit spatial filters are used to model background error covariances which can be isotropic or flow-dependent and spatially anisotropic. In the latter case, a flow-dependent anisotropic filter is constructed based on the true or estimated error field of the analysis background, following the approach of Riishojgaard (1998). Three-dimensional variational retrieval experiments are conducted using the simulated data and the quality of the analyses is evaluated through comparisons with the truth. The results are summarized as follows:

1) The 3DVAR system, even with isotropic background error covariances, can retrieve from surface moisture and GPS *SWV* observations the 3D mesoscale moisture structure reasonably

well, and the analysis is able to capture major features of the model simulated dryline.

2) The use of flow-dependent background error covariances realized through an anisotropic spatial filter always improves the analysis. The best analysis is obtained in the ideal situation where the background error structure is known. Otherwise, the background error structure can be estimated from a first-pass analysis obtained using isotropic background error covariances, and in this case, a significant improvement in the analysis can also be obtained. The role of flow-dependent covariances is enhanced at regions near the lateral boundaries and at the low levels, where the model grid cells are intersected by few if any slant paths. In these data sparse regions, the retrieval benefits more from the properly spread observation increments via background error covariances.

3) The retrieval is still feasible even with an artificial vertically logarithmic background that is homogeneous along the model levels when flow-dependent background error covariance is applied. The quality of analysis is not as good as the smoothed background case but still reasonable except for regions near the boundaries where few slant paths go through. This suggests that our 3DVAR method is rather robust, and the analysis derives most of the water vapor structure information from the observations and the background error statistics.

4) Sensitivity experiments indicate that surface moisture observations are important for accurate analysis of water vapor at low levels, and more so when no good information on the background error covariances is available or used. When there is no surface observation, flow-dependent background error has an even more positive impact on the analysis. The vertical component of the spatial filter is shown to be very beneficial, especially in the low-level data-sparse region, where its main effect is in the upward spread of surface moisture information. Improved low-level moisture analysis also leads to better upper-level analysis through an

improved vertical distribution of moisture.

5) Sensitivity tests on the observation errors show that our analysis system is also robust in the presence of realistic errors in the surface moisture and *SWV* observations. Main structures of the dryline can still be recovered with reasonable accuracy. With the density of ground-based receiver stations halved, the dryline structure can be reasonably recovered when flow-dependent background error covariances are used but the problems near the boundaries are worse. Such deterioration is more dramatic when an isotropic spatial filter is used. In other words, the positive impact of flow-dependent background error covariances increases when the density of ground-based GPS receiver stations decreases.

In our current analysis system, an explicit spatial filter is used to model the background error covariances as well as to reduce computer memory requirements as compared to storing the full **B** matrix. This treatment cannot guarantee the positive definiteness of the modified covariance matrix, however. Meanwhile, the larger cutoff radii are, the more expensive the algorithm becomes. A computationally more efficient alternative is the recursive filter, which can be used to model both isotropic and anisotropic background error (Wu et al. 2002; Purser et al. 2003a; 2003b), although the realization of the latter with recursive filter is much more complicated. We plan to implement and test recursive filters in our system in the near future. Further, we will use the retrieved moisture field to initialize a mesoscale model, such as the ARPS, and examine the impact of assimilating GPS *SWV* data on short-range precipitation forecasts. The assimilation and examination of the impact of real GPS *SWV* data collected during the IHOP_2002 field experiment are also planned. Further extensions of this work include analyzing tropospheric total delay data along slant paths in a multivariate 3DVAR system, where the contributions of temperature and pressure to the signal delay are also included and the

moisture field is analyzed together with all other state variables.

Acknowledgement

References

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75,** 111-130.

Bengtsson, L., G. Robinson, R. Anthes, K. Aonashi, A. Dodson, G. Elgered, G. Gendt, R. Gurney, M. Jietai, C. Mitchell, M. Mlaki, A. Rhodin, P. Silvestrin, R. Ware, R. Watson, and W. Wergen, 2003: The Use of GPS Measurements for Water Vapor Determination. *Bull. Amer. Meteor. Soc.*, **84,** 1249-1258.

Braun, J., C. Rocken, and R. Ware, 2001: Validation of line-of-sight water vapor measurements with GPS. *Radio Sci.*, **36,** 459–472.

Brewster, K., 1996: Application of a Bratseth analysis scheme including Doppler radar data. *Preprints, 15th Conf. Wea. Anal. Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 92-95.

Businger, S., S. R. Chiswell, M. Bevis, J. Duan, R. A. Anthes, C. Rocken, R. H. Ware, M. Exner, T. VanHove, and F. S. Solheim, 1996: The Promise of GPS in Atmospheric Monitoring. *Bull. Amer. Meteor. Soc.*, **77,** 5-18.

Cucurull, L., F. Vandenberghe, D. Barker, E. Vilaclara, and A. Rius, 2004: Three-dimensional variational data assimilation of ground-based GPS ZTD and meteorological observations during the 14 December 2001 storm event over the western Mediterranean Sea. *Mon. Wea. Rev.*, **132,** 749-763.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.

Davis, J. L., T. A. Herring, I. I. Shapiro, A. E. Rogers, and G. Elgered, 1985: Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio Sci.*, **20,** 1593-1607.

De Pondeca, M. S. F. V. and X. Zou, 2001a: Moisture retrievals from simulated zenith delay

"observations" and their impact on short-range precipitation forecasts. *Tellus*, **53,** 192.

De Pondeca, M. S. F. V. and X. Zou, 2001b: A case study of the variational assimilation of GPS zenith delay observations into a mesoscale model. *J. Appl. Meteor.*, **40,** 1559-1576.

Duchon, C. E., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.*, **18,** 1016-1022.

Emanuel, K., D. Raymond, A. Betts, L. Bosart, C. Bretherton, K. Droegemeier, B. Farrell, J. M. Fritsch, R. Houze, M. LeMone, D. Lilly, R. Rotunno, M. Shapiro, R. Smith, and T. A, 1995: Report of the first Prospectus Development Team of the U.S. Weather Research Program to NOAA and the NSF. *Bull. Amer. Meteor. Soc.*, **76,** 1194-1208.

Falvey, M. and J. Beavan, 2002: The impact of GPS precipitable water assimilation on mesoscale model retrievals of orographic rainfall during SALPEX'96. *Mon. Wea. Rev.*, **130,** 2874-2888.

Guo, Y.-R., Y.-H. Kuo, J. Dudhia, D. Parsons, and C. Rocken, 2000: Four-dimensional variational data assimilation of heterogeneous mesoscale observations for a strong convective case. *Mon. Wea. Rev.*, **128,** 619-643.

Ha, S.-Y., Y.-H. Kuo, Y.-R. Guo, and G.-H. Lim, 2003: Variational assimilation of slant-path wet delay measurements from a hypothetical ground-based GPS network. Part I: Comparison with precipitable water assimilation. *Mon. Wea. Rev.*, **131,** 2635-2655.

Harris, B. A. and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Roy. Meteor. Soc.*, **127,** 1453-1468.

Hayden, C. M. and J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34,** 3-15.

Huang, X.-Y., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.*, **128,** 2588–2600.

Kuo, Y.-H., Y.-R. Guo, and E. R. Westerwater, 1993: Assimilation of precipitable water measurements into a mesoscale numerical model. *Mon. Wea. Rev.*, **121,** 1215-1238.

Kuo, Y.-H., X. Zou, and Y. R. Guo, 1996: Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model. Part I: Moisture retrieval and sensitivity experiments. *Mon. Wea. Rev.*, **124,** 122-147.

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of the future generation of integrated observing systems. *Preprint volume, 1st Symposium on Integrated Observation Systems*, Long Beach, CA, Amer. Meteor. Soc., 45-47.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109,** 701-721.

MacDonald, A. E., Y. Xie, and R. H. Ware, 2002: Diagnosis of three-dimensional water vapor using a GPS network. *Mon. Wea. Rev.*, **130,** 386-397.

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 covariances. *Mon. Wea. Rev.*, **131,** 1524-1535.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131,** 1536-1548.

Riishojgaard, L. P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, **50A,** 42-57.

Riishøjgaard, L. P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, **50A,** 42-57.

Rocken, C., R. H. Ware, T. v. Hove, F. Solheim, C. Alber, J. Johnson, M. Bevis, and S. Businger,

1993: Sensing atmospheric water vapor with global positioning system. *Geophys. Res. Lett.*, **20,** 2631-2634.

Ware, R., C. Alber, C. Rocken, and F. Solheim, 1997: Sensing integrated water vapor along GPS ray paths. *Geophys. Res. Lett.*, **24,** 417-420.

Ware, R. H., D. W. Fulker, S. A. Stein, D. N. Adnderson, S. K. Avery, R. D. Clark, K. K. Droegemeier, J. P. Kuettner, J. B. Minster, and S. Sorooshian, 2000: SuomiNet: A real-time national GPS network for atmospheric research and education. *Bull. Amer. Meteor. Soc.*, **81,** 677-694.

Weckwerth, T. M., D. B. Parsons, S. E. Koch, J. A. Moore, M. A. LeMone, B. B. Demoz, C. Flamant, B. Geerts, J. Wang, and W. F. Feltz, 2004: An overview of the International H2O Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.*, **85,** 253-277.

Wolfe, D. E. and S. I. Gutman, 2000: Developing an Operational, Surface-Based, GPS, Water Vapor Observing System for NOAA: Network Design and Results. *J. Atmos. Oceanic Tech.*, **17,** 426-440.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130,** 2905-2916.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics*, **75,** 161-193.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and

applications. *Meteor. Atmos. Phy.*, **76,** 143-165.

Xue, M., K. Brewster, D. Weber, K. W. Thomas, F. Kong, and E. Kemp, 2002: Realtime storm-scale forecast support for IHOP 2002 at CAPS. *Preprint, 15th Conf. Num. Wea. Pred. and 19th Conf. Wea. Anal. Forecasting*, San Antonio, TX, Amer. Meteor. Soc., 124-126.

Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Physics*, **82,** 139-170.

List of Tables

Table 1. List of moisture retrieval experiments. In the table, *SWV* denotes the GPS slant-path water vapor observation data and 'sfc' is for the surface moisture observation data.

Table 1. List of moisture retrieval experiments. In the table, *SWV* denotes the GPS slant-path

water vapor observation data and 'sfc' is for the surface moisture observation data.

Experiment	background	Flow-dependent background error?	Obs used	Obs error	Obs resolution	Filter	Correlation coefficient
CNTL or STF	smoothed truth	Yes, based on true background error	*SWV*+sfc	no	1 obs / 4 grid intervals	3D	0.926
SNF	smoothed truth	No	*SWV*+sfc	no	same as above	3D	0.830
SUF	smoothed truth	Yes, on updated analysis	*SWV*+sfc	no	same as above	3D	0.832
LTF	logarithmic	Yes, on true background error	*SWV*+sfc	no	same as above	3D	0.827
LNF	logarithmic	No	*SWV*+sfc	no	same as above	3D	0.821
STFNSFC	smoothed truth	Yes, on true background error	*SWV*	no	same as above	3D	0.894
SNFNSFC	smoothed truth	No	*SWV*	no	same as above	3D	0.668
STFNVF	smoothed truth	Yes, on true background error	*SWV*+sfc	no	same as above	2D	0.801
STF_ER	smoothed truth	Yes, on true background error	*SWV*+sfc	yes	same as above	3D	0.790
SNF_LR	smoothed truth	No	*SWV*+sfc	no	1 obs / 8 grid intervals	3D	0.679
STF_LR	smoothed truth	Yes, on true background error	*SWV*+sfc	no	1 obs / 8 grid intervals	3D	0.870

List of Figures

CNTL. Dashed lines represent negative values and solid lines positive values.

Figure 7. (a) East-west vertical cross-section of specific humidity field (g kg^{-1}) at y = 234 km where solid lines are for 'nature' and dotted lines for experiment SNF. (b) Analysis increment of specific humidity (g kg^{-1}) at the surface from experiment SNF, where dashed lines are for negative values and solid line for positive values.

Figure 8. As Fig. 7 but for experiment SUF.

Figure 9. As Fig. 5 but dotted lines are from experiment LNF.

Figure 10. As Fig. 5 but dotted lines are from experiment LTF.

Figure 11. As Fig. 6 but (a) is for experiment STFNSFC (b) is for experiment SNFNSFC.

Figure 12. Profiles of correlation coefficient of specific humidity increment (difference from background, in g kg^{-1}) between those of 'nature' and 3DVAR analysis from experiments CNTL, STFNSFC, SNFNSFC, and SNF, plotted for different model levels. Mean height of each level is given in a footnote in the paper.

Figure 13. As Fig. 12 but for experiments CNTL (solid line) and STFNVF (dotted line).

Figure 14. As Fig. 5 but dotted lines are for experiment STF_ER.

Figure 15. As Fig. 7 but for experiment STF_LR.

MIN=3.191 MAX=18.70 inc=1.0   MIN=0.1438 MAX=18.64 inc=2.0
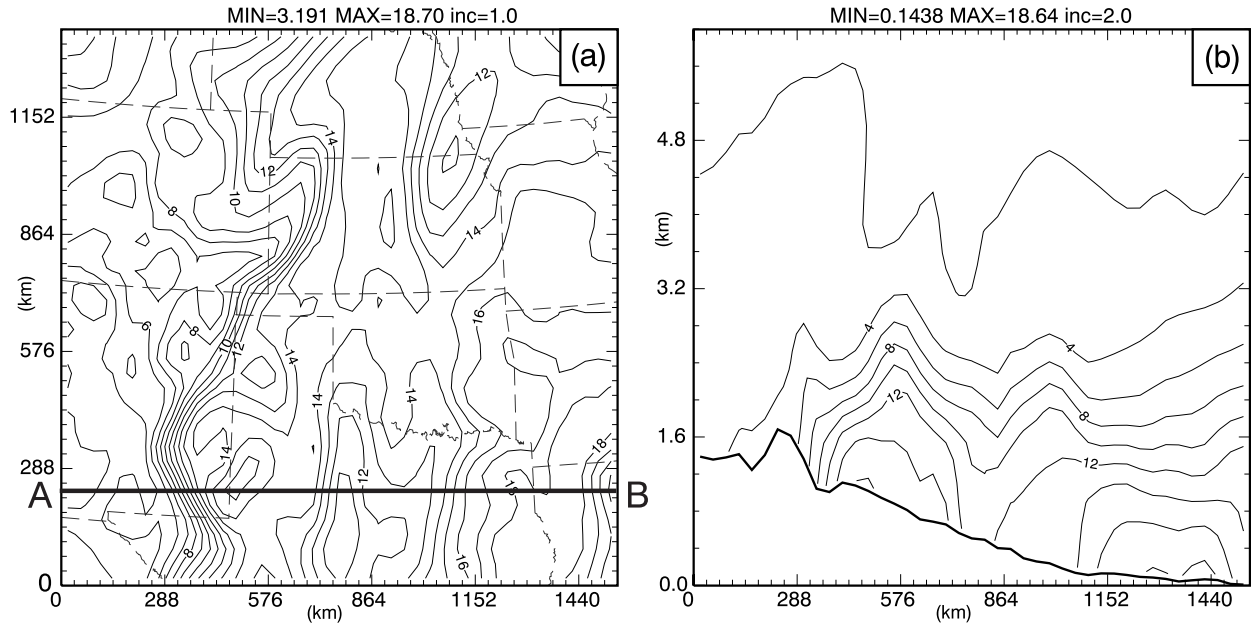
Figure 1. Specific humidity field (g kg^{-1}) from the 'nature' derived from an ARPS simulation for the IHOP case of 19 June 2002, at 20 UTC, (a) at the surface and (b) in the east-west vertical cross-section at y = 234 km (along thick line A-B). A roughly north-south zone of strong horizontal moisture gradient is located to the west of Kansas, Oklahoma and Texas, representing the dryline. In vertical cross-section, a boundary between the dry and moist air is oriented nearly vertically in the lowest 1.5 km then turns horizontal to the east.
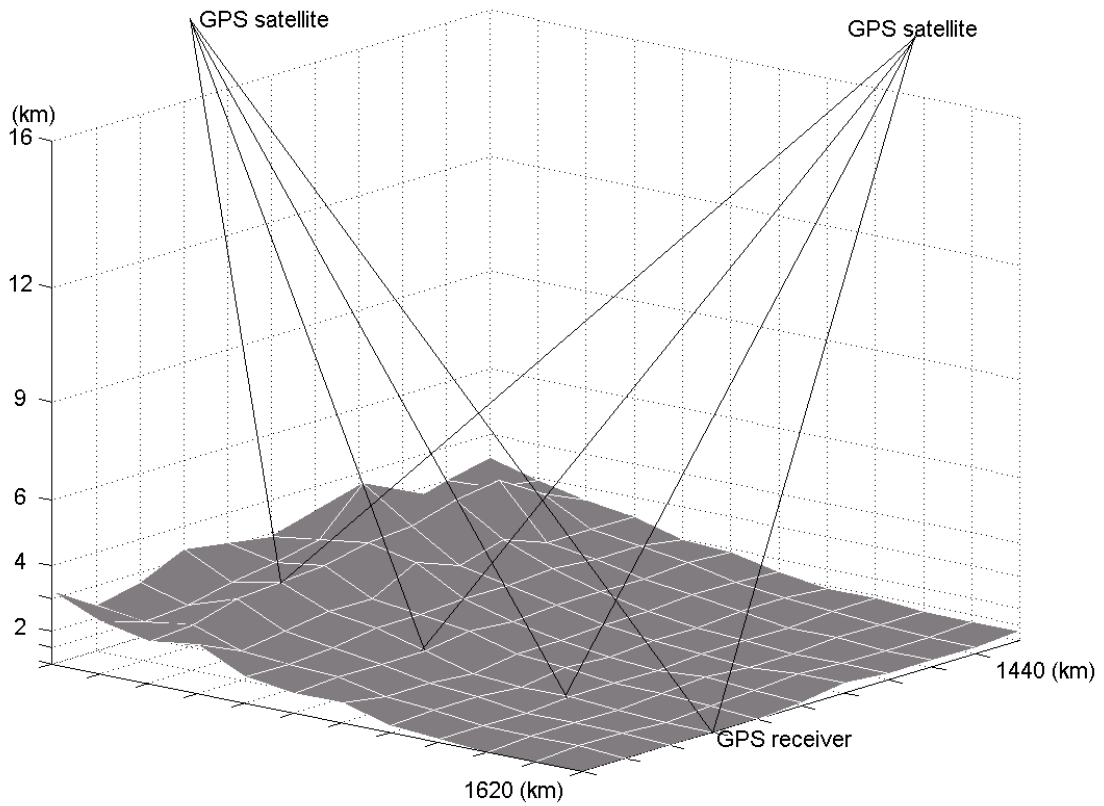
Figure 2. A schematic of a ground-based GPS observation network whose data are analyzed using 3DVAR. Shaded surface represents terrain. Dark solid lines are slant paths between ground-based GPS receivers and GPS satellites. Dotted lines give a sense of the vertically stretched grid although the actual grid levels are in terrain-following coordinate.

MIN=-4.412 MAX=0.1396E-01 inc=0.5   MIN=-4.410 MAX=0.1984E-02 inc=0.5
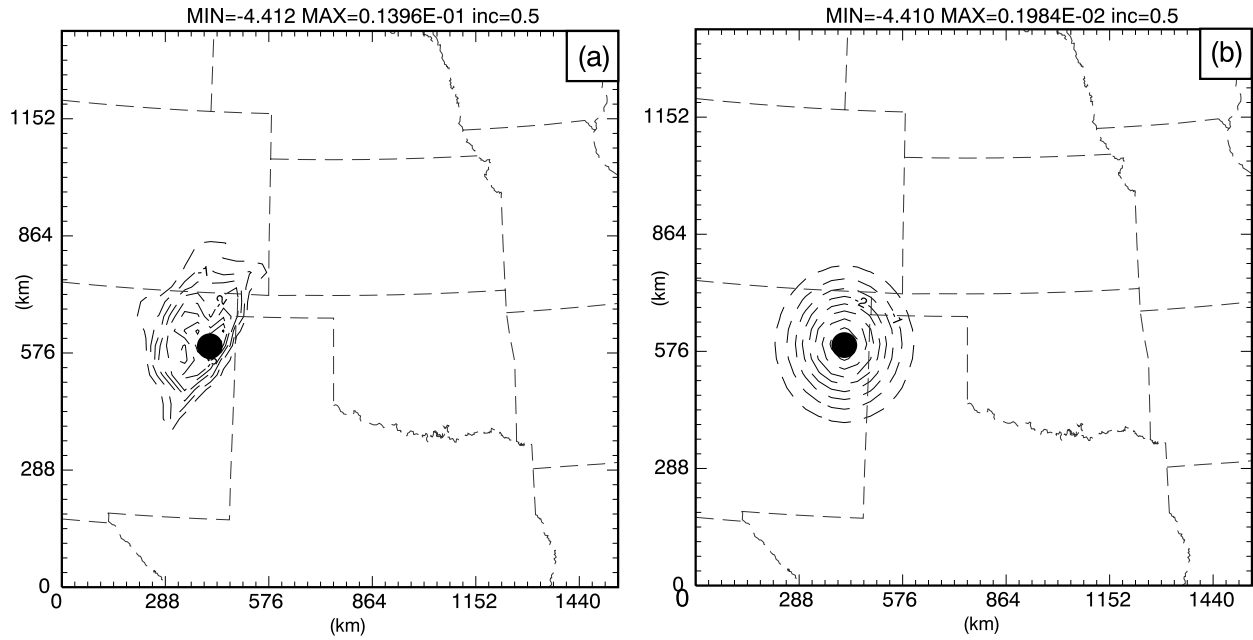
Figure 3. Specific humidity increment field in g kg^{-1} at the surface from single moisture observation tests, for 3DVAR analysis (a) with anisotropic flow-dependent background error covariance and (b) with isotropic covariance. The location of the single specific humidity observation at the surface is marked by the black dot. Contour interval is 0.5 g kg^{-1}.
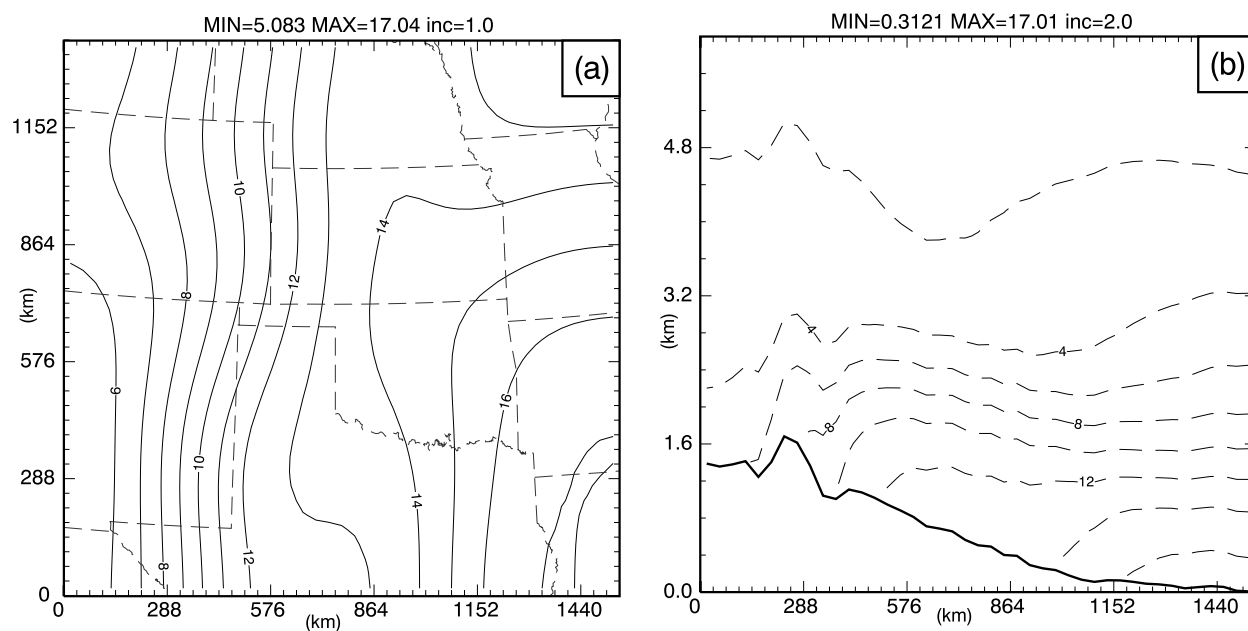
Figure 4. Background specific humidity field in g kg^{-1}, obtained by smoothing 'nature' 50 times using a 9-point filter in the horizontal, (a) at the surface and (b) in the east-west vertical cross-section at y = 234 km.
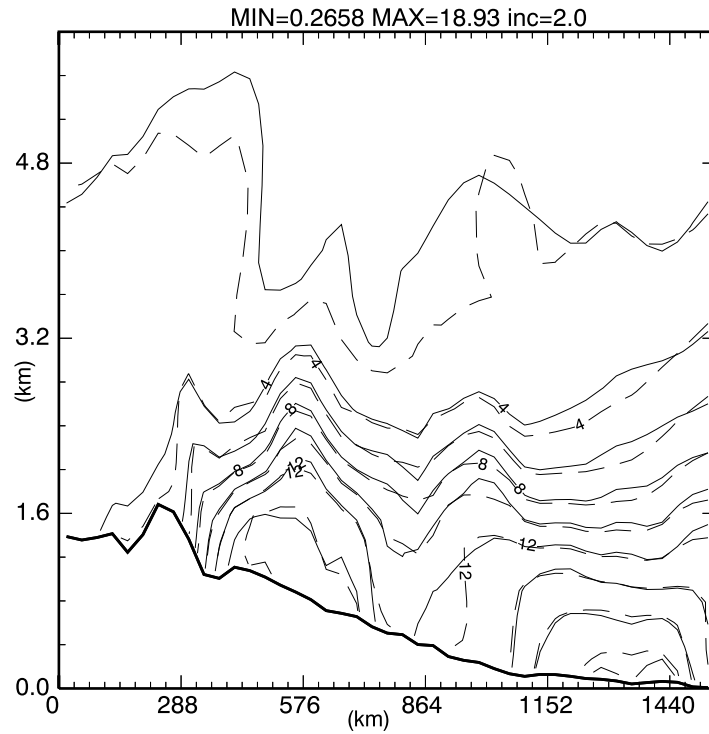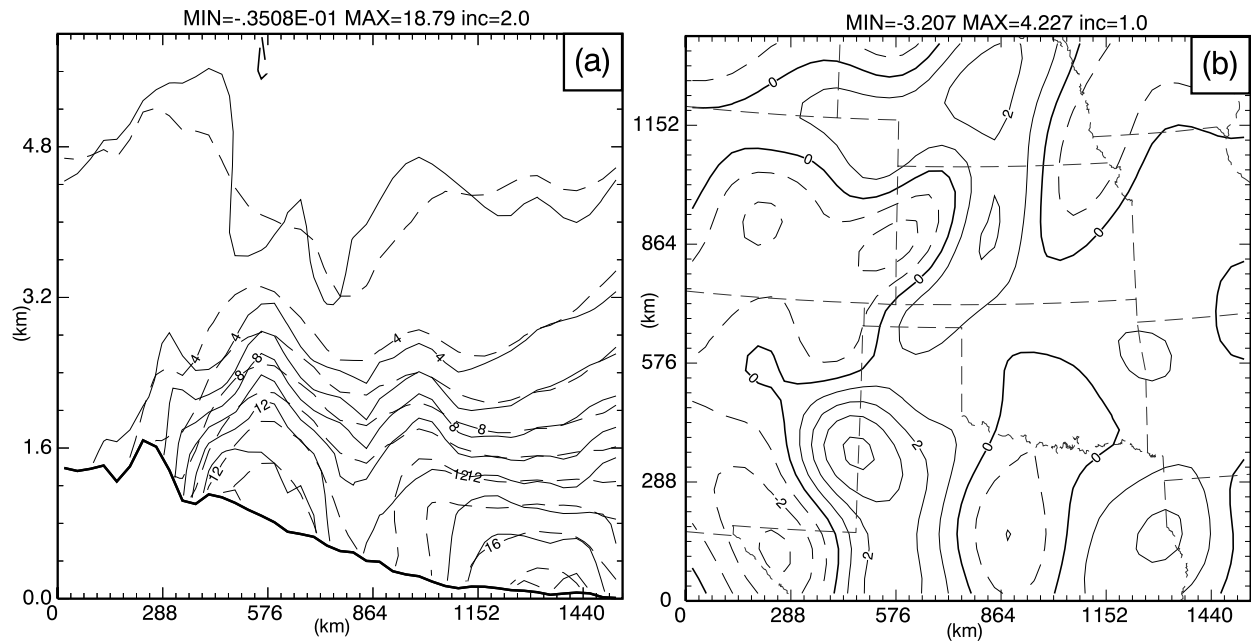
Figure 5. East-west vertical cross-section of specific humidity field (g kg^{-1}) at y = 234 km. The solid lines are for 'nature' and the dotted lines are for CNTL.

Figure 6. Specific humidity increment field in g kg^{-1} at the surface (a) from 'nature' and (b) from CNTL. Dashed lines represent negative values and solid lines positive values.

MIN=-.3508E-01 MAX=18.79 inc=2.0

MIN=-3.207 MAX=4.227 inc=1.0

Figure 7. (a) East-west vertical cross-section of specific humidity field (g kg^{-1}) at y = 234 km where solid lines are for 'nature' and dotted lines for experiment SNF. (b) Analysis increment of specific humidity (g kg^{-1}) at the surface from experiment SNF, where dashed lines are for negative values and solid line for positive values.

MIN=0.2092 MAX=18.65 inc=2.0   MIN=-3.731 MAX=4.589 inc=1.0

Figure 8. As Fig. 7 but for experiment SUF.

MIN=1.292 MAX=18.51 inc=2.0

Figure 9. As Fig. 5 but dotted lines are from experiment LNF.

MIN=-.4949 MAX=20.41 inc=2.0

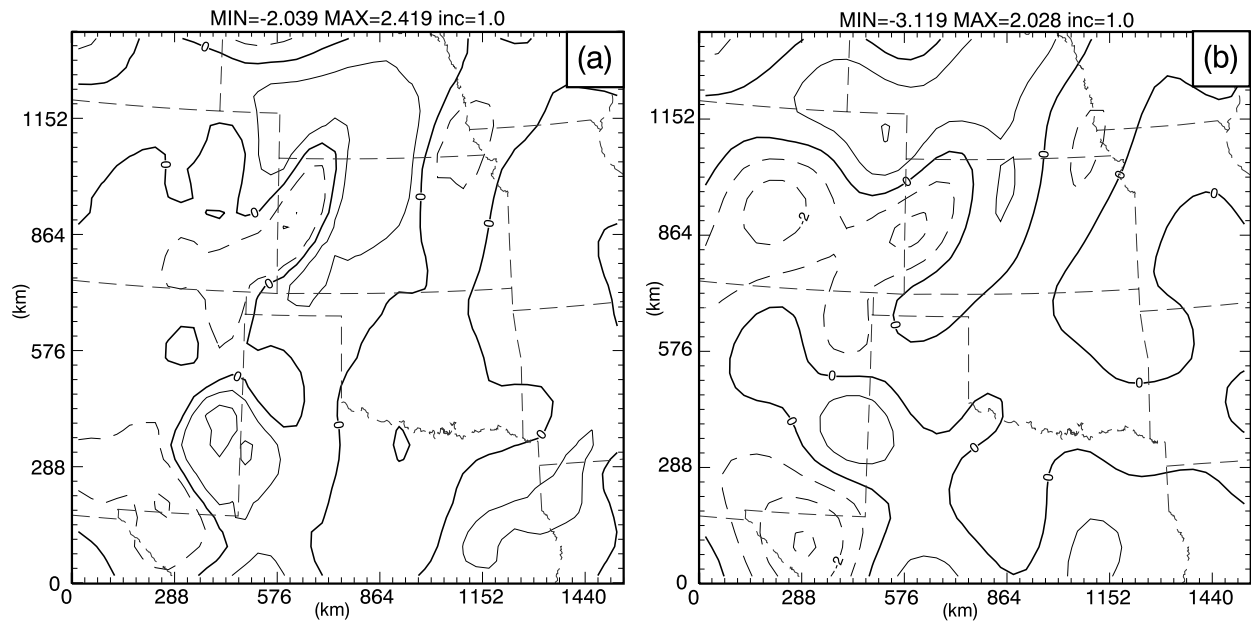Figure 10. As Fig. 5 but dotted lines are from experiment LTF.

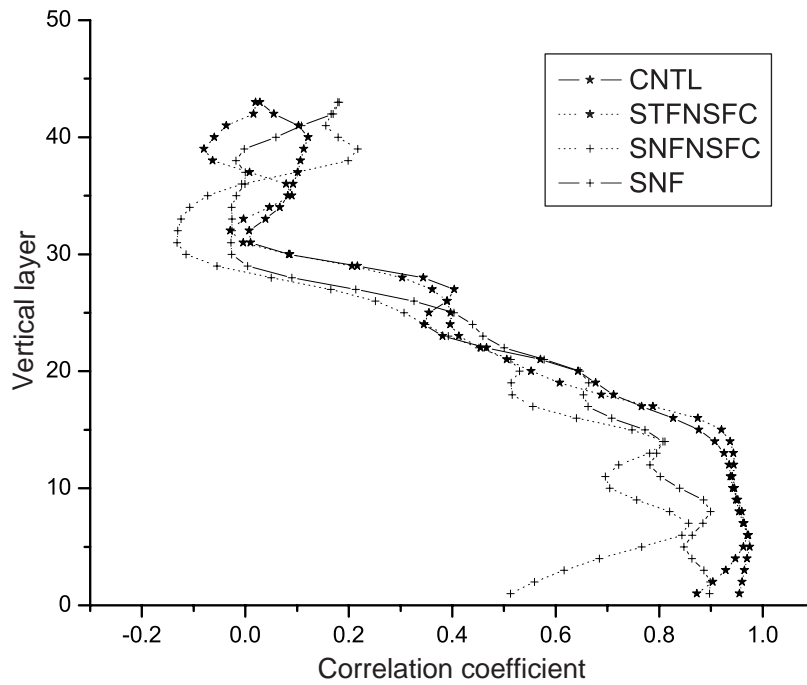Figure 11. As Fig. 6 but (a) is for experiment STFNSFC (b) is for experiment SNFNSFC.

Figure 12. Profiles of correlation coefficient of specific humidity increment (difference from background, in g kg^{-1}) between those of 'nature' and 3DVAR analysis from experiments CNTL, STFNSFC, SNFNSFC, and SNF, plotted for different model levels. Mean height of each level is given in a footnote in the paper.
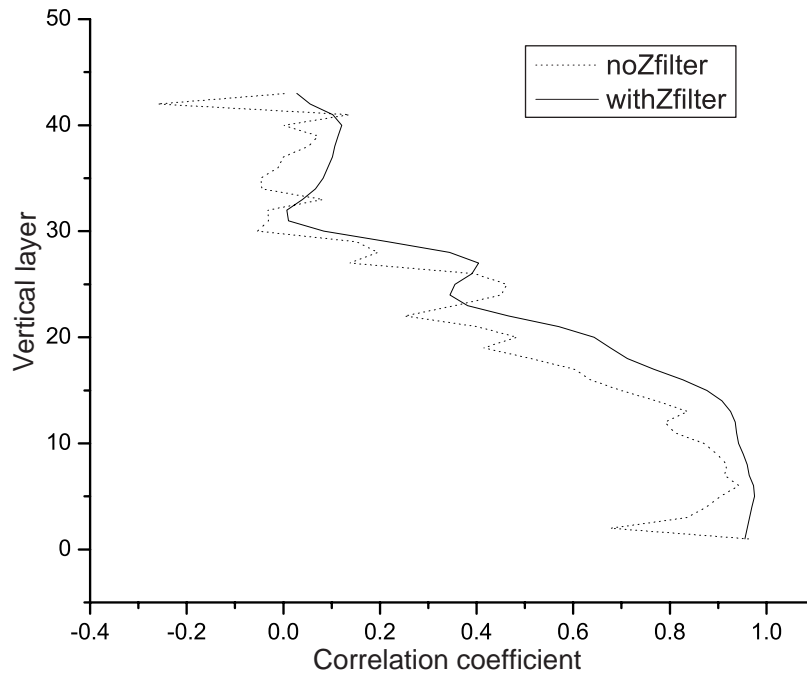
Figure 13. As Fig. 12 but for experiments CNTL (solid line) and STFNVF (dotted line).
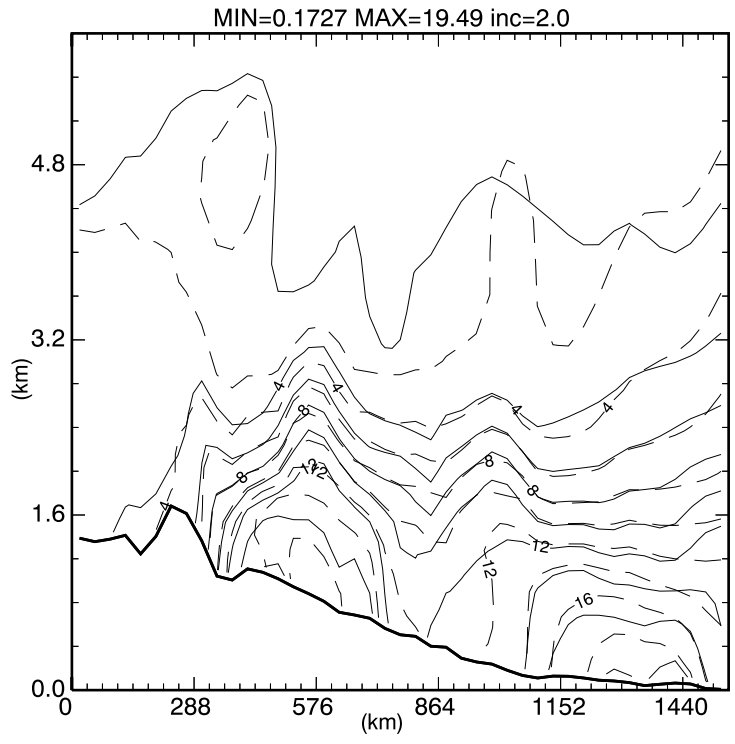
MIN=0.1727 MAX=19.49 inc=2.0

Figure 14. As Fig. 5 but dotted lines are for experiment STF_ER.

Figure 15. As Fig. 7 but for experiment STF_LR.