QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. **133**: 969–980 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.79

RMetS
ROYAL METEOROLOGICAL SOCIETY

A variational data assimilation system for ground-based GPS slant delays

Heikki Järvinen,[a] Reima Eresmaa,[a]* Henrik Vedel,[b] Kirsti Salonen,[a] Sami Niemelä[a] and John de Vries[c]

[a] *Finnish Meteorological Institute, Helsinki, Finland*
[b] *Danish Meteorological Institute, Copenhagen, Denmark*
[c] *Royal Netherlands Meteorological Institute, de Bilt, The Netherlands*

ABSTRACT: Tropospheric delay affects the propagation of the microwave signals broadcast by the Global Positioning System (GPS) satellites. Geodetic processing software enable estimation of this effect on the slanted signal paths connecting the satellites with the ground-based receivers. These estimates are called slant delay (SD) observations and they are potentially of benefit in numerical weather prediction.

The three-dimensional variational data assimilation system of the High-Resolution Limited-Area Model (HIRLAM 3D-Var) has been modified for data assimilation of the SD observations. This article describes the ground-based GPS observing system, the SD observation operator, the estimation of the observation- and background-error standard deviations, the methodology of accounting for the observation-error correlation, and the tuning of the background quality control for this observation type.

The SD data assimilation scheme is evaluated with experiments utilizing hypothetical observations from a single receiver station, as well as a single-case experiment utilizing real observations from a regional GPS receiver network. The ability of the data assimilation system to extract the asymmetric information from the SD observations is confirmed. In terms of analysis increment structure and magnitude, SD observations are found to be comparable with other observation types currently in use, provided that the observation-error correlation is taken into account.

KEY WORDS GPS meteorology; numerical weather prediction; observing system

Received 17 October 2006; Revised 1 March 2007; Accepted 12 March 2007

1. Introduction

The geodetic applications of the Global Positioning System (GPS) are based on measuring the carrier phases of the microwave signals broadcast by the satellites and received by the ground-based receivers (e.g. Hofmann-Wellenhof *et al*., 2001). The measurements are affected by a number of uncertainty factors. One of these is the tropospheric effect on the signal propagation, the so-called tropospheric delay. In the case of a sufficiently dense network of geodetic GPS receivers, the tropospheric delay can be estimated as a by-product of the geodetic processing. The tropospheric delay estimates can be considered as meteorological observations containing signatures of the atmospheric humidity field, and they form the basis for the concept known as GPS meteorology (Bevis *et al*., 1992). These scalar-valued observations have dimension of length. For a near-zenith satellite and a receiver at mean sea level, the tropospheric delay varies between two and three metres.

GPS meteorology provides potential for both subjective forecasting and nowcasting as well as for the numerical weather prediction (NWP). The latter application is strongly coupled with the process of data assimilation. The data assimilation produces the maximum likelihood estimate of the atmospheric state, which is used as the initial condition for the numerical forecasts. Development of NWP systems towards fine-scale atmospheric modelling sets requirements for the resolution of the observing systems. Ground-based GPS receivers are recognized as a potential source of high spatio-temporal resolution observations of atmospheric humidity. The European meteorological and geodetic communities are collaborating to advocate the use of the ground-based GPS receiver networks as an observing system for future operational NWP systems (e.g. Elgered *et al*., 2005; Vedel *et al*., 2006). Such a system is now being set up in the E-GVAP project (EUMETNET GPS Water Vapour Programme, http://egvap.dmi.dk).

It is not yet known precisely how the GPS measurements should be processed in order to provide the highest benefit to NWP models. The currently used geodetic software support processing of zenith total delay (ZTD), which is an estimate of the tropospheric delay in a vertical

* Correspondence to: Reima Eresmaa, P.O. Box 503, FI-00101, Helsinki, Finland. E-mail: Reima.Eresmaa@fmi.fi

column above each receiver station (e.g. Elgered *et al*., 2005). By using surface meteorological data of pressure and temperature, ZTD can be separated into hydrostatic and wet contributions (zenith hydrostatic delay, ZHD and zenith wet delay, ZWD), and the latter further converted into integrated water vapour (IWV) (Bevis *et al*., 1992). These estimates are essentially weighted means of the vertical projections of the corresponding quantities along the individual slanted signal paths.

It is obvious that part of the meteorological information is lost in the ZTD processing. A single ZTD estimate gives no indication about the azimuthal asymmetry, information that is contained in the raw GPS measurements. This asymmetry is typically most significant in the vicinity of strong horizontal humidity gradients, such as frontal regions, often accompanying severe weather phenomena (e.g. Koch *et al*., 1997). In such cases, the use of the ZTD estimates can be considered suboptimal. It would be more beneficial to process the raw GPS measurements in a manner which conserves the asymmetry. One possibility to achieve this is to estimate the tropospheric delay along the individual signal paths. Efficient meteorological use of such slant delay (SD) observations requires further development of the data assimilation systems.

So far, the majority of the ground-based GPS data assimilation experiments have been conducted using the ZTD, ZWD or IWV estimates (e.g. De Pondeca and Zou, 2001; Pacione *et al*., 2001; Vedel and Huang, 2004; Vedel and Sattler, 2006 and references therein). In general it is found that the data assimilation of the ZTD has a weak, but mainly positive, impact on NWP forecasts of humidity and precipitation. Ha *et al*. (2003) studied data assimilation of slant wet delays with hypothetical data in a mesoscale NWP system. Hypothetical SD observations were used also by MacDonald *et al*. (2002) for diagnosing a three-dimensional humidity field. These studies show promising results for retrieving the humidity field on the basis of the SD measurements, even though it has not yet been demonstrated that the data assimilation of real SD observations actually improves the NWP forecast scores.

Methodology for data assimilation of SD observations has been implemented in the High-Resolution Limited-Area Model (HIRLAM; Undén *et al*., 2002) as preparation for impact studies. This article reviews the implementation. The variational data assimilation in the HIRLAM framework is described in Section 2. Section 3 discusses SD as an observation type. The data assimilation aspects specific to SD are covered in Section 4. The experiments evaluating the SD data assimilation system are described in Section 5. A summary and the main conclusions are given in Section 6.

2. The HIRLAM 3D-Var

The hydrostatic HIRLAM limited-area NWP system is developed in a co-operation between the national meteorological services of Denmark, Finland, Iceland, Ireland, the Netherlands, Norway, Spain and Sweden (Undén *et al*., 2002). There is ongoing co-operation between the HIRLAM and ALADIN (Bubnová *et al*., 1993) consortia aiming at the joint development of a non-hydrostatic mesoscale NWP system.

The analysis system of HIRLAM is based on the variational data assimilation technique (Gustafsson *et al*., 2001; Lindskog *et al*., 2001), in either three (3D-Var) or four dimensions (4D-Var). While the currently operational implementations use 3D-Var, the 4D-Var set-up is being fine-tuned for operational use in the near future.

The incremental formulation (Courtier *et al*., 1994) is applied in the HIRLAM 3D-Var in order to reduce the computational load of the analysis problem. The background field $\mathbf{x}_\mathrm{b}$ is updated with the analysis increment $\delta\mathbf{x}$ corresponding to the minimum of the incremental cost function

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^\mathrm{T}\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(H[\mathbf{x}_\mathrm{b}] + \mathbf{H}\delta\mathbf{x} - \mathbf{y})^\mathrm{T}\mathbf{R}^{-1}(H[\mathbf{x}_\mathrm{b}] + \mathbf{H}\delta\mathbf{x} - \mathbf{y}), \quad (1)$$

where $\mathbf{y}$ represents the observations, $\mathbf{B}$ and $\mathbf{R}$ are the error covariance matrices for the background and the observations, respectively, and H is the observation operator producing the model state in observation space. The incremental formulation allows one to update the high-resolution background with a low-resolution analysis increment, i.e. the dimensions of $\mathbf{x}_\mathrm{b}$ and $\delta\mathbf{x}$ need not be the same.

The incremental cost function (1) is based on the assumption that H is linear for realistic $\delta\mathbf{x}$ in the neighbourhood of $\mathbf{x}_\mathrm{b}$. The linearized version of the observation operator $\mathbf{H}$ is thus applied for projecting the analysis increment to observation space. Moreover, the iterative minimization of (1) makes use of the adjoint of the tangent-linear observation operator $\mathbf{H}^\mathrm{T}$, as seen from the cost function gradient

$$\nabla_{\delta\mathbf{x}} J = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^\mathrm{T}\mathbf{R}^{-1}(H[\mathbf{x}_\mathrm{b}] + \mathbf{H}\delta\mathbf{x} - \mathbf{y}), \quad (2)$$

which is calculated at each iteration step.

The analysis increment $\delta\mathbf{x}$ of the HIRLAM 3D-Var consists of the increments of horizontal wind components (u and v), temperature (T), specific humidity (q) and logarithm of surface pressure ($\ln p_\mathrm{s}$). Introduction of a new observation type requires the development of the observation operator (H, $\mathbf{H}$, $\mathbf{H}^\mathrm{T}$) and the estimation of the observation-error statistics for $\mathbf{R}$.

3. Processing of the SD observations

The nominal constellation of the GPS system consists of 24 satellites in circular orbits with a radius of about 26 600 km. The satellites broadcast a modulated signal on two carrier-wave frequencies in the microwave domain.

DOI: 10.1002/qj

The most accurate positioning applications make use of the phase pseudoranges, i.e. phases of the received carrier waves. The charged particles at high atmospheric altitudes affect the wave propagation causing the ionospheric refraction. This frequency-dependent effect can be largely eliminated by computational methods in case the receiver has a capacity of measuring the carrier-wave phases at the two frequencies (e.g. Hofmann-Wellenhof *et al*., 2001).

The GPS processing deals with the raw pseudorange measurements. Within the processing, a number of unknown parameters are either estimated or eliminated. The tropospheric delay is one of the contributing effects in the measurements. It is generally not possible to directly estimate the SD along each individual signal path. Therefore, SD processing is usually done as a two-step procedure (de Haan *et al*., 2002).

In the first step, ZTD is estimated as part of the geodetic network solution simultaneously with the satellite and receiver positions and the clock errors. Accurate estimation of ZTD requires separation between ZHD and ZWD. The separation makes use of surface pressure data at the receiver locations. Pre-defined hydrostatic and wet mapping functions (m_{h} and m_{w}) are used for relating the ZHD and ZWD to the corresponding contributions of slant delay at zenith angle ζ. Once the zenith delays are solved, the *symmetric* part of SD, SD^{s}, is obtained as

$$SD^{\mathrm{s}} = m_{\mathrm{h}}(\zeta)\ ZHD + m_{\mathrm{w}}(\zeta)\, ZWD. \tag{3}$$

In the second step, the least-squares residuals of the network solution, obtained in the first step, are interpreted as the asymmetric part of SD, SD^{a}. The final SD is then

$$SD = SD^{\mathrm{s}} + SD^{\mathrm{a}}. \tag{4}$$

The undesirable slowly varying effects of (e.g.) signal multipath propagation and antenna phase-centre variations can be reduced by utilizing multipath maps that are generated by averaging the data over several days prior to the measurements. Due to the observing geometry, in high latitudes there are relatively more data available at large zenith angles.

The SD processing method is still subject to critical discussions among GPS scientists. The simulation study performed by Elosegui and Davis (2004) shows that a large part of the asymmetric information, contained in the raw measurements, is lost during the processing of SD. This is a consequence of the least-squares fitting method, which aims at finding the parameter values that provide a best fit to all observational data. Obviously, this is a severe limitation for the information content of the SD observations and possibly also for their value to the NWP models.

As the information content of the SD observations is somewhat questionable, it is critical to assess the asymmetry properties of the observations before drawing conclusions from extensive assimilation experiments. Such an assessment has been conducted recently by Eresmaa *et al*. (2007) from the observation modelling perspective. The conclusions suggest that the asymmetric contribution to the SD observations is of the order of a few parts per thousand of the total SD, and it is most significant at large zenith angles. Moreover, the asymmetry contained in the SD observations is found to be related to real atmospheric asymmetry in the vicinity of the GPS receiver stations.

Another point of concern is the observation-error correlation, which is expected to be significant. The simultaneous processing of all raw delay measurements makes the ZTD observations in the receiver network mutually dependent, implying correlated observation errors. Furthermore, several SD observations are generated from a single ZTD observation. Consequently, the ZTD observation errors propagate through the mapping functions to the SD observations making the SD observation errors correlated at one time instant at one receiver station. This implies that the corresponding SD observations are mutually dependent.

4. SD data assimilation characteristics

This section reviews the specific actions taken in order to enable the SD data assimilation with the HIRLAM 3D-Var system. Since the SD observation operator has already been discussed in detail by Eresmaa and Järvinen (2006), it is described only briefly here.

4.1. Observation operator

The observation operator is built upon the definition

$$SD = \int_s (n-1)\,\mathrm{d}s = 10^{-6} \int_s N\,\mathrm{d}s, \tag{5}$$

where n is the real part of the atmospheric refractive index, $N \equiv 10^6\,(n-1)$ is refractivity, s is the signal path from the satellite to the receiver and $\mathrm{d}s$ is a signal path element in length units. This definition for tropospheric delay is commonly used in geodetic literature (e.g. Hofmann-Wellenhof *et al*., 2001).

The observation operator consists essentially of three algorithms. First, the slanted signal path, s, through the NWP model grid is determined. The signal path is defined as a set of coordinates (latitude φ, longitude λ and height above the mean sea level h) for the intersections with the NWP model levels. The algorithm for the signal path determination makes use of the concept of geometrical path (GP) given by the satellite azimuth and zenith angles corresponding to the satellite–receiver geometry. The refractive bending of the signal path is taken into account by an explicit correction on top of the GP approximation. (Initial tests with an alternative approach for the signal path determination by ray tracing (Rodgers, 2000; Healy, 2001) seem to produce similar results when applied to HIRLAM.)

Second, the NWP model state is projected from the model grid to the signal path intersections using bi-linear

DOI: 10.1002/qj

interpolation in the horizontal. Once the model variables are interpolated, refractivity N at the intersections is obtained through

$$N = \frac{k_1 p}{T} + \frac{(k_2 - k_1)\, qp}{(0.622 + 0.378q)\, T} + \frac{k_3 qp}{(0.622 + 0.378q)\, T^2}, \quad (6)$$

where p is pressure, T is temperature, q is specific humidity and the empirical coefficients are $k_1 = 77.60$ K hPa^{-1}, $k_2 = 70.4$ K hPa^{-1} and $k_3 = 3.739 \times 10^{-5}$ K^2 hPa^{-1} (Bevis *et al.*, 1994). Relation (6) is derived from the ideal gas assumption.

Third, the refractivity is integrated using the model state variables to yield SD. The refractivity integration is done layer by layer in the vertical. For each layer of depth h, the parameters a and b of the exponential function

$$N = \exp(a + bh) \quad (7)$$

are uniquely determined from values of N at the bottom and the top of the layer. Use of these parameters allows piecewise analytical integration.

The observation operator has been validated against observational SD data. The model counterparts are produced for an SD dataset consisting of nearly 400 000 observations with satellite zenith angle cutoff of 80°. The observations are processed for 17 permanent receiver stations in Northwest Europe and they cover the time period 1–24 May 2003. The dataset is processed at the Technical University of Delft, the Netherlands. The root-mean-square difference between the observed and the modelled SD is about 2 cm near the zenith and about 10 cm at the zenith angle of 80°. Also the tangent-linear observation operator and its adjoint are implemented and evaluated with appropriate tests.

One practical issue is recognized as a potential problem specific for the SD observations. Due to the high demands on the computational efficiency, the current NWP systems are usually implemented on parallel computers, such that each processor handles only one geographical area, i.e. subdomain, at a time. It can occasionally happen that the SD signal path at a large zenith angle intersects two or more subdomains. For such observations, the forward model and the data assimilation code will be significantly more complicated. In the current HIRLAM implementation, those observations are simply rejected in the data assimilation. The drawback from this simple solution is that the details of parallelization affect the selection of observations and further the resulting analysis. The parallelization of the SD data assimilation in HIRLAM is still an open technical question.

4.2. Estimation of the error standard deviations

The data assimilation algorithms are derived from the assumption that the observation- and the background-error distributions are Gaussian with expectation value zero. The observation- and background-error standard deviations (σ_o and σ_b) need to be explicitly specified. In case the assumption of unbiased observations cannot be justified, a bias correction algorithm also needs to be implemented. The HIRLAM data assimilation system does not yet include a bias correction for SD.

The statistical properties of the observation minus model background (OmB) departures provide the basis for estimation of σ_o and σ_b. The standard deviation of the OmB departure contains contributions of both the observation and the background errors. The contribution of the background error is determinated by the randomization method (Andersson *et al.*, 2000). This method makes use of a sample of Gaussian random vectors (here 500), each representing the background error in the model grid-point space. The randomization method simulates the propagation of the background errors from the model grid-point space to observation space by applying the tangent-linear observation operator to each of the Gaussian random vectors. The estimation method assumes horizontal homogeneity and independence of azimuth angle.

The standard deviation of the OmB departure shows a strong dependence on the satellite zenith angle (Eresmaa and Järvinen, 2006). Therefore, σ_o and σ_b are estimated separately for nominal zenith angles 0°, 15°, 30°, 40°, 50°, 60°, 65°, 70° and 75°. The estimation procedure is as follows:

(1) The variance of the OmB departure is calculated at each nominal zenith angle. Only the departures for which the zenith angle is within 0.5° of the nominal zenith angle are used.
(2) The randomization method is applied in order to obtain the background-error variance at each nominal zenith angle.
(3) At each nominal zenith angle, the background-error variance is subtracted from the OmB variance in order to obtain the observation-error variance. The observation and background errors are assumed to be uncorrelated.
(4) σ_o and σ_b at the nominal zenith angles are calculated as the square roots of the corresponding variances.

Finally, functional models for σ_o and σ_b are constructed assuming the form

$$f(\zeta) = \frac{c}{\cos\zeta + d} \quad (8)$$

as a function of zenith angle ζ. The parameters c and d are determined separately for σ_o and σ_b using the least-squares fitting method. The input data for the fitting is obtained in the form of σ_o and σ_b at the nominal zenith angles.

The results of the fitting are $c = 11.27$ mm and $d = -5.669 \times 10^{-2}$ for σ_o, and $c = 7.550$ mm and $d = 2.654 \times 10^{-3}$ for σ_b. Figure 1 shows the fitted curves as a function of zenith angle. For all zenith angles, σ_o exceeds σ_b. Moreover, the ratio σ_o/σ_b increases with increasing

DOI: 10.1002/qj

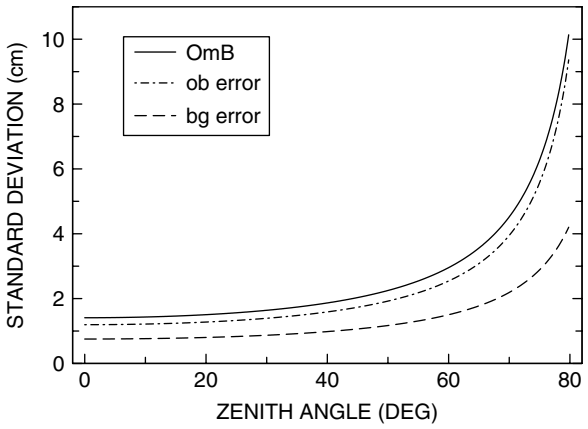

Figure 1. Standard deviations of SD OmB departure (solid line), SD background error (dashed line) and SD observation error (dash-dotted line) as functions of zenith angle.

zenith angle. This is interpreted as follows. With increasing zenith angle, the SD observation becomes increasingly sensitive to various noise terms, such as the signal multipath. Such sensitivity is absent from the SD model counterparts. Therefore, the increase with zenith angle is faster for σ_o than for σ_b. Consequently, the SD observations receive relatively smaller weight in the data assimilation than the model background, especially at large zenith angles.

It should be noted that the randomization method only projects the background-error statistics from grid-point space to observation space. The obtained σ_b values are therefore realistic only as far as the background-error covariance matrix is realistic. An alternative procedure to compute the error standard deviations based on the NMC method by Parrish and Derber (1992) has been applied by de Vries (2006) with generally supportive results.

4.3. Observation-error correlation

As pointed out in section 3, the observation-error correlations of SD are likely to be significant. Observation errors are assumed to be uncorrelated for most observation types currently assimilated in the HIRLAM system, the only exception being the Advanced TIROS Operational Vertical Sounder radiances (Schyberg *et al.*, 2003). The observation-error covariance in the HIRLAM context is estimated for the ZTD (Eresmaa and Järvinen, 2005). A covariance model for the SD based on elevation angle and site separation has been derived by de Vries (2006). The SD data assimilation system allows specification of the local error correlation properties between the SD observations at the same receiver station.

The dimension of the correlated blocks in the **R** matrix is of the order of 20 to 200, depending on the assimilation time-window length and the observation thinning details. The observation-error correlation is accounted for by following the procedure introduced in Järvinen *et al.* (1999).

Let us define the departure **d** between the analysis and the observations as

$$\mathbf{d} = H[\mathbf{x}_b] + \mathbf{H}\delta\mathbf{x} - \mathbf{y},$$

and the effective departure $\widehat{\mathbf{d}}$ as

$$\widehat{\mathbf{d}} = \mathbf{R}^{-1}\mathbf{d}.$$

The observation part of the incremental 3D-Var cost function (1) then becomes

$$J_o = \frac{1}{2}\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d} = \frac{1}{2}\mathbf{d}^T\widehat{\mathbf{d}}, \tag{9}$$

and the observation part of the cost function gradient (2) becomes

$$\nabla_{\delta\mathbf{x}} J_o = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{d} = \mathbf{H}^T\widehat{\mathbf{d}}. \tag{10}$$

Once the observation-error covariance matrix has been specified, the effective departure is obtained by solving the linear system

$$\mathbf{R}\widehat{\mathbf{d}} = \mathbf{d}, \tag{11}$$

for $\widehat{\mathbf{d}}$. In the HIRLAM implementation of SD data assimilation, (11) is solved by making use of the Cholesky decomposition for **R**, which is positive-definite. This procedure allows calculation of the cost function and its gradient in a general case of correlated observation errors. The sparseness of **R**, due to (e.g.) the lack of temporal correlations in the 3D-Var context, makes the decomposition tractable even in the case of many hundreds if not thousands of observations.

4.4. Observation quality control

The background quality control (BgQC) aims at identification and rejection of observations contaminated with non-Gaussian gross errors. The BgQC will reject the observation y_i if it does not satisfy the inequality

$$\frac{(H_i[\mathbf{x}_b] - y_i)^2}{\sigma_{o,i}^2 + \sigma_{b,i}^2} \leq L, \tag{12}$$

where $H_i[\mathbf{x}_b]$ is the non-linear model counterpart to y_i, $\sigma_{o,i}$ and $\sigma_{b,i}$ are the error standard deviations of the observation and the model counterpart, respectively, and L is the rejection limit (Järvinen and Undén, 1997; Lindskog *et al.*, 2001). The BgQC is tuned in HIRLAM for the SD observations by specifying L for this observation type.

Andersson and Järvinen (1999) showed how the rejection limit L can be tuned using a large number of OmB departures. First, a normalized departure z_i is calculated for all observations y_i as

$$z_i = \frac{y_i - H_i[\mathbf{x}_b]}{\sqrt{\sigma_{o,i}^2 + \sigma_{b,i}^2}}. \tag{13}$$

DOI: 10.1002/qj

A frequency histogram of the normalized departures is then constructed and it is compared with the normal distribution. Deviations from the normal distribution are expected to occur at the tails of the histogram; the rejection limit L will be chosen to correspond to the normalized departure value at which the frequencies considerably exceed the expected frequency of the normal distribution. In order to emphasize these deviations, a transformation

$$\widehat{f} = \sqrt{-2 \ln \left\{ \frac{f}{\max(f)} \right\}}, \tag{14}$$

is applied on the regular histogram. In (14), f and $\widehat{f}$ are the numbers of data in each bin of the regular and transformed histograms, respectively. A normal distribution with zero mean transforms into two straight lines intersecting at the origin.

Figure 2 shows the frequency histogram and the transformed histogram for the normalized departures of SD at the zenith angle intervals 25–35° and 65–75°. The number of observations in the latter interval is almost twice the number of observations in the former interval. The transformed histograms coincide with each other. In fact, the histograms are nearly identical for all zenith angle intervals (not shown). A simple test (not shown) indicates that a combination of a normal and a flat probability distributions (5% of sample from the flat distribution) results in a statistically similar sample to Figure 2. On the basis of Figure 2, the range of accepted normalized departures is chosen to be $-3 \leq z_i \leq 3$, implying $L = 9$ as a rejection limit.

Variational quality control (VarQC; Andersson and Järvinen, 1999) is not applied in the experiments described in this paper.

5. Evaluation of the SD data assimilation system

The SD data assimilation system is evaluated using two approaches. First, analysis increments resulting from data assimilation of hypothetical SD observations are studied. Second, real SD observations are assimilated and the analysis increments are compared with the analysis increments obtained from assimilating either radiosonde or ZTD observations. The data assimilation experiments are conducted using HIRLAM with a 9 km horizontal grid spacing and 40 levels in vertical. The analysis increment resolution is 44 km. The adopted physical parametrizations include the Soft Transition Condensation (STRACO) scheme for cloud condensation in large and convective scales (Undén *et al.*, 2002). As part of the system evaluation, a rough estimate is obtained for the SD observation-error correlation.

An arbitrarily chosen single case, 10 May 2003 at 00 UTC, is used for both experiments. The synoptic chart for this case (not shown) is dominated by a low pressure south of Iceland accompanied by a trough at the 500 hPa level extending to the North Sea. A weak cold front extends from the south coast of Sweden to the Netherlands separating the relatively dry air in Scandinavia from the moister air in Central Europe.

5.1. Data assimilation experiments with hypothetical observations

In the first evaluation approach, the analysis increments obtained from data assimilation of hypothetical SD observations are studied. A hypothetical GPS receiver is located at 57.40 °N, 11.93 °E. In the first run, a single SD observation at zenith angle zero is passed to the system. The observation is chosen such that the OmB is −30 mm, within the acceptable limits of the BgQC. The data assimilation of this single observation is expected to result in negative specific humidity analysis increments in the region of the observation.

Figure 3 shows the vertical profiles of the specific humidity background and the analysis increment at the hypothetical GPS receiver. Most of the humidity lies below the model level 30. Nevertheless, the analysis increment is spread throughout the lower troposphere. The strongest impact occurs at the model level 28 (≈800 hPa). The vertical range of the impact extends from the surface up to the model level 20, (≈540 hPa).

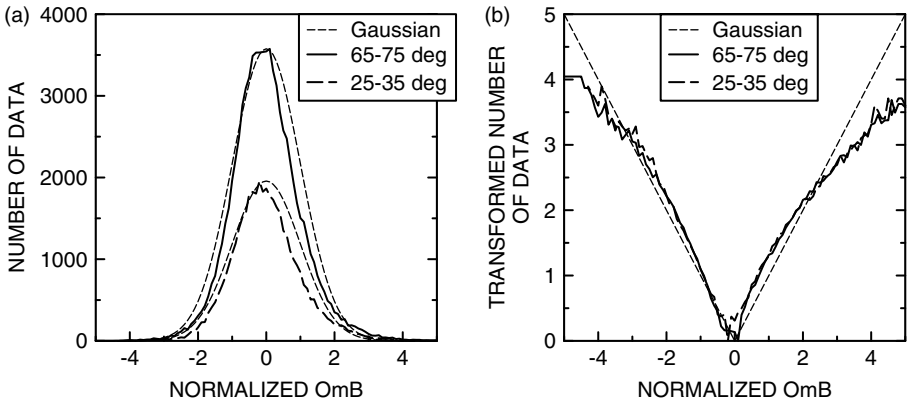

Figure 2. (a) regular and (b) transformed frequency histogram of the normalized OmB departure for SD at zenith angle intervals 25–35° (bold dashed line) and 65–75° (bold solid line). The Gaussian distributions (thin dashed lines) are shown for reference.

DOI: 10.1002/qj

Since the observation is a line integral through the atmosphere and only one observation is assimilated, the vertical distribution of the analysis increment reflects the specification of the background-error standard deviations in the data assimilation system. In the HIRLAM system, the maximum of specific humidity background-error standard deviation is close to 800 hPa, in agreement with Figure 3.

Figure 4 shows the distribution of the specific humidity analysis increment on the model level 28. As expected, the maximum impact coincides with the horizontal location of the hypothetical receiver. The analysis increment covers a circular area with radius of about 450 km, which is interpreted as the effective radius of influence for a single humidity observation in the HIRLAM 3D-Var data assimilation system. The analysis increments of other analyzed quantities are negligible (not shown) due to the univariate nature of the humidity analysis in the HIRLAM 3D-Var. Even though the SD observation operator imposes a dependency between humidity, pressure and temperature, the impact of SD data assimilation is mainly in the specific humidity field. This is because the background-error standard deviation for specific humidity is relatively large compared with those for pressure and temperature.

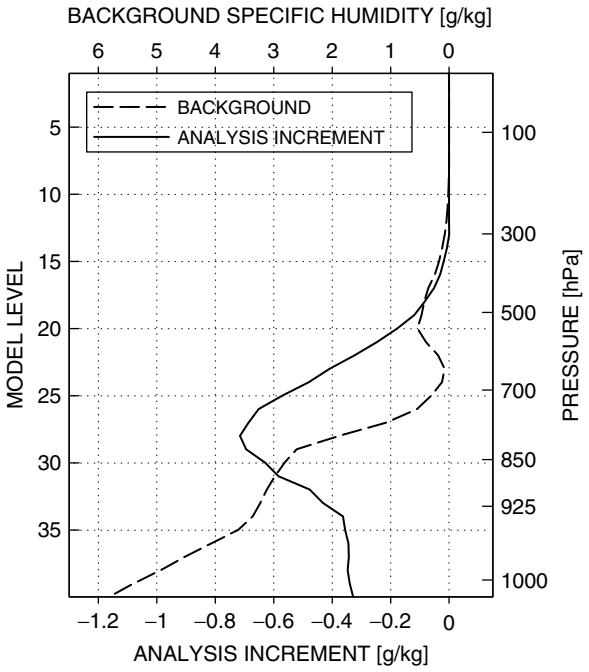

Figure 3. Vertical profiles of the background (dashed) and the analysis increment (solid) of specific humidity at the hypothetical GPS receiver location.

The second experiment aims at demonstrating the ability of the data assimilation system to produce asymmetric analysis increments from a set of SD measurements at a single receiver station. It is assumed that eight GPS satellites are simultaneously observed at the hypothetical receiver station. The properties of these observations and the corresponding OmB departures are listed in Table I. Note that the real GPS satellite constellation will never provide exactly this observation geometry; it is constructed only for the purpose of illustrating the capability of the SD data assimilation system to extract asymmetric information. With the NWP model grid spacing of 9 km, the signal paths of a pair of SD observations at opposite azimuth angles (e.g. observations 3 and 7 in Table I) are separated from each other by one grid point at 730 hPa and two grid points at 520 hPa. The observations are constructed such that they indicate existence of a sharp gradient in the humidity background error; in the east (west), the observed delay is larger (smaller) than the background delay. The observation errors are assumed uncorrelated.

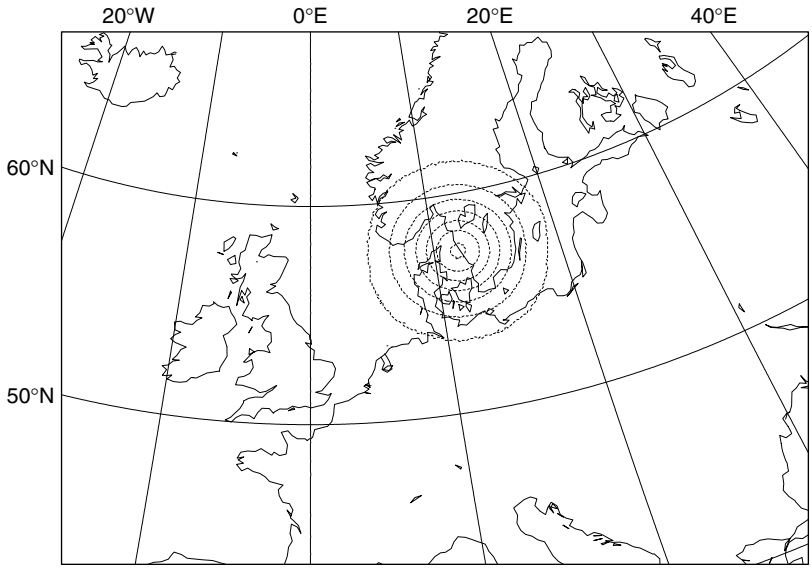

Figure 4. The specific humidity analysis increment on the model level 28 resulting from data assimilation of a hypothetical single SD observation at (57.40 °N, 11.93 °E). The contour interval is 0.1 g kg^{-1}, negative contours are dashed and the zero contour is omitted.

DOI: 10.1002/qj

Table I. Azimuth (α) and zenith angles (ζ) and OmB departures for the eight hypothetical observations.

	α (°)	ζ (°)	OmB (mm)
1	0	60	0
2	45	60	99
3	90	60	70
4	135	60	99
5	180	60	0
6	225	60	−99
7	270	60	−70
8	315	60	−99

Figure 5 shows the vertical cross-section in the zonal direction through the hypothetical receiver station. The analysis increment has a dipole structure with a zero increment along the vertical at the receiver station. The signal path of all hypothetical observations is within the influence radius of the other observations. Therefore, the positive increments in the east are largely cancelled out by the negative increments in the west, and *vice versa*. The maximum and minimum increments now occur at the model level 27, i.e. slightly higher in the model atmosphere than in the single observation experiment described above. This is understood as a consequence of the fact that the increments mostly cancel out at low altitudes, where the signal paths are relatively close to each other.

The distribution of the specific humidity analysis increment at the model level 27 (≈770 hPa) is plotted in Figure 6. The analysis impact is consistent with the observations indicating the sharp gradient in the background error of specific humidity. It is concluded that the SD data assimilation system is capable of extracting asymmetric information from asymmetric SD observations. Whether this information is present in real SD observations is another matter.

5.2. Data assimilation experiments with real observations

The second approach for the data assimilation system evaluation makes use of real SD, ZTD, and radiosonde observations. The objective in this experiment is to compare the different observation types in terms of overall magnitude and structure of specific humidity analysis increments. Given comparable observing networks, the analysis increments corresponding to each observation type should be comparable in a properly tuned data assimilation system.

The SD observations are gathered over a 3-hour period, centred at the analysis time 00 UTC, 10 May 2003. In the original dataset, the observations are made every 30 s. In order to reduce the number of data, only the hourly observations, valid at the analysis time and ±1 hour, are used. As a result there are 336 SD observations, generated from 47 ZTD observations, at 16 receiver stations. The background is taken at appropriate time (the so-called FGAT option) using 2-, 3- and 4-hour forecasts initiated at 9 May 2003, 21 UTC.

Concerning the radiosonde reports, all the observed quantities, including geopotential, temperature, specific humidity and wind, are assimilated. Only a suitable subset of all available radiosonde data is used. The assimilated reports are selected subjectively in such a way that the resulting radiosonde observing network covers approximately the same geographical area in Northwest

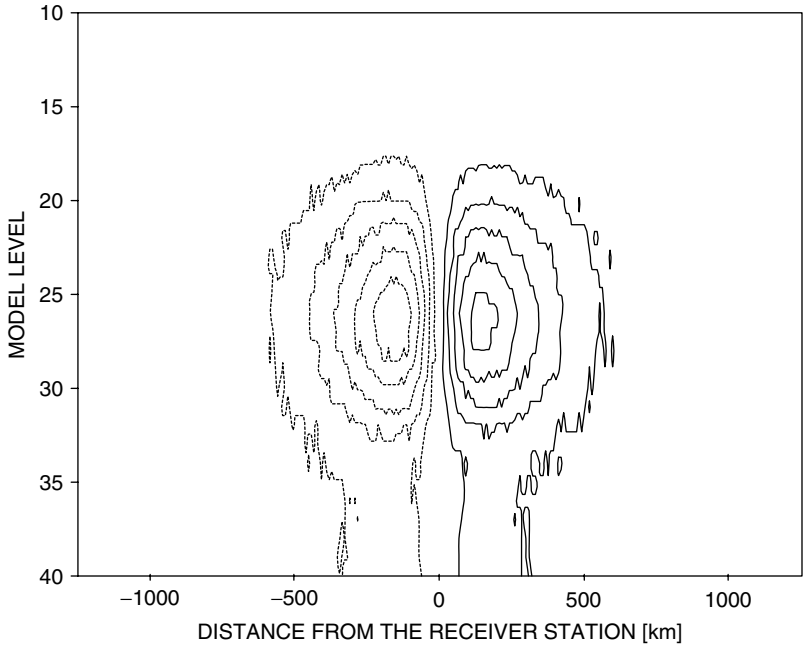

Figure 5. Vertical cross-section of the specific humidity analysis increment resulting from the data assimilation of eight hypothetical SD observations observed at a single receiver station. The contour interval is 0.02 g kg^{-1}.

DOI: 10.1002/qj

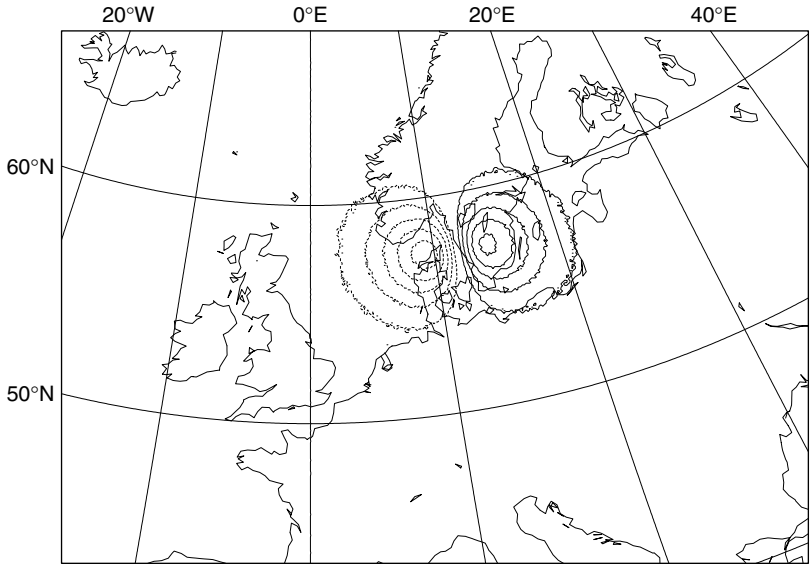

Figure 6. The specific humidity analysis increment on model level 27 resulting from the data assimilation of eight hypothetical SD observations observed at a single receiver station at (57.40 °N, 11.93 °E). The contour interval is 0.02 g kg^{-1}, negative contours are dashed and the zero contour is omitted.

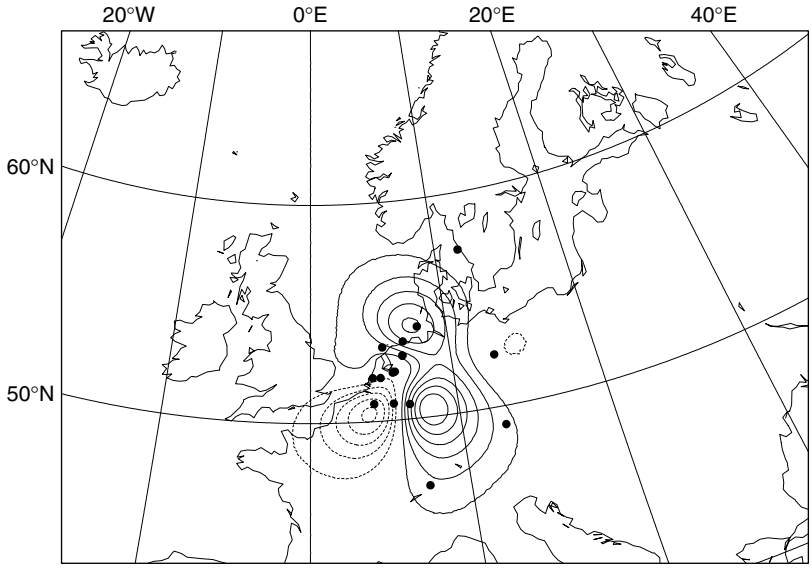

Figure 7. The specific humidity analysis increment on model level 27. Only the SD observations are assimilated, and observation errors are assumed uncorrelated. The contour interval is 0.2 g kg^{-1}, negative contours are dashed and the zero contour is omitted. The black dots indicate the ground-based GPS receiver station locations.

Europe and has a station density similar to the available GPS receiver network. The total number of assimilated radiosonde reports is 13 with 450 *in situ* humidity observations on different pressure levels.

5.2.1. Tests with uncorrelated observation errors

The experiment consists of performing a set of analyses with different observations but with the same background. In the first analysis, only the SD observations are assimilated, assuming uncorrelated observation errors. In the second analysis only the radiosonde observations are assimilated, also assuming uncorrelated observation errors. The specific humidity analysis increments corresponding to the two analyses are shown in Figures 7 and 8 for the model level 27. The dots in these figures indicate the locations of the GPS receiver stations and the radiosonde stations, respectively. The two analysis increments show similar structures. For instance, the positive and negative increments in the middle of the observing network occur in nearly the same locations. It is concluded that the two observing systems recognize the same error in the background humidity field. However, the magnitude of the SD analysis increment is larger than the magnitude of the radiosonde analysis increment throughout the area.

In the third analysis, only the ZTD observations are assimilated, again assuming uncorrelated observation

DOI: 10.1002/qj

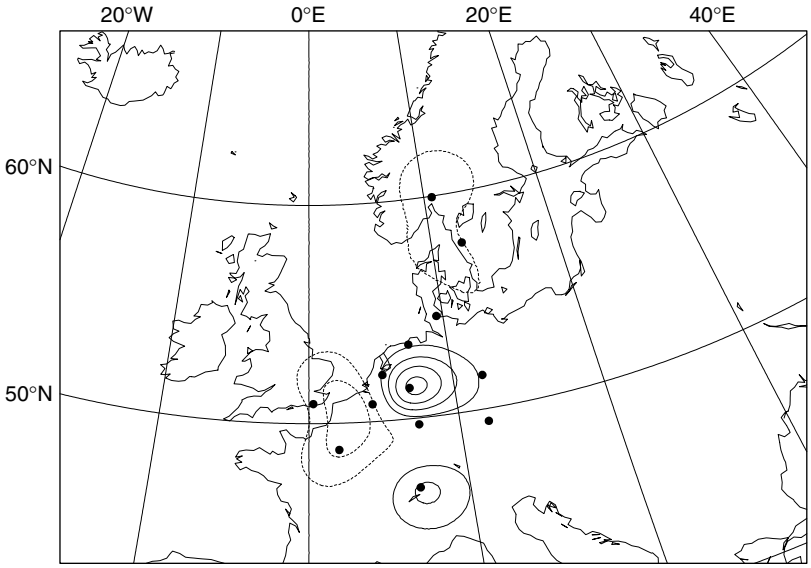

Figure 8. As Figure 7, but only the radiosonde observations are assimilated. The black dots indicate the radiosonde observation locations.

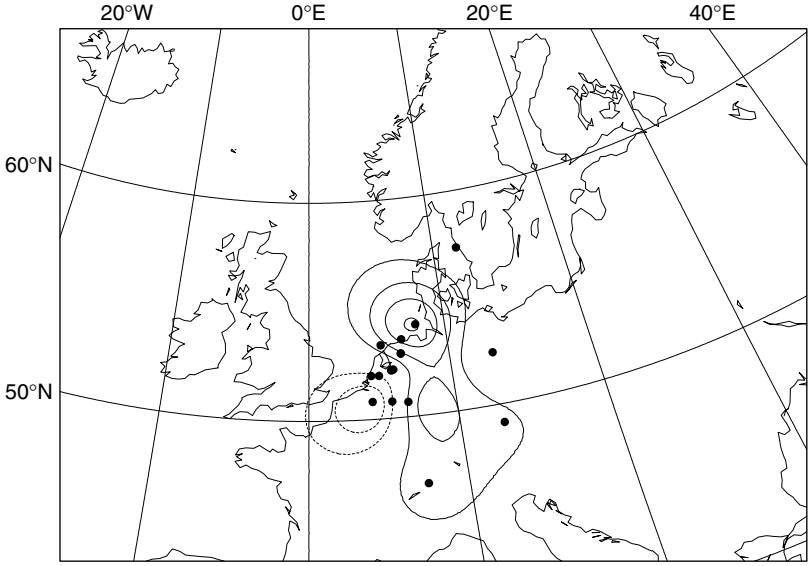

Figure 9. As Figure 7, but only the ZTD observations are assimilated.

errors. The magnitude of the resulting specific humidity analysis increment (Figure 9) is reduced compared with the SD analysis increment (Figure 7), but the general shape remains similar. There are variations in the relative magnitudes of the maxima and minima, which are probably due to the variations of the ratio between the numbers of available SD and ZTD observations in different parts of the observing network.

5.2.2. Tests with correlated observation errors

The effect of specifying the SD observation-error correlation is studied next by performing a number of analyses. The observation-error correlations between different receiver stations are assumed zero. The observation-error correlation between any two observations from the same receiver station is accounted for and characterized by a positive correlation coefficient ρ. The elements of the observation-error covariance matrix $\mathbf{R}$ are determined by

$$\begin{cases} R_{i,j} = \rho\sigma_{\mathrm{o},i}\sigma_{\mathrm{o},j} & \text{if} i \neq j, \\ R_{i,j} = \sigma^2_{\mathrm{o},i} & \text{if} i = j, \end{cases}$$

where $\sigma_{\mathrm{o},i}$ and $\sigma_{\mathrm{o},j}$ are the specified observation-error standard deviations for observations i and j, respectively. Setting $\rho = 0$ implies uncorrelated and $\rho = 1$ fully correlated observation errors.

The effect of varying ρ on the analysis is studied. With increasing ρ, the magnitude of the specific humidity analysis increment decreases systematically (not shown). The decrease is not uniform; in this single analysis, it is steeper above southern Germany than above the North Sea. The structure of the SD analysis increment with increasing SD observation-error correlation approaches that of the ZTD analysis increment.

DOI: 10.1002/qj

By comparing the analysis increments corresponding to different values of ρ with the analysis increment corresponding to radiosonde or ZTD data assimilation, it is concluded that a reasonable value for ρ appears to be somewhere between 0.3 and 0.5. The value 0.3 is chosen here as the most conservative value, i.e. $\rho = 0.3$ is the smallest observation-error correlation that appears to reduce the unrealistically large magnitude of the SD analysis increment. Figure 10 shows the corresponding analysis increment. The magnitude of the increment is now comparable with both the radiosonde and the ZTD analysis. Note that this correlation coefficient estimation is not intended to be conclusive, but rather a test of the sensitivity of the SD data assimilation system on assumed local SD observation-error correlation. Error correlation between SD observations from different receiver stations is also neglected at this stage.

Finally, the experiment with eight hypothetical observations from a single receiver station is repeated with $\rho = 0.3$. The dipole analysis increment structure is preserved. It is concluded that by taking the observation-error correlation into account, the data assimilation system is equally capable of extracting asymmetric information from SD observations.

6. Conclusions

This article describes the system development for data assimilation of the SD observations in the framework of the HIRLAM 3D-Var system. The system development consists of constructing the observation operator, estimation of the standard deviations for the observation and the background errors in observation space, implementation of the method to account for the observation-error correlations and tuning of the background quality control. The evaluation of the SD data assimilation system is performed with experiments using hypothetical and real SD observations as well as radiosonde observations.

The experiments with hypothetical data show that the data assimilation of the SD observations results in reasonable specific humidity analysis increments. As one of the main benefits of the SD observations lies in their ability to capture fine-scaled asymmetric atmospheric humidity structures, it is encouraging to note that the data assimilation system is able to retrieve the asymmetric information contained in the hypothetical observations (Figures 5 and 6).

Real SD observations are assimilated in an arbitrarily chosen single case. The resulting specific humidity analysis increments have a similar spatial structure to those obtained with a comparable radiosonde or ZTD observing networks. However, assuming uncorrelated SD observation errors leads to too large analysis increments. The analysis increment is decreased to a proper level, as determined by the analysis increments corresponding to radiosonde and ZTD observations, by accounting for the local error correlation between the SD observations from a single receiver station. As a rough estimate, a value of 0.3 is concluded to be appropriate for the correlation coefficient. The error correlation between the SD observations from different receiver stations is neglected in this study.

The data assimilation system will be subject to further development issues. These are planned to include a detailed investigation of the SD observation-error statistics, implementation of a bias-correction algorithm, and modification of the program code in order to enable parallel computing in a proper way. Furthermore, it is becoming evident that the current methodology for SD processing does not fully support the SD data assimilation (Elosegui and Davis, 2004). Emphasis will therefore be put on the processing methods in order to make best possible use of the inherently asymmetric atmospheric measurements at the ground-based GPS receiver stations.

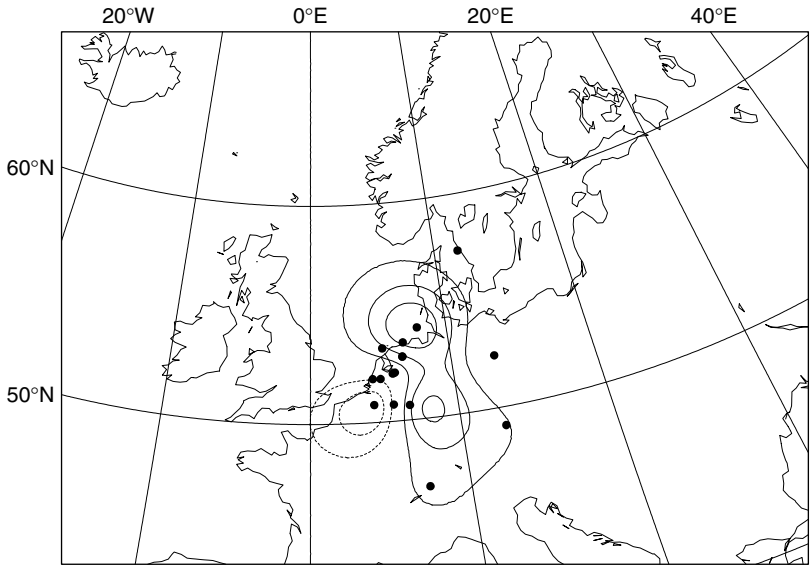

Figure 10. As Figure 7, but the SD observation-error correlation coefficient is set to 0.3.

DOI: 10.1002/qj

Acknowledgements

Two anonymous reviewers and the journal are acknowledged for a speedy revision process and for a number of useful comments and suggestions to improve the manuscript. The funding from the EU FP5 project 'Targeting Optimal Use of GPS Humidity Measurements in Meteorology' (TOUGH) in 2003–2006 and from the TEKES project 'Geophysically Assisted Satellite Positioning' in 2004–2006 is thankfully acknowledged. TOUGH is a shared-cost project (contract EVG1-CT-2002-00080) co-funded by the Research DG of the European Commission within the RTD activities of the Environment and Sustainable Development sub-programme (5th Framework Programme). TEKES is the Finnish Funding Agency for Technology and Innovation. The authors are grateful to Hans van der Marel (Technical University of Delft, the Netherlands) for fruitful discussions and for provision of the SD observations.

References

Andersson E, Järvinen H. 1999. Variational quality control. *Q. J. R. Meteorol. Soc.* **125**: 697–722.

Andersson E, Fisher M, Munro R, McNally A. 2000. Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.* **126**: 1455–1472.

Bevis M, Businger S, Herring T, Rocken C, Anthes R, Ware R. 1992. GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *J. Geophys. Res.* **97**: 15 787–15 801.

Bevis M, Businger S, Chiswell S, Herring TA, Anthes RA, Rocken C, Ware RH. 1994. GPS meteorology: Mapping zenith wet delays onto precipitable water. *J. Appl. Meteorol.* **33**: 379–386.

Bubnová R, Horányi A, Malardel S. 1993. 'International project ARPEGE/ALADIN'. *EWGLAM Newsletter* **22**: 117–130.

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**: 1367–1387.

de Haan S, van der Marel H, Barlag S. 2002. Comparison of GPS slant delay measurements to a numerical model: case study of a cold front passage. *Phys. Chem. Earth* **27**: 317–322.

De Pondeca M, Zou X. 2001. A case study of the variational assimilation of GPS zenith delay observations into a mesoscale model. *J. Appl. Meteorol.* **40**: 1559–1576.

de Vries J. 2006. Validation and error modelling of GPS tropospheric slant delays. Available as TOUGH project deliverable no D37, at http://tough.dmi.dk.

Elgered G, Plag H-P, van der Marel H, Barlag S, Nash J. 2005. 'Exploitation of ground-based GPS for operational numerical weather prediction and climate applications'. COST Action 716: Final report. European Union: Brussels.

Elosegui P, Davis J. 2004. 'Accuracy assessment of GPS slant-path determinations'. Proc. International Workshop on GPS meteorology, Tsukuba, Japan, 14–17 Jan 2003. Iwabuchi T, Shoji Y. (eds).

Eresmaa R, Järvinen H. 2005. Estimation of spatial GPS zenith delay observation error covariance. *Tellus* **57A**: 194–203.

Eresmaa R, Järvinen H. 2006. An observation operator for ground-based GPS slant delays. *Tellus* **58A**: 131–140.

Eresmaa R, Järvinen H, Niemelä S, Salonen K. 2007. Azimuthal asymmetry in ground-based GPS slant delay observations and their NWP model counterparts. *Atmos. Chem. Phys. Discuss.* **7**: 3179–3202.

Gustafsson N, Berre L, Hörnquist S, Huang X-Y, Lindskog M, Navascués B, Mogensen KS, Thorsteinsson S. 2001. Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus* **53A**: 425–446.

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

Healy S. 2001. Radio occultation bending angle and impact parameter errors caused by horizontal refractive index gradients in the troposphere: A simulation study. *J. Geophys. Res.* **106**(D11): 11 857–11 889.

Hofmann-Wellenhof B, Lichtenegger H, Collins J. 2001. *GPS. Theory and practice.* Springer-Verlag: Vienna, Austria.

Järvinen H, Undén P. 1997. 'Observation screening and background quality control in the ECMWF 3D-Var data assimilation system'. Tech. Memo. 236. ECMWF: Reading, UK.

Järvinen H, Andersson E, Bouttier F. 1999. Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus* **51A**: 469–488.

Koch A, Aksakal A, McQueen J. 1997. The influence of mesoscale humidity and evapotranspiration fields on a model forecast of a cold-frontal squall line. *Mon. Weather Rev.* **125**: 384–409.

Lindskog M, Gustafsson N, Navascués B, Mogensen KS, Huang X-Y, Yang X, Andræ U, Berre L, Thorsteinsson S, Rantakokko J. 2001. Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus* **53A**: 447–468.

MacDonald AE, Xie Y, Ware RH. 2002. Diagnosis of three-dimensional water vapor using a GPS network. *Mon. Weather Rev.* **130**: 386–397.

Pacione R, Sciarretta C, Faccani C, Ferretti R, Vespe F. 2001. GPS PW assimilation into MM5 with the nudging technique. *Phys. Chem. Earth (A)* **26**: 481–485.

Parrish D, Derber J. 1992. The National Meteorological Centre's spectral statistical interpolation scheme. *Mon. Weather Rev.* **120**: 1747–1763.

Rodgers CD. 2000. *Inverse methods for atmospheric sounding. Theory and practice*. World Scientific: Singapore.

Schyberg H, Landelius T, Thorsteinsson S, Tveter FT, Vignes O, Amstrup B, Gustafsson N, Järvinen H, Lindskog M. 2003. 'Assimilation of ATOVS data in the HIRLAM 3D-Var System'. HIRLAM Technical Report No. 60. Available at http://hirlam.org/open/publications/TechReports/TR60.pdf.

Undén P, Rontu L, Järvinen H, Lynch P, Calvo J, Cats G, Cuxart J, Eerola K, Fortelius C, Garcia-Moya JA, Jones C, Lenderlink G, McDonald A, McGrath R, Navascues B, Nielsen NW, Ødegaard V, Rodriguez E, Rummukainen M, Rööm R, Sattler K, Sass BH, Savijärvi H, Schreur BW, Sigg R, The H, Tijm A. 2002. 'HIRLAM-5 Scientific Documentation'. Available at http://www.knmi.nl/hirlam/SciDoc_Dec2002.pdf or from Hirlam-5 Project, SMHI, S-60176, Norrköping, Sweden.

Vedel H, Huang X-Y. 2004. Impact of ground-based GPS data on numerical weather prediction. *J. Meteorol. Soc. Japan* **82**: 459–472.

Vedel H, Sattler K. 2006. Comparison of TOUGH impact studies with ground-based GPS observations. TOUGH project deliverable No. D49, at http://tough.dmi.dk.

Vedel H, Sattler K, Huang X-Y, Navascues B, Sanchez Arriola J, Garcia-Moya J, Faccani C, Ferretti R, Jupp A, Offiler D, Ridal M, Gustafsson N, de Vries J, Eresmaa R, Järvinen H. 2006. GPS Data Recommendations for European Numerical Weather Prediction. TOUGH project deliverable No. D73, at http://tough.dmi.dk.

DOI: 10.1002/qj