Journal of the Meteorological Society of Japan, Vol. 82, No. 1B, pp. 459–472, 2004

Impact of Ground Based GPS Data on Numerical Weather Prediction

Henrik VEDEL and Xiang-Yu HUANG

Danish Meteorological Institute, Copenhagen, Denmark

(Manuscript received 4 April 2003, in revised form 10 November 2003)

Abstract

A study on the impact of zenith total delays (ZTD's) obtained from observations by ground based Global Positioning System (GPS) receivers in Europe on the skill of numerical weather prediction (NWP) is presented.

The ZTD depends mainly on the local pressure at the GPS site and the integrated water vapour in the column above. Both pressure and humidity information are vital to NWP models, but it is mainly as a source of humidity information the ZTD's are expected to be significant to NWP models. Given the currently meagre amount of humidity observations the inclusion of ZTD's in the data assimilation, determining the starting states for NWP forecasts, is expected to improve forecast skill. Groups in Europe, the US, and Japan are carrying out impact and other studies to assess whether such expectations can be fulfilled.

We here present results from parallel runs, with and without ZTD data in the data assimilation, for the period of February 2002, using European ZTD data from the COST Action 716 data sample. Data from 117 GPS stations are included in the assimilations. The runs are performed using the spectral version of the HIRLAM local area NWP model and its variational data assimilation system, HIRVDA, in 3DVar mode. The simulations have been carried out at both 0.45 and 0.15 degree resolution.

It is found that statistical verification against observations (heights, winds, temperatures, humidities) from stations on the EWGLAM list indicates mainly neutral impact of GPS ZTD's in this period, with the exception that systematic improvements are seen in the geopotential heights.

A detailed comparison of NWP forecasts of precipitation is performed against 12 hour rain gauge observations. Both contingency tables and in particular detailed comparisons (by eye) of precipitation maps show that for this period inclusion of ground based GPS ZTD observations improve the prediction of strong precipitation.

In combination with our previous results, (Vedel and Huang 2003), and those reported by other groups, this solidifies the conclusion that GPS ZTD data will improve NWP forecast skill of precipitation.

Questions remain to be solved before ground based GPS data can be used in an optimal way in NWP models. These include better knowledge of the errors and error correlations of both the ZTD observations and of the NWP first guess field, as well as better ways of comparing model and observed precipitation. The new European project TOUGH (Targeting Optimal Use of GPS Humidity Observations in Meteorology) will address some of these issues.

1. Introduction

Within the framework of the MAGIC (Meteorological Applications of Global Positioning System Integrated Column Water Vapor Measurements in the Western Mediterranean) project the Danish Meteorological Institute (DMI) has been partaking in the validation of GPS derived atmospheric delay estimates against independent data from radiosondes and from the numerical weather prediction (NWP) model HIRLAM, which is used operationally at the DMI. The essence of this work, reported in

---

Corresponding author: Henrik Vedel and Xiang-Yu Huang, Danish Meteorological Institute, Lyngbyvej 100, DK 2100 Copenhagen, Denmark.
E-mail: hev@dmi.dk and xyh@dmi.dk

Vedel et al. (2001) and references therein (see also Pugnaghi et al. (2002) for a comparison of GPS data with sun photometer observations), is that GPS zenith total delays (ZTD's) compare well with other independent estimates. The GPS delays are in general better matched by the other measurements, with a standard deviation $\sigma = 11.7$ mm against nearby radiosondes in the MAGIC sample (average distance = 22 km), than by the NWP model data with $\sigma = 17.1$ mm (numbers from Vedel et al. 2001), demonstrating that the observations contain information not fully known by the model.

The ZTD can be split into a dry component, which to a fine approximation depends solely on local pressure and the position of the GPS site, and a wet component which depends mainly on the total amount water vapour per area in the column above the site. GPS observations themselves do not allow this splitting, but it is possible for radiosonde and NWP model data. While the dry delay is one order of magnitude larger than the wet delay, the NWP and radiosonde dry delay correlate far better than the corresponding wet delays, $\sigma_{dry} = 3.7$ mm, $\sigma_{wet} = 14.7$ mm (Vedel et al. 2001). The belief is therefore that assimilation of ground based GPS data can improve the analysis fields mainly by providing auxiliary humidity information.

In the following we describe the results of an impact experiment comparing forecasts with and without using the GPS data as well as to meteorological observations. In section 2, the assimilation system is described with focus on the GPS specific parts; section 3 describes the simulations and observations, while section 4 contains the discussions and section 5 the conclusion.

The study reported here should be seen in connection with the impact study previously carried out by us. The results of that were reported at the GPS Met workshop in Tsukuba 2003 and are published in the proceedings of the workshop (Vedel and Huang 2003).

2. Assimilation of ground based GPS data

At DMI we use a high resolution limited area model (as opposed to global) called HIRLAM for our operational forecasts. This model was developed and is being maintained in a collaboration between a number of meteorological institutes in Europe. Further details about the DMI implementation of the HIRLAM model can be found in Sass et al. (2000).

The data assimilation is performed using the variational data assimilation tool HIRVDA (HIRLAM variational data assimilation) designed for HIRLAM. It has been developed for both 3D and 4D data assimilation, here we use it in 3DVar mode, similar to what is currently done operationally at DMI. Further details about the HIRVDA system itself can be found in Gustafsson et al. (2001) and Lindskog et al. (2001).

Software enabling assimilation of ground based GPS observations in the form of zenith total delay (ZTD), zenith wet delay (ZWD), or integrated water vapour (IWV) has been made for HIRVDA. The version used here contains improvements made by H. Vedel and N. Gustafsson, including the knowledge about how to best determine delays from meteorological data gained in the MAGIC project from comparison of GPS data with the HIRLAM model and radiosondes (Vedel et al. 2001).

2.1 *Observation operators for ZTD, ZWD, and IWV*

The purpose of an observation operator is to determine the estimated value of an observable given a model state. For observables depending on a number of model variables this will typically consists of two steps. A fairly general one determining model state properties at or above a given site, and a specific one determining the value of the observable given by the model estimate of the atmospheric properties at the site. It is the latter which is here referred to as the 'observation operator', in accord with the fact that the general part is handled by the existing, core HIRVDA software.

For the ground based GPS observations, the following is done regarding determination of model estimates of the GPS observables.

First, interpolations are performed in the NWP model grid in order to determine the estimated atmospheric properties at the location of the site of interest, resulting in a profile of temperature and humidity above the site as well as estimates of surface pressure and surface (geoid) altitude. Secondly, the profile obtained is ‘shifted’ up or down, shifting the surface height of the model to the altitude of the GPS antenna. This is done using the Majewski scheme (Majewski 1985), which shifts the whole atmospheric profile up or down, rather than just interpolating/extrapolating in the profile, thereby maintaining the boundary layer structure.

In NWP models like HIRLAM, geopotential heights and geopotentials are measured in a special way. The surface height is the geometric geoid height of the surface, while the bit further up is determined by adding the geopotential height increment from the surface and up. Prior to using the Majewski scheme, we therefore transform the GPS station height to the corresponding ‘pseudo’ geopotential height in the HIRLAM system of reference. Notice that this transformation is model specific, and it changes with the resolution and orography of the NWP model. The importance of the transformation increases with the offset in height between the GPS site and the model (poor resolution of actual orography), with the altitude itself and varies with latitude. It is largest for equatorial and polar regions. Further details can be found in Vedel (2000). After the vertical shift, the ZTD, ZWD or IWV for the GPS site can be calculated.

The observation operator for ZTD is written as,

$$ZTD = ZHD + ZWD, \tag{1}$$

where ZHD is the zenith hydrostatic delay. This we determine as

$$\begin{aligned} ZHD &= 2.2768\ 10^{-5}[\mathrm{m/Pa}]p_a/z_f \\ z_f &= 1 - 2.66\ 10^{-3}\cos(2\theta) \\ &\quad - 2.8\ 10^{-7}[\mathrm{m}^{-1}]h_a \end{aligned} \tag{2}$$

where p_a is the pressure at the antenna, h_a the antenna altitude, and θ the latitude of the GPS site. The above is based on the Saastamoinen approach (Saastamoinen 1972), which neglects the effects of temporal temperature variations upon the relation between proper height and pressure. However, it has been found that in most cases the Saastamoinen approach provides a very fine approximation to ZHD’s which are determined using more elaborate methods, e.g., Vedel et al. (2001). The form above is particularly simple to deal with when it comes to coding of the necessary derivative and corresponding adjoint of the observation operator.

The observation operator for ZWD is written as,

$$ZWD = \frac{R}{g_s(\theta)}\sum_{i=1}^{N} q_i\left(k' + \frac{k_3}{T_i}\right)(p_{i+1/2} - p_{i-1/2}) \tag{3}$$

where q is specific humidity, T is temperature, p is pressure, g_s is the gravitational acceleration at the surface as function of latitude. (We have verified that the low scale height of humidity in the atmosphere merits neglecting the variation of g with height in determination of ZWD for realistic humidity distributions and significant ZWD’s.) Notice that in the HIRLAM model pressure levels are counted from the top of the atmosphere and down and that pressure is known at grid cell boundaries while temperature and humidity are known at grid centres. R is the gas constant for dry air. For the k’s we use $k' = 2.21\ 10^{-7}$ K/Pa and $k_3 = 3.7\ 10^{-3}$ K^2/Pa (Bevis et al. 1994).

The observation operator for IWV in HIRVDA is listed here for completeness, it is not used in the experiments presented,

$$IWV = \frac{1}{g_s(\theta)}\sum_{i=1}^{N} q_i(p_{i+1/2} - p_{i-1/2}) \tag{4}$$

In HIRVDA, the best estimate of the atmospheric state is found by minimisation of the cost function,

$$\begin{aligned} J = J_b + J_o &= \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(H(\mathbf{x}_b) \\ &\quad + \mathbf{H}\delta\mathbf{x} - \mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x}_b) + \mathbf{H}\delta\mathbf{x} - \mathbf{y}), \end{aligned} \tag{5}$$

with respect to the analysis increment $\delta\mathbf{x}$. Here $\mathbf{B}$ is the error covariance matrix of the model first guess field and $\mathbf{R}$ is the error covariance matrix for the observations. The latter is normally assumed diagonal, and includes errors such as the error of representativeness, which takes into account effects arising from the offset in resolution between the NWP model grid and the various observation platforms. H maps the model state vector $\mathbf{x}$ onto observation space, $\mathbf{H}$ is the derivative of that with respect to the state vector, and is assumed linear. $\mathbf{x}_b$ is the first guess model state (an old forecast based on the previous analysis and valid at the time of the new analysis), and **y** is a vector representing all the observations.

3. The simulations

Using the HIRVDA data assimilation package described above and the spectral version of the HIRLAM forecast model we have conducted a number of simulations in order to study the impact of ground based GPS observations on forecasts.

The simulations cover the period February 2002. This period was chosen solely because a control run, without GPS data, had already been performed for other purposes with the same set of observations and boundary files are available. Weather wise the period contains nothing unusual.

The simulations have been made at two different resolutions covering two different areas, named G and E, respectively, in Fig. 1. The G model simulations provide boundaries for the E-model, whereas ECMWF (European Centre for Medium-Range Weather Forecasts) analyses are used as boundaries of the G-model. The resolutions of the models are 0.45 degrees and 0.15 degrees, respectively. Both have 31 layers in the vertical, using a hybrid sigma coordinate system. The Eulerian semi-implicit time steps were 3 and 1 minutes, respectively. The G-model has 202 × 190 grid points in horizontal and the E-model has 272 × 282 grid points.

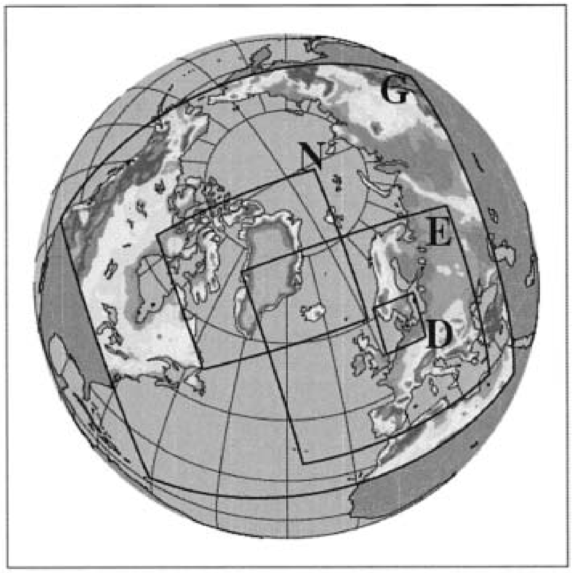


Fig. 1. Simulations were made for the domains G and E.

Data assimilation was carried out every 3 hours, starting at 00 UTC. First guess fields were 3-hour forecasts started from the previous analysis.

For each model domain data assimilation was performed with and without GPS ZTD data in the observations. The 'other' observations comprise traditional observational data, such as SYNOP (surface pressure), TEMP (radiosonde data), BUOY (surface pressure from the sea), AIREP (data from aircraft), SHIP (data from ships), and PILOT (wind data, as radiosonde winds). In addition, radiance data and (A)TOVS are assimilated, which is a significant difference in comparison with our previous study for June 2000 (Vedel and Huang 2003). We expect this to make it harder for the GPS data to have an impact.

The GPS observations are the COST Action 716 ZTD's available for the period. Further details about COST Action 716 and the GPS and ZTD data can be found at www.oso.chalmers.se/geo/cost716.html. When more than one ZTD value was available from a given GPS site within the time window of the data assimilation (3 hours), the ZTD closest to the central time of the data analysis time window was chosen.

Longer forecasts were made from the 00, 06, 12, and 18 UTC analyses, out to 48 hours, with forecasts being output for every six hours.

The GPS sites from which data have been assimilated are shown in Fig. 2. It is important to have this map in mind while considering the precipitation maps. However, it is to be noted that data were not available from all the GPS sites shown on the map for all data analyses.

The optimal error estimates to use for the GPS observables in data assimilation are currently not well known. It is clear that they have to be significantly larger than the 'instrumental' GPS ZTD error, in order to include the error of representativeness, which essentially corrects for the fact that a given type of observation often correspond to a property which varies itself within the volume sampled by each NWP model gridbox. For some observations, the error of representativeness can be the dom-

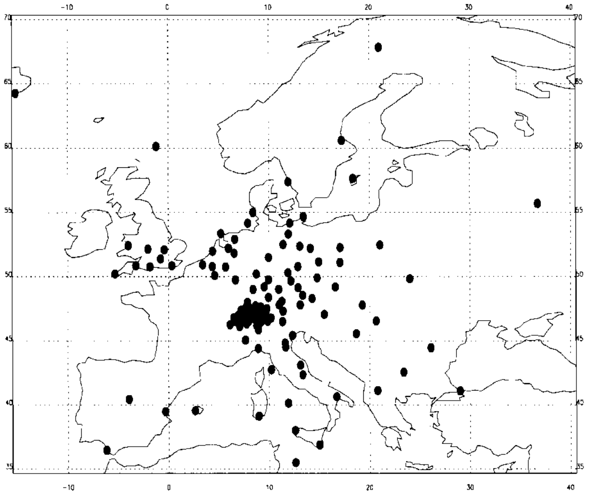


Fig. 2. Location of the 117 GPS sites providing data.

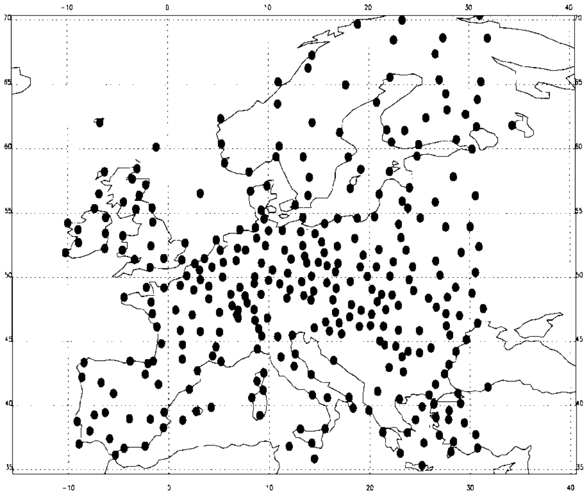


Fig. 3. Locations of the EWGLAM stations providing verification data.

inating observational error term. In the present study, the observational error of ZTD was set to 3 times the error estimated by the GPS ZTD processing software. This results in errors of the order 10 mm, which was found in Vedel et al. (2001) to be the correct order of magnitude. The errors of the GPS ZTD's are expected to be correlated, resulting in the proper R matrix being non diagonal as regards the ZTD part. This has been neglected.

4. Results

In the following sections we will study the impact of the GPS data on the forecasts both objectively through observational verification and subjectively by visually comparing forecasts with observations.

4.1 Statistical verification against standard observations

We have made observational verifications for each of the simulations. This is a standard way to measure performance skill of the meteorological models. It is made by comparing observations from a number of selected standard sites, in this case the sites from the EWGLAM (European Weather Group for Limited Area Modelling) list, to the NWP model estimates of the observations. Fig. 3 shows the locations of the EWGLAM stations. With a few exceptions, they are all land stations. Most of them provide near surface data (pressure, 2 metre temperature and 10 meter wind) and about 90 also provide radiosonde data, typically twice a day. GPS observations are thus not included in the verification of the present set of simulations.

The properties verified are geopotential heights, winds, temperatures and humidities at 850, 700, 500, 300 and 200 hPa, as well as mean sea level pressure (mslp), 2 m temperature, 10 m wind and 2 m humidity. Due to space limitations, Fig. 4 shows only three of the properties most commonly presented in verification studies, along with a verification plot for humidity. It is seen that the bias of mslp is slightly reduced, whereas the rms is slightly increased. The 850 hPa temperature bias is slightly reduced with the rms unchanged. The 500 hPa geopotential height bias is reduced, while the rms remains unchanged. The 700 hPa relative humidity bias is slightly reduced when using GPS ZTD data, while the rms is neutral, except at the analysis time. From the other statistics (not shown here), we found that the biases of the geopotential heights are systematically reduced and the geopotential rms' slightly reduced or unchanged. In contrast to this, the effect of the GPS data on the statistics of the wind, temperature and humidity is sometimes positive and sometimes negative, and often changing (in terms of positive/negative) over the 48 hour forecast period. (E-model statistics are not yet available).

In conclusion, the impact is neutral on most of the above properties, but positive on geopotential heights. The reason for this is not yet well understood, but it is a fine demonstration of the fact that the multivariate nature of the data assimilation and the many physical balances of the atmosphere represented in the analysis system enables one type of observation to improve many different types of the state variables in the region near the observation.

4.2 *Statistical verification against precipitation measurements*

4.2.1 *The measured precipitation data*

The precipitation data come from the operational database at DMI. This has a rather uneven spatial coverage from country to country and from time to time. The best coverage obtainable is found at 06 and 18 UTC in the form of 12 hour precipitation measurements. It would be interesting to compare precipitation over other times, in particular short term precipitation, but that has to await access to more detailed precipitation data.

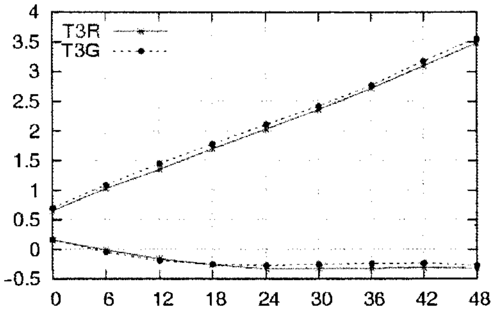
Twice a day, at 06 and 18 UTC, for February 2002 12 hour precipitation is extracted from the database along with the coordinates of the sites. No high level quality checking has been applied to the observations, but observations larger than 100 mm/12 hours are rejected.

4.2.2 *Calculation of model precipitation*

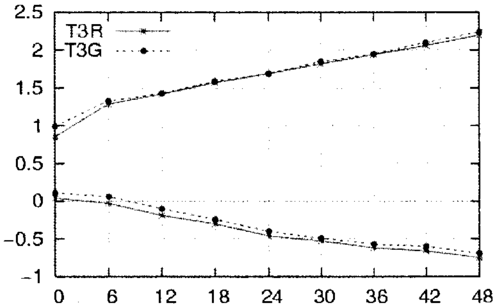
12 hour HIRLAM precipitation estimates for the rain gauge sites are produced in the following way: HIRLAM 12 hour forecasts valid at 06 or 18 UTC are selected. In the HIRLAM fields, precipitation is divided into 'large scale' and 'convective' precipitation. These are summed together and a horizontal interpolation is applied for the resulting field to the geographical locations of the rain gauge sites.

4.2.3 *Contingency tables for precipitation*

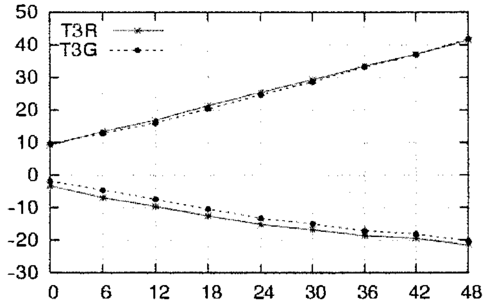
The observation and model precipitation data have been 'paired' in contingency tables of 12 hour precipitation. The tables are a count of how many precipitation pairs (predicted and observed) fall in which cell of the table. The tables are divided horizontally in five classes according to the observation and vertically in five classes according to the model prediction. The limits for the classes are the same in both directions: $0 \leq pr \leq 0.2$, $0.2 < pr \leq 1$, $1 < pr \leq$



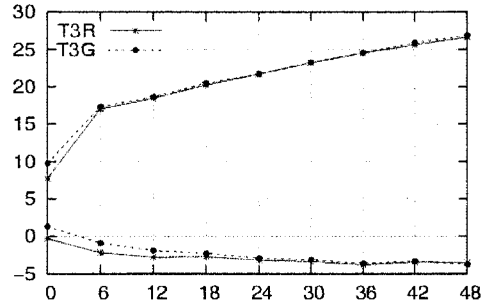
Mean sea level pressure [hPa]



850 hPa temperature [K]



500 hPa geopotential hieght [m]



700 hPa relative humidity

Fig. 4. Verification against EWGLAM observations for the G-model simlations. Filled circles represent run with GPS data, asterisks the run without GPS data. Forecast length in hours.

Table 1. Contingency table for 12 hour precipitation.

GPS, G-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	27.93	0.59	0.15	0.01	0.00	97
F1	20.14	3.05	1.50	0.14	0.04	12
F2	14.70	7.32	10.89	2.36	0.62	30
F3	1.79	0.63	2.76	2.27	0.98	27
F4	0.35	0.04	0.30	0.58	0.84	40
FO	43	26	70	42	34	88335

Table 2. Contingency table for 12 hour precipitation.

NO GPS, G-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	30.71	0.80	0.23	0.02	0.01	97
F1	20.40	3.97	2.16	0.20	0.03	15
F2	12.16	6.48	11.21	2.77	0.74	34
F3	1.40	0.35	1.84	2.02	1.05	30
F4	0.24	0.03	0.16	0.36	0.66	45
FO	47	34	72	38	26	88355

Table 3. Contingency table for 12 hour precipitation.

GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	26.76	0.83	0.33	0.06	0.03	96
F1	20.72	3.69	2.16	0.26	0.07	14
F2	13.46	6.81	10.61	2.41	0.68	31
F3	1.88	0.66	2.79	2.14	0.86	26
F4	0.54	0.13	0.48	0.74	0.94	33
FO	42	30	65	38	37	81891

Table 4. Contingency table for 12 hour precipitation.

NO GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	29.91	1.18	0.49	0.08	0.03	94
F1	19.92	4.32	2.75	0.28	0.08	16
F2	11.56	6.12	10.83	2.78	0.74	34
F3	1.56	0.44	2.00	1.96	0.97	28
F4	0.39	0.06	0.30	0.49	0.76	38
FO	47	36	66	35	29	81891

5, $5 < pr \leq 10$, $10 < pr$, where pr is the 12-hour precipitation and the units are mm (Similar limits are used for the verification of the operational HIRLAM forecasts at DMI). Perfect forecasts of precipitation would result in diagonal contingency tables. The numbers are turned into relative measures in order to ease inter comparison of the tables.

In the contingency tables presented the numbers in the body of the tables are given in percent of the total number of pairs, which are listed in the lower right corner. The row named 'FO' lists the percentage of the forecasted values falling in the same class as the named observation class. The column named 'OF' lists the percentage of the observations falling in the same class as the named forecast class. Ideally OF and FO should all be 100 and the inner part of the tables are to be diagonal.

Tables 1, 2, 3, and 4 are for the G and E model runs with and without GPS data. One can notice that the number of correct forecasts of class 4 and 5 is markedly better with GPS data (higher FO scores). On the other hand, the number of false alarms is also higher, resulting in lower OF scores. We shall return to this in the discussion. Further, all models runs are too wet when low precipitation is observed, and this problem increases when GPS data is used.

There have been reports on biases between NWP model estimates of ZTD, other observations of ZTD, and GPS ZTD's (e.g., Gustafsson 2002; Vedel et al. 2001). This has led to speculations whether the indications on an improved skill in NWP prediction of medium to large precipitation when ZTD data are included is due to a general increase of the precipitation at all precipitation levels.

To access the importance of such effects, we have determined the offsets in 12 hour precipitation between the two NWP data sets and the observations as well as between the two NWP data sets themselves. The results are given in the tables 5 and 6.

The median offset between the two NWP data sets is very small, while the average offset is of the order 0.28 mm. We have made a second set of contingency tables for precipitation for the NWP GPS forecasts in which a 'bias correction' was included, such that $NWP_{precip} \rightarrow \max\{(NWP_{precip} - \Delta_{precip}), 0\}$, where Δ_{precip} is taken from tables 5 and 6.

The results when bias correcting for the median of the offsets between the NWP model

Table 5. Statistics of 12 hour precipitation offsets [mm].

G-model			
	gps-nogps	gps-obs	nogps-obs
mean	0.27096	0.71493	0.44397
median	0.00663	0.38610	0.24947

Table 6. Statistics of 12 hour precipitation offsets [mm].

E-model			
	gps-nogps	gps-obs	nogps-obs
mean	0.28493	0.78174	0.49681
median	0.01293	0.33185	0.22654

Table 7. Contingency table for 12 hour precipitation. Biascor = gps-nogps median.

GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	27.09	0.86	0.34	0.06	0.03	95
F1	20.98	3.90	2.30	0.28	0.08	14
F2	12.89	6.56	10.48	2.39	0.67	32
F3	1.86	0.65	2.77	2.14	0.86	26
F4	0.53	0.13	0.48	0.73	0.94	34
FO	43	32	64	38	36	81891

Table 8. Contingency table for 12 hour precipitation. Biascor = gps-nogps mean.

GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	36.58	1.94	0.89	0.12	0.05	92
F1	14.09	3.86	2.69	0.34	0.10	18
F2	10.48	5.64	9.91	2.43	0.67	34
F3	1.71	0.55	2.46	2.04	0.85	27
F4	0.49	0.12	0.42	0.67	0.90	35
FO	58	32	61	36	35	81891

Table 9. Contingency table for 12 hour precipitation. Biascor = gps-obs median.

GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	39.99	2.51	1.13	0.15	0.06	91
F1	11.04	3.48	2.59	0.33	0.10	20
F2	10.16	5.47	9.82	2.44	0.68	34
F3	1.68	0.54	2.42	2.02	0.84	27
F4	0.48	0.12	0.41	0.66	0.89	35
FO	63	29	60	36	35	81891

Table 10. Contingency table for 12 hour precipitation. Biascor = nogps-obs median.

NO GPS, E-model, 200202						
obs→ ↓for	O1	O2	O3	O4	O5	OF
F0	37.64	2.10	1.00	0.13	0.04	92
F1	14.96	4.64	3.53	0.40	0.10	20
F2	8.94	4.93	9.78	2.77	0.77	36
F3	1.43	0.38	1.78	1.83	0.94	29
F4	0.38	0.06	0.27	0.46	0.73	38
FO	59	38	60	33	28	81891

precipitations are given in table 7. Corresponding to the small medians, the statistics is almost not changed compared to the case with bias correction. While it is statistically more correct to correct for the median offset, which represents the typical offset, for completeness we also made a bias correction based on the average offsets. The results of these are given in table 8. One can notice that now the results for precipitation of class 4 and 5 is less good than without the correction, but it is still better than for the 'NO GPS' model predictions.

We conclude that the improved GPS based prediction of heavy precipitation is *not* just due to a general lift of precipitation due to a bias of the ZTD observations.

Both with and without GPS observations included the NWP model simulations are in general too wet, producing too much light precipitation when there should be less or none.

A very simple attempt to correct this in a forecast situation would be applying a bias correction to the model predictions. To see the effects of this, we have bias corrected all the NWP model predictions by subtracting from the predicted precipitations, the medians of the offsets between the NWP models and the observations. The results are presented in the tables 9, and 10.

Comparing table 3 to 9 and 4 to 10, one can notice that the bias corrections lead to a dramatic improvement for observation class 1, some improvement for class 2 and a mild reduction in quality for class 3–5. Comparing table 9 to 10, the forecasts with GPS are still superior to those without for class 4 and 5 after the bias correction of both types of forecasts. Further, one can see that after the bias correction the 12 hour precipitation forecasts with GPS data are now better than those without for class 1 and 3.

Obviously more elaborate correction schemes could be invented depending on forecast class. To explore this is beyond the scope of the present paper.

In Figs. 5 and 6 we plot the probability of detection score ($\mathrm{POD} = hits/(hits + misses)$) and equitable threat score ($\mathrm{ETS} = (hits - randomhits)/(hits + misses + falsealarms - randomhits)$) as function of forecast time. They are made by summing up the forecasts and observations of precipitation at the rain gauge sites according to whether they are above or below a certain limit and interrelating the outcome. Such scores are a means to simplify the information in contingency tables by reducing it to single numbers. The limits used are 5 and 10 mm, corresponding to the highest two of the limits used for the contingency tables. To make a valid comparison between different forecast lengths we have restricted the comparison to the times for which 12 hour precipitation fields are available for the same time at both 12, 18, 24, 30, 36, 42, and 48 h forecast lengths, resulting in about 8 percent fewer forecasts being available per forecast length than for the contingency tables presented above. At 12 hour forecast length the PODs are highest for the forecasts with GPS data, but this superiority is lost already at 18 hours in most cases. The ETS is highest with GPS at 12 hour for the 10 mm limit, but not so for the 5 mm limit. The POD's are highest for the runs with GPS data at the shortest forecast length (12 h), with a degradation relative to the forecasts without at larger forecast times. Unfortunately it has turned out that, for technical reasons, the absolute precision of the precipition stored in the output files of this particular HIRLAM version becomes poorer the larger the difference between the minimum and the maximum precipitation in the field. This may limit the usefulness of Figs. 5 and 6 at large forecast times.

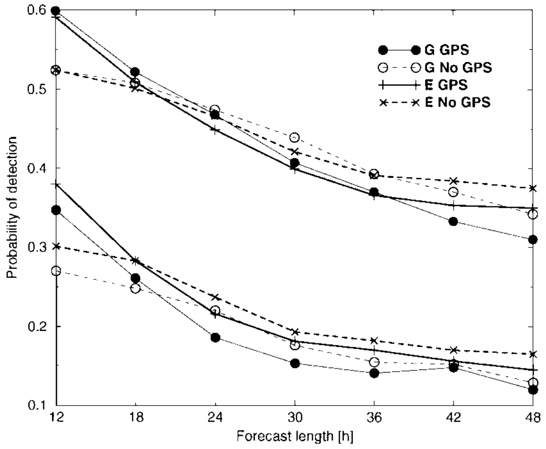


Fig. 5. Probability of detection of 12 hour precipitation versus forecast length. Upper set of curves for 5 mm limit, lower for 10 mm.

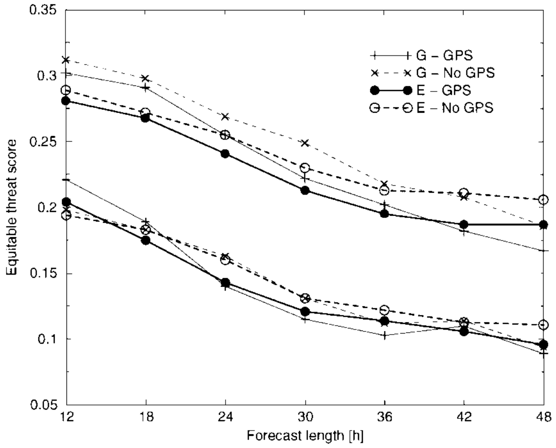


Fig. 6. Equitable threat score of 12 hour precipitation versus forecast length. Upper set of curves for 5 mm limit, lower for 10 mm.

In general, we expect the detailed information about humidity to be lost relatively fast as function of forecast length, and it would be interesting to make also contingency tables and determine scores for shorter forecast lengths. Unfortunately we currently do not have access to the necessary observational data.

4.3 Subjective impact study of precipitation

We have made sets of precipitation maps for the simulations, showing side by side the observed model forecast with and without GPS data. For both G and E model runs, there are two such sets per day corresponding to the 12 hour precipitation measurements/

predictions at 06 and 18 UTC. After a careful inspection of the maps, we find that most often the differences between the forecasts with and without GPS data are quite small. When small differences are found sometimes the forecast with GPS is the better in a certain region and sometimes the forecast without. However, when we find large differences in the prediction of heavy precipitation the forecasts with GPS seem systematically better than the forecasts without. A number of these cases are shown in the Figs. 7 and 8 (The white regions toward south-east on Fig. 8 are caused by the boundary of the E-model).

It is noteworthy that these cases of significant improvement are found over and near areas with good GPS coverage (see Fig. 2) like Germany, the Netherlands, southern England, Switzerland, and the Shetland Islands.

On the other hand, one sees a tendency for the precipitation to spread too far, e.g. to the stations in the North Sea and between the Shetland Islands and Norway. This may be an indication that the background error correlations used in the data assimilation are spatially too wide. The current background error correlations are determined using the NMC method (Parrish and Derber 1992), which is based on the differences between forecasts of different age but valid at the same time. Work is under way to determine background error correlations for HIRVDA by other means.

5. Discussion

The precipitation contingency tables revealed that the use of GPS data improved the prediction of actual strong precipitation. In connection with that an increase in the 'false alarm' rate was also found.

Possibly part of the false alarms can be attributed to the strong local variation of precipitation (and humidity) compared to other meteorological properties modelled by NWP models. In our studies we have seen gauges located next to one another on the precipitation maps (sometimes within one NWP model grid cell) reporting widely different values, e.g. 0 and more than 80 mm for 12 hour precipitation in extreme cases (see maps of observed precipitation in Vedel and Huang 2003). In Figs. 7 and 8 this phenomena is clearly visible (left columns), while it is also clear that the NWP models are far from resolving it (middle and right columns). If, therefore, a NWP forecast is right in predicting strong precipitation in a certain region, it is almost guarantied that the contingency tables will show a number of false alarms in connection with that prediction, reflecting that some gauges in the region received little or no rain.

An indication that such effects provides at least part of the reason for the increase in false alarms is the fact that our by eye inspection of the precipitation maps did not reveal cases where the inclusion of GPS data lead to degradations as clear as the improvements found in the cases with GPS data, like those presented in Figs. 7 and 8. We are attempting to develop a new statistical tool to access this problem.

In comparison with our previous study for June 2000 (Vedel and Huang 2003), the improvement found in the present study when we use GPS ZTD data is more solid. Again we found marked improvements in the subjective verification of precipitation maps, but this time we also found some improvement in the statistical verification against synop and radiosonde observations, and improvements in the contingency tables for strong precipitation (such tables were not made in the previous study). Much more GPS ZTD data were available for Feb. 2002 and the assimilation system has been slightly improved. However, now radiance data are included in the data assimilation, which diminish the impact of other observational data, and the period simulated is less prone to the type of local weather phenomena on which ground based GPS data are expected to have the largest impact.

It would be highly interesting to verify against precipitation and humidity measurements over shorter time scales which is unfortunately not yet possible. Gutman et al. (2003) reports substantial improvements in both NWP humidity and precipitation for the first few hours of forecasts followed by no improvements in the 6–12 hour interval. Further, they report increasing improvements in forecast skill as the size of the GPS network expands, which may be what we see here as well.

Other groups report impact results, which are either positive or neutral overall (see articles in Shoji 2003 and this journal). In gen-

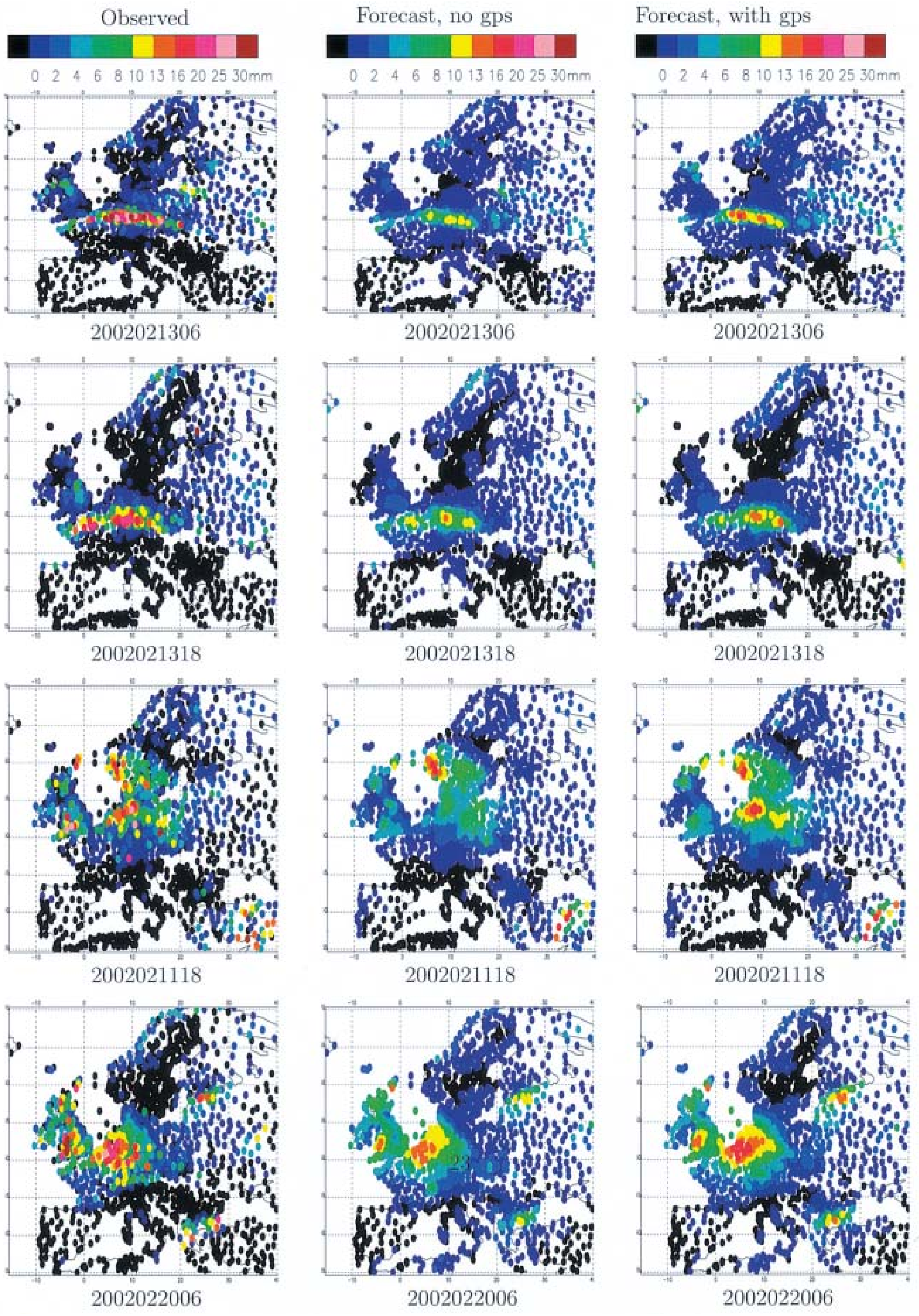


Fig. 7. Comparisons of observed and predicted G-model 12 hour precipitation.

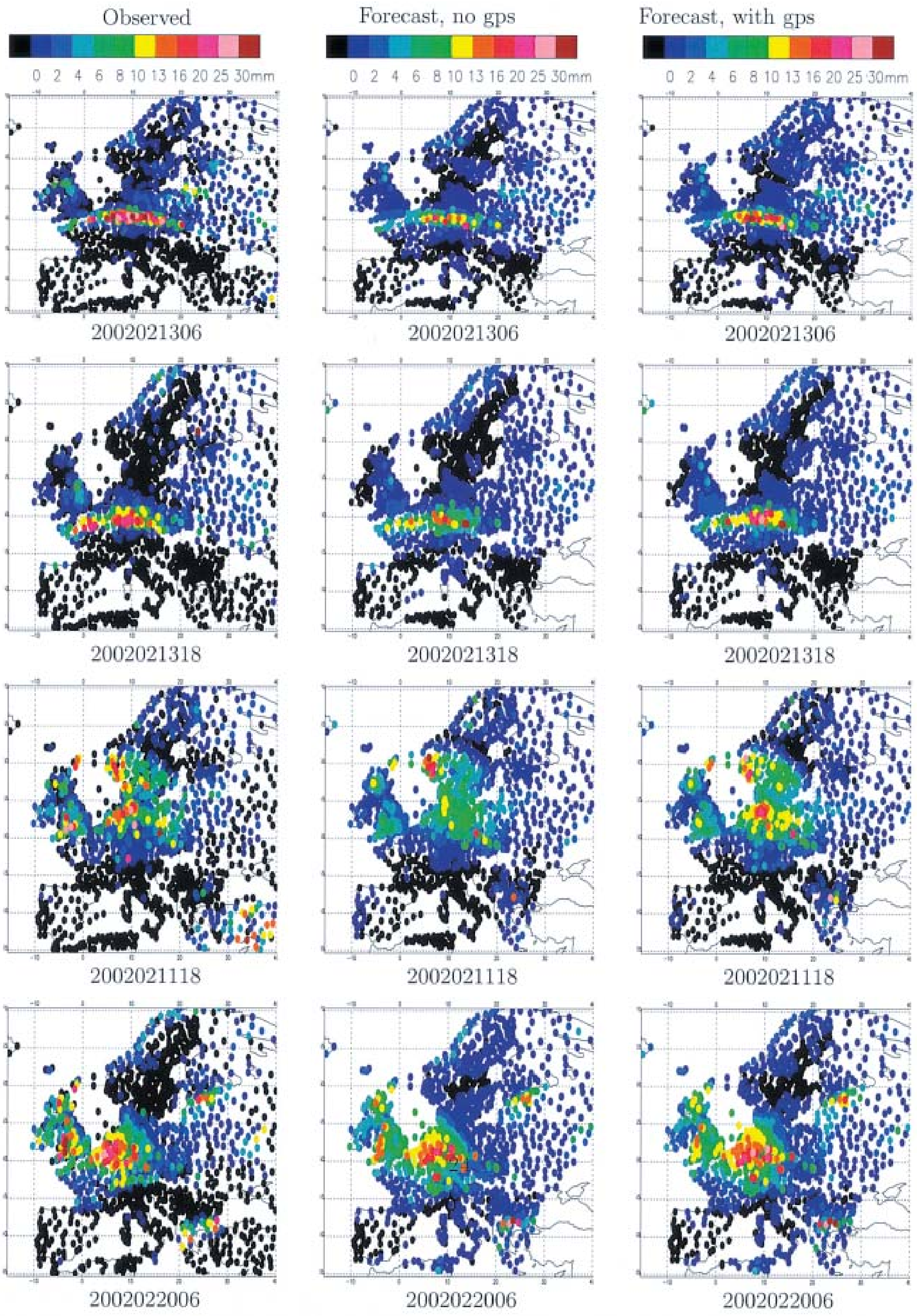


Fig. 8. Comparisons of observed and predicted E-model 12 hour precipitation.

eral, the results are ‘mixed’, meaning that independently of whether the impact is overall positive or neutral seen over a period, there are occasions where the forecasts with GPS data are less good than those without.

6. Conclusion

We have made two sets of parallel simulations, with and without GPS ZTD data in the otherwise standard 3DVar data assimilation, at two different model resolutions (0.45 and 0.15 degrees). The conclusions reported below hold for both resolutions. The period covered is February 2002, during which Europe did not experience unusual weather. We use the spectral version of the NWP model HIRLAM and its data assimilation system HIRVDA. The GPS ZTD data come from the COST Action 716 data sample. Other data assimilated include SYNOP, TEMP, SHIP, BUOY, AIREP, and (A)TOVS data. Statistical verification of geopotential heights, temperatures, winds, and humidities reveals improvements of the geopotential heights when using GPS data, whereas the impact on the other properties is neutral.

Statistical verification of 12 hour precipitation against rain gauge measurements from DMIs database shows improved skill in the prediction of strong precipitation (> 5 mm/12 h) when we use GPS data. The ‘false alarm’ rate also increased when we use GPS data, which could be due to the fact that the real precipitation field varies much more locally than the NWP fields. Thus, in a region receiving strong precipitation, there will often be some rain gauges receiving little or no precipitation, which is not resolved properly by the NWP model.

It is demonstrated that the improvements found in forecasting strong precipitation are *not* due to a common bias of the GPS ZTD data lifting all precipitation predictions when we use GPS data.

An experiment with bias correcting for the median offsets of 12 hour precipitation led to large improvements in the prediction of zero to weak precipitation at the cost of a small reduction of forecast quality of stronger precipitation.

Subjective verification of precipitation maps indicates a marked improvement in the prediction of strong precipitation when we use GPS data. Rarely the forecasts without GPS is better in predicting strong precipitation, and in the cases where large differences were found, the GPS forecasts were the best.

In connection with our previous work and the work by other groups this solidifies the claims that inclusion of ground based GPS data in NWP simulations do improve the forecasts.

More work needs to be done before GPS ZTD data can be utilised optimally in NWP simulations. This includes a better understanding and determination of the errors and error correlations of the GPS ZTD observations, possibly also improved background error covariances for the NWP assimilation system, as well as further impact studies to identify the best assimilation algorithms. A 3 year inter European project called TOUGH (Targeting optimal use of GPS humidity measurements in meteorology) project has just been started for that purpose.

Acknowledgements

The two referees and editor Hiromu Seko are thanked for comments resulting in an improved manuscript. H.V. acknowledges the invitation to take part in the International Workshop on GPS Meteorology, Jan. 14–17, 2003, in Tsukuba, Japan. We are grateful to the organisations providing the GPS data as well as to the institutes processing them into ZTDs for the COST Action 716 database. We thank K.E. Christensen, DMI, for help in extracting humidity data from the DMI database and B. Amstrup for providing the (A)TOVS radiance data. Thanks also to N. Gustafsson, M. Lindskog, and L. Berre (SMHI), who made the original version of the HIRVDA GPS assimilation software, J. Haase (University of Purdue), and K.S. Mogensen (DMI). This work is supported in part by the European Commission, through the 5th Framework Programme, contract no EVG1-CT-2002-00080 in support of the TOUGH project.

References

Bevis, M., T.A. Herring, R.A. Anthes, C. Rocken, and R.H. Ware, 1994: GPS meteorology: Mapping zenith wet delays onto precipitable water, *J. Appl. Meteor.*, **33**, 379–386.

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

Gutman, S.I., K.L. Holub, S.R. Sahm, J.Q. Stewart, T.L. Smith, S.G. Benjamin, and B.E. Schwartz, 2003: Rapid Retrieval and Assimilation of Ground Based GPs-met Observations at the NOAA Forecast Systems Laboratory: Impact on Weather Forecasts, *Proceedings of the International Workshop on GPS Meteorology—GPS Meteorology: Ground-Based and Space-Borne Applications—*.

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

Majewski, D., 1985: Balanced initial and boundary values for a limited area model, *Beitr. Phys. Atmosph.*, **58**, 147–159.

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

Pugnaghi, S., M. Boccolari, R. Pacione, R. Santangelo, H. Vedel, and F. Vespe, 2002: Comparison of independent integrated water vapour estimates from GPS and sun photometer measurements and a meteorological model. *Physics and Chemistry of the Earth*, **27**, 355–362.

Saastamoinen, J., 1972: Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites, *The Use of Artificial Satellites for Geodesy*, **15** 247pp.

Sass, B.H., N.W. Nielsen, J.U. Jørgensen, B. Amstrup, and M. Kmit, 2000: The Operational DMI—HIRLAM system, *Danish Meteorological Institute*, *DMI Technical Report*, 00–26.

Vedel, H., 2000: Conversion of WGS84 geometric heights to NWP model HIRLAM geopotential heights, Danish Meteorological Institute, *DMI scientific report*, 00–04.

——— and X.-Y. Huang, 2003: An NWP impact study with ground based GPS data, *Proceedings of the International Workshop on GPS Meteorology—GPS Meteorology: Ground-Based and Space-Borne Applications—*.

———, K.S. Mogensen, and X.-Y. Huang, 2001: Calculation of delays from meteorological data, comparison of NWP model and observed delays., *Physics and Chemistry of the Earth*, **26**, 497–502.