Q. J. R. Meteorol. Soc. (2005), **131**, pp. 3013–3036 doi: 10.1256/qj.05.27

Validation of precipitable water from ECMWF model analyses with GPS and radiosonde data during the MAP SOP

[1]*Service d'Aéronomie/CNRS, Université Paris VI, France*
[2]*DLR-Institut fuer Physik der Atmosphaere, Oberpfaffenhofen, Germany*
[3]*Laboratoire d'Aérologie/CNRS, Université Paul Sabatier, Toulouse, France*
[4]*LAREG, Institut Géographique National, Marne-la-Vallée, France*

(Received 25 February 2005; revised 14 June 2005)

SUMMARY

Precipitable water vapour contents (PWCs) from European Centre for Medium-Range Weather Forecasts (ECMWF) analyses have been compared with observations from 21 ground-based Global Positioning System receiving stations (GPS) and 14 radiosonde stations (RS), covering central Europe, for the period of the Mesoscale Alpine Programme experiment special observing period (MAP SOP). Two model analyses are considered: one using only conventional data, serving as a control assimilation experiment, and one including additionally most of the non-operational MAP data. Overall, a dry bias of about -1 kg m^{-2} (-5.5% of total PWC), with a standard deviation of $\sim$2.6 kg m^{-2} (13% of total PWC), is diagnosed in both model analyses with respect to GPS. The bias at individual sites is quite variable: from -4 to $\sim$0 kg m^{-2}. The largest differences are observed at stations located in mountainous areas and/or near the sea, which reveal differences in representativeness. Differences between the two model analyses, and between these analyses and GPS, are investigated in terms of usage and quality of RS data. Biases in RS data are found from comparisons with both model and GPS PWCs. They are confirmed from analysis feedback statistics available at ECMWF. An overall dry bias in RS PWC of 4.5% is found, compared to GPS. The detection of RS biases from comparisons both with the model and GPS indicates that data screening during assimilation was generally effective. However, some RS bias went into the model analyses. Inspection of the time evolution of PWC from the model analyses and GPS occasionally showed differences of up to 5–10 kg m^{-2}. These were associated with severe weather events, with variations in the amount of RS data being assimilated, and with time lags in the PWCs from the two model analyses. Such large differences contribute strongly to the overall observed standard deviations. Good confidence in GPS PWC estimates is gained through this work, even during periods of heavy rain. These results support the future assimilation of GPS data, both for operational weather prediction and for mesoscale simulation studies.

KEYWORDS: Data assimilation Mesoscale Alpine Programme

1. INTRODUCTION

The MAP SOP (Mesoscale Alpine Programme Special Observing Period, Bougeault *et al.* 2001) took place between 7 September and 15 November 1999 in central Europe (mainly centred on the Alps). One of main goals of this experiment was to provide a database suitable for the evaluation and improvement of high-resolution numerical models, especially for the prediction of heavy precipitation in complex topography. The experiment lasted 70 days, during which 17 intensive observing periods (IOPs) were conducted, among which 15 heavy to moderate precipitation events were documented. The experiment produced a huge dataset from special observing systems comprising: a high-density surface rain-gauge network; increased radiosonde network; ground-based fixed and mobile radars; wind profilers; and airborne radars. These data have been used to investigate orographic precipitation mechanisms and to validate numerical weather prediction (NWP) model simulations. They are now intended to be used for mesoscale assimilation experiments. A number of such results were published in the MAP special issue of the Quarterly Journal in January 2003 (*Q. J. R. Meteorol. Soc.*, **129,** 341–895) and elsewhere.

* Corresponding author: Service d'Aéronomie/CNRS, Université Paris VI, 4, place Jussieu, 75252 Paris cedex, France. e-mail: Olivier.Bock@aero.jussieu.fr

A good description of the initial conditions, and especially the three-dimensional (3D) water vapour field, is crucial for the simulation of convective systems and situations leading to heavy precipitation events (Koch *et al.* 1997; Romero *et al.* 1998; Ducrocq *et al.* 2002; Faccani *et al.* 2003). Mesoscale model simulations of different MAP events were performed using the European Centre for Medium-Range Weather Forecasts (ECMWF) operational analyses of 1999 and, more recently, the re-analysis of the MAP SOP data performed at ECMWF in 2002, referred to hereafter as MAPRA. Few experiments have been conducted so far with MAPRA. Buzzi *et al.* (2004) report one case where MAPRA clearly has a positive impact on high-resolution simulations (based on precipitation forecast scores), while in general they state that evidence is insufficient to conclude whether MAPRA or the operational analysis leads to better simulations. Richard *et al.* (2003) and Lascaux *et al.* (2004) performed several mesoscale simulation experiments with both MAPRA and the operational analysis, and tested a number of physical parametrizations for the IOP2a event (a squall line associated with moderate rainfall passing over the Lago Maggiore area). They show that in this special case MAPRA fields do not allow convection to develop, while the operational analysis leads to much better results in the simulations. A comparison of both analyses revealed significant differences in the low-level moisture fields (MAPRA being much drier than the operational analysis). Reasons for this difference are the higher resolution of topography leading to stronger flow-blocking in MAPRA, and the assimilation of special observational data (Keil and Cardinali 2004).

The present paper focuses on the validation of MAPRA and a control analysis experiment (referred to as CTRL) which did not include the special MAP observations. Both were performed with the operational model and 4D-Variational (4D-Var) assimilation system of ECMWF as operated in 2002. The operational analysis from 1999 has not been considered here, since numerous changes were implemented in the model between 1999 and 2002 and the analysis of the impact of these changes is beyond the scope of this paper. Precipitable water vapour content (PWC) estimates from a ground-based network of Global Positioning System (GPS) receivers are used to conduct a verification of the quality of these two analyses for the whole MAP SOP period. Although GPS data, as opposed to radiosonde (RS) data, only provide the integrated amount of humidity, they are fully model-independent observations since they are not assimilated. GPS as a humidity observing technique is quite recent, but it has proven to be as accurate as conventional techniques (Bevis *et al.* 1992; Rocken *et al.* 1995; Emardson *et al.* 1998; Klein Baltink *et al.* 2002) and has already been used for the validation of numerical models (Yang *et al.* 1999; Köpken 2001; Hagemann *et al.* 2003).

This paper is organized as follows: in section 2 the dataset used for this study and the methodology for the PWC comparisons are described, and error sources associated with the data are discussed. In section 3, statistical results of PWC comparisons (model versus GPS and RS) are presented. In section 4, the overall time evolution of model and GPS PWCs and their differences is first analysed; this is followed by a focus on a few target areas with heavy precipitation. Conclusions and perspectives are drawn in section 5.

2. Data, methodology and error sources

(*a*) *GPS data*

For the present work, zenith tropospheric delay (ZTD) estimates from 20 permanent European GPS stations and one special station (referred to as MILA) deployed in Milan, Italy, are used (see Fig. 1 and Table 1). ZTD estimates are generally

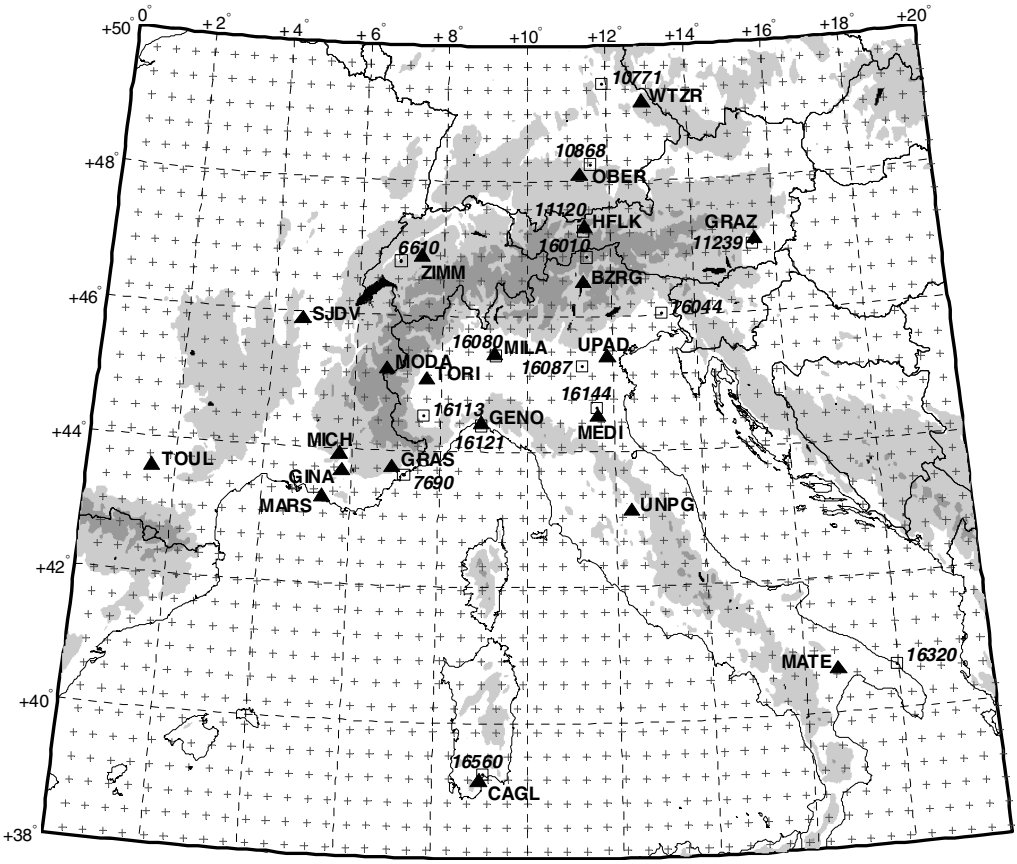

Figure 1. Domain of study showing high-resolution topography as dark grey where greater than 1500 m, and light grey from 500 to 1500 m above sea level. Superimposed is the European Centre for Medium-Range Weather Forecasts reduced Gaussian grid as grey pluses (256 latitude lines between 0 and 90°N, with quasi-regular resolution of 40 km). Global Positioning System stations are given as black triangles with four-letter short names, and radiosonde sites as open squares with WMO station numbers in italics.

produced through the processing of GPS data using geodetic analysis software. The ZTD estimates for the 20 permanent stations were retrieved from the archive of the MAGIC project (Meteorological Applications of GPS Integrated Column water vapour measurements in the Western Mediterranean, see www.acri.fr/magic). These ZTDs were processed using the GAMIT (GPS at Massachusetts Institute of Technology) software as described in Haase *et al.* (2003). For MILA we processed the data using the GPS software of the University of Berne, Switzerland (referred to as Bernese software) as described in Bock *et al.* (2001). Bernese software, like GAMIT, performs a network analysis. For validation purposes, stations common with the MAGIC network were thus included in the Bernese analysis. The results obtained from the two versions of GPS software showed a root mean square (RMS) difference in ZTD of ∼5 mm, which is accounted for by slightly different analysis procedures. This uncertainty in ZTD converts into ∼0.8 kg m^{-2} of PWC, which is a typical level of uncertainty associated with PWC retrieval from GPS data (Emardson *et al.* 1998; Fang *et al.* 1998).

ZTD is an integral of the refractivity of the air at the frequency of the GPS signal transmitted from the GPS satellites. The refractivity itself depends on the concentration of water vapour molecules. Hence, ZTD can be related quite easily to PWC.

TABLE 1. LOCATION OF GPS STATIONS, RS SITES AND THE NEAREST GRIDPOINT FROM THE ECMWF MODEL IN 2002

GPS station				RS site				ECMWF model		
Name	Lat. (°)	Lon. (°)	Altitude (m)	Name	WMO station number	Altitude (m)	Distance to GPS (km)	Gridpoint altitude (m)	Distance (km) to GPS	to RS
BZRG	46.5	11.3	280	Vipiteno/Sterzing	16010[1]	942	44	1375	6	39
CAGL	39.1	9.0	192	CAGLIARI/ELMAS	16560[1]	18	14	98	16	14
GENO	44.4	8.9	111	Genova	16121[1]	10	2	316	3	4
GINA	43.7	5.8	331					455	14	
GRAS	43.8	6.9	1269	Nice	07690[1]	4	25	527	13	16
GRAZ	47.1	15.5	491	GRAZ/THALERHOF	11239	340	9	657	15	27
HFLK	47.3	11.4	2336	INNSBRUCK	11120	581	5	1609	10	12
MARS	43.3	5.4	12					73	14	
MATE	40.6	16.7	490	BRINDISI	16320	10	105	195	6	110
MEDI	44.5	11.7	10	BOLOGNA	16144[1]	11	15	68	12	26
MICH	43.9	5.7	577					732	20	
MILA	45.5	9.3	116	MILANO/LINATE	16080[1]	103	4	147	17	14
MODA	45.2	6.7	1129					2191	8	
OBER	48.1	11.3	596	MUENCHEN	10868	84	27	652	23	35
SJDV	45.9	4.7	383					241	11	
TORI	45.1	7.7	263	Cuneo	16113	382	59	379	4	67
TOUL	43.6	1.5	159					189	16	
UNPG	43.1	12.4	303					470	18	
UPAD	45.4	11.9	39	Verona	16087[1]	67	49	−52	12	61
WTZR	49.1	12.9	619	KUEMMERSBRUCK	10771	418	78	623	14	93
ZIMM	46.9	7.5	908	PAYERNE	06610[1]	490	40	799	7	36

Altitudes are m above mean sea level. Latitude/longitude coordinates are only reported for GPS. Uppercase names of radiosondes indicate operational stations. See text for definitions and further details.
[1]High-resolution data are used for these RS stations.

The conversion of GPS ZTD estimates into PWC (i.e. PWC_{GPS}) is performed in two steps (see, e.g. Bevis *et al.* 1994):

- Formation of zenith wet delay: $ZWD = ZTD - ZHD$, where ZHD is the zenith hydrostatic delay and is estimated from the surface pressure, P_{surf}, at the height of the GPS receiver: $ZHD = 2.279\,P_{\mathrm{surf}}$.
- Conversion of ZWD into PWC using a conversion factor $\kappa(T_{\mathrm{m}})$, depending on the mean temperature, T_{m}, in the column of atmosphere above the GPS antenna: $PWC_{\mathrm{GPS}} = \kappa(T_{\mathrm{m}})\,(ZTD - ZHD)$. An evaluation from radiosonde data at Milan, during the MAP SOP, yielded values for $\kappa(T_{\mathrm{m}})$ in the range 152–163 kg m^{-3}.

Using the definitions for ZHD, ZWD, and $\kappa(T_{\mathrm{m}})$, as given, for example, in Bevis *et al.* (1994), PWC_{GPS} can be expressed as:

$$PWC_{\mathrm{GPS}} = \frac{1}{k_3/T_{\mathrm{m}} + k_2'} \int_{z_{\mathrm{surf}}}^{\infty} \left(k_3 \frac{1}{T(z)} + k_2' \right) \rho_{\mathrm{v}}(z)\, \mathrm{d}z, \tag{1}$$

where k_2' and k_3 are refractivity constants at the GPS carrier frequencies (Bevis *et al.* 1992), $\rho_{\mathrm{v}}(z)$ and $T(z)$ are water vapour density and temperature, respectively, and z is the altitude within the atmosphere. Equation (1) shows that the PWC_{GPS} estimate is a weighted integral of water vapour density in the column of atmosphere above the height of the GPS antenna, z_{surf}. The impact of the weighting is generally very weak and can be made negligible using a proper model for T_{m}. A number of models have been proposed for T_{m}, as a function of atmospheric temperature close to the surface, T_{surf}. In the present work the linear model of Bevis *et al.* (1994) is used.

This model writes $T_m = aT_{surf} + b$, with $a = 0.72$ and $b = 72.0$ K. Though coefficients a and b are known to be season and latitude dependent, their variability over the domain and period of interest is quite small. An evaluation at Milan (Cagliari), based on high-resolution RS data, yielded values of $a = 0.74$ (0.74) and $b = 64.0$ K (64.4 K), and a RMS difference compared with the model of Bevis *et al.* (1994) of 2.6 K (2.8 K). When converted into PWC, this uncertainty represents only 0.18 kg m^{-2} (0.22 kg m^{-2}). However, since T_m varies with T_{surf}, it is important to use measurements of T_{surf} close to the GPS station in order to follow the large diurnal variations in temperature close to the surface over the period (e.g. 5 degC $< T_{surf} <$ 30 degC at Milan during the MAP SOP).

The conversion of GPS ZTD estimates into PWC requires two additional observations: surface pressure and temperature. The impact of errors in these variables on PWC can be assessed from the partial derivatives, computed from the definition of PWC and the formulae given above: $\partial PWC/\partial P_{surf} = 0.35$ kg m^{-2} hPa^{-1} and $\partial PWC/\partial T_{surf} =$ 0.05 kg m^{-2}K^{-1}. At most of the GPS stations operated in 1999, surface pressure and temperature were not measured. Hence, values from the nearest meteorological surface stations (e.g. from the WMO network) were retrieved from the MAP database (http://www.map.ethz.ch/). These data were corrected for differences in altitude from the GPS stations (<200 m for most stations) using hydrostatic equilibrium and a temperature lapse rate of -6.5 degC km^{-1}. The uncertainty associated with a 200 m extrapolation has been evaluated from 2 months of radiosonde data from Milan. Differences of $\sim$0.4 hPa and $\sim$2 degC RMS are found (converting into 0.24 kg m^{-2} in PWC). Some authors use surface pressure and temperature from NWP models instead of observations. A comparison of model values from MAPRA and observations for surface pressure and temperature in Milan showed differences of $\sim$0.8 hPa and $\sim$2 degC RMS (converting into 0.45 kg m^{-2} in PWC). Similar results were obtained from CTRL. Since observations have smaller uncertainties than model surface values, the former are used in this work. Hagemann *et al.* (2003) have reported similar results uncertainties in model surface values.

The combination of the different errors from sources estimated above yields an overall uncertainty in PWC_{GPS} smaller than 1 kg m^{-2} RMS. This uncertainty is dominated by errors in ZTD estimates rather than errors in the conversion from ZTD into PWC. Comparison with independent observing techniques such as RS and microwave radiometers has been shown to yield an agreement on the level of 1–2 kg m^{-2} RMS (Rocken *et al.* 1995; Emardson *et al.* 1998; Köpken 2001; Klein Baltink *et al.* 2002).

(*b*) *Radiosondes*

RS data (see Fig. 1 and Table 1 for details) are used for a comparison with both GPS and ECMWF analyses. Quality-checked RS data were retrieved from the MAP database (http://www.map.ethz.ch/). At some of these stations (see Table 1) data are available at a high vertical resolution ($\sim$50 m). However, none of the available data is corrected for biases in humidity measurement. Vaisala RS80 sensors, which were extensively used in Europe at the time of the MAP SOP, are known to exhibit dry biases (Wang *et al.* 2002). According to Häberli (2003, 2005) these bias corrections amount on the average to $\sim$5–7%, depending on the site and correction scheme. This uncertainty converts into an absolute bias in PWC on the order of 1 kg m^{-2} (5% of an average PWC of 20 kg m^{-2}, typical over Europe during the SOP). Larger values are estimated at some sites, among which is Cagliari with 13% uncertainty, i.e. 2.7 kg m^{-2} PWC bias. This dry bias is thus not only a major error source associated with the RS data used here for comparison purposes but also during data assimilation.

For the purpose of the comparison of RS estimates of PWC, PWC_{rs}, with PWC_{GPS} or model analyses PWC, PWC_{mod}, PWC_{rs} is computed in the following way:

$$PWC_{\mathrm{rs}} = \int_{z_0}^{z_1} \rho_{\mathrm{v}}(z)\,\mathrm{d}z, \tag{2}$$

where $\rho_{\mathrm{v}}(z)$ is the water vapour density measured by the RS as a function of altitude between the surface altitude z_0, corresponding either to the height of a GPS station or that of the nearest model grid point, and z_1, the highest altitude where humidity data are reported by the RS. In the case when $z_1 < 10$ km, a correction term for the missing part above the profile is added (based on climatology). When $z_1 = 5$ km, this correction is accurate to $\sim$0.3 kg m^{-2} RMS (based on 2 months of RS data from Milan).

Depending on the site, it can happen that the altitude of GPS station or the model's grid point is below the altitude of the lowest RS data. In this case RS data are extrapolated downwards assuming a constant relative humidity, hydrostatic equilibrium and a -6.5 degC km^{-1} temperature lapse rate. The error in the extrapolated surface specific humidity is estimated to be $\sim$1 g kg^{-1} RMS for a 200 m height difference, producing a small error in PWC of 0.2 kg m^{-2} RMS (based on 2 months of radiosonde data from Milan).

Finally, the difference in vertical resolution between radiosonde profiles and the model might contribute to differences in PWC. A comparison of low- and high-resolution profiles at Milan ($\sim$40 and $\sim$600 levels per ascent, respectively) yields a difference in PWC of $\sim$0.4 kg m^{-2} RMS. This gives an idea of the level of accuracy expected at best between PWC_{mod} and PWC_{rs}.

(*c*) *ECMWF analyses*

In this paper two different analyses are compared: MAPRA, the final reanalysis, and CTRL, a control analysis experiment. Both analyses were obtained from ECMWF and used the assimilation system available in 2002; thus they benefited from substantial upgrades in the global data assimilation system since the MAP field phase (Rabier *et al.* 2000). The horizontal resolution of both analyses amounts to 40 km on 60 vertical levels (T511L60, see the representation of this grid in Fig. 1). MAPRA and CTRL differ only in data usage; extra MAP observations were assimilated in MAPRA but not in CTRL. The additional MAP observations assimilated in MAPRA comprise data from: surface stations; high–resolution radiosoundings; European wind profiler; dropsondes; and research aircraft. Both the improved model resolution and the use of special data in MAPRA have been shown to produce slightly moister conditions in the southern Alpine region, southern France and over the Adriatic Sea, and drier conditions on the Alpine crest and the islands of Corsica and Sardinia (Keil and Cardinali 2004).

In both analyses, the following humidity data are assimilated:

- RS: relative humidity up to 300hPa (with observation error standard deviation of 17% in the lower troposphere);
- SYNOP surface stations: relative humidity when the station is within 4 hPa of the orography in the model (with observation error of 13%);
- PWC from Special Sensor Microwave Imager (SSM/I) instrument over the oceans only.

Hence, RS are the major source of conventional tropospheric humidity data assimilated over land.

The verification of model fields with observational data is a difficult task mainly for two reasons. The first is that it needs a proper handling of model variables and/or

observational data in order to ensure that the variables that are compared are consistent in space and time. Especially, the comparison of model fields to sparse observational data is limited by factors such as: differences in altitude between the real orography and its representation in the model (especially for integrated variables such as PWC), model representativeness, and sampling errors (especially when only a few observational sites are used). The second reason is that observational data can contain outliers and observations that might be inconsistent with the model's variables (e.g. due to improper physical balance imposed in the model). Thus, observational data need to be correctly quality checked. During the assimilation process, the ECMWF data assimilation system provides much information on background departures (differences between the model forecast and observations) and analysis departures (difference between the final analysis and observations). These data are stored in feedback files. Feedback files for both MAPRA and CTRL assimilation experiments are used in subsection 3(b).

PWC from the model humidity field is integrated over model pressure levels, P, between the top of model, $P_{\text{top}} = 0.1$ hPa, and the surface pressure, P_{surf}:

$$PWC_{\text{mod}} = \frac{1}{g_0} \int_{P_{\text{top}}}^{P_{\text{surf}}} q(P)\, \mathrm{d}P, \tag{3}$$

where $g_0 = 9.80665$ m s^{-2} is the standard acceleration due to gravity at mean sea level, and $q(P)$ is the specific humidity of air as a function of atmospheric pressure.

In order to minimize comparison errors due to height differences between GPS and model topography, the model surface variables are interpolated or extrapolated from the nearest model levels to the height of the GPS station and integration is actually performed above this height. The procedure is similar to that used for the RS data (cf. subsection 2(b)). In a previous work, model PWC fields were not corrected for the altitude difference, and biases in PWC showed up between GPS observations and model (Bock *et al.* 2004). The relative PWC bias had a linear relationship with an altitude difference of about -40% PWC (1000m)$^{-1}$.

Model representativeness is best preserved when using data on the original 'reduced Gaussian grid' (Hortal and Simmons 1991), which is used during assimilation and forecast. Although some authors have employed horizontally interpolated data, using surrounding grid points (e.g. Köpken 2001), only data from the nearest grid-point (in horizontal coordinates) are used in the present work (as recommended at ECMWF for verification with sparse observational data). To give an idea of the impact of such an interpolation, we have compared PWC fields extracted on a regular latitude/longitude grid to those of the nearest reduced Gaussian grid points. Differences as large as ± 3 kg m^{-2} were observed in regions of steep orography.

3. Comparison of model, GPS and RS PWCs

(*a*) *ECMWF analyses compared to GPS*

In this section, PWCs from 3-hourly model analyses are compared to PWCs from GPS observations. Model fields were extracted at the nearest grid-point to the GPS sites, and each PWC was corrected for the difference in altitude (subsection 2(c)).

For the comparison, model PWC data are paired with the nearest in time GPS PWC data. Tests made with GPS ZTD averaged over 1 to 3 h intervals led to similar results ($<2\%$ difference in bias and standard deviation); hence GPS data were not averaged for the subsequent comparisons. The number of GPS–model data pairs is unevenly distributed in time, due to gaps in the GPS data series. These gaps vary from one station

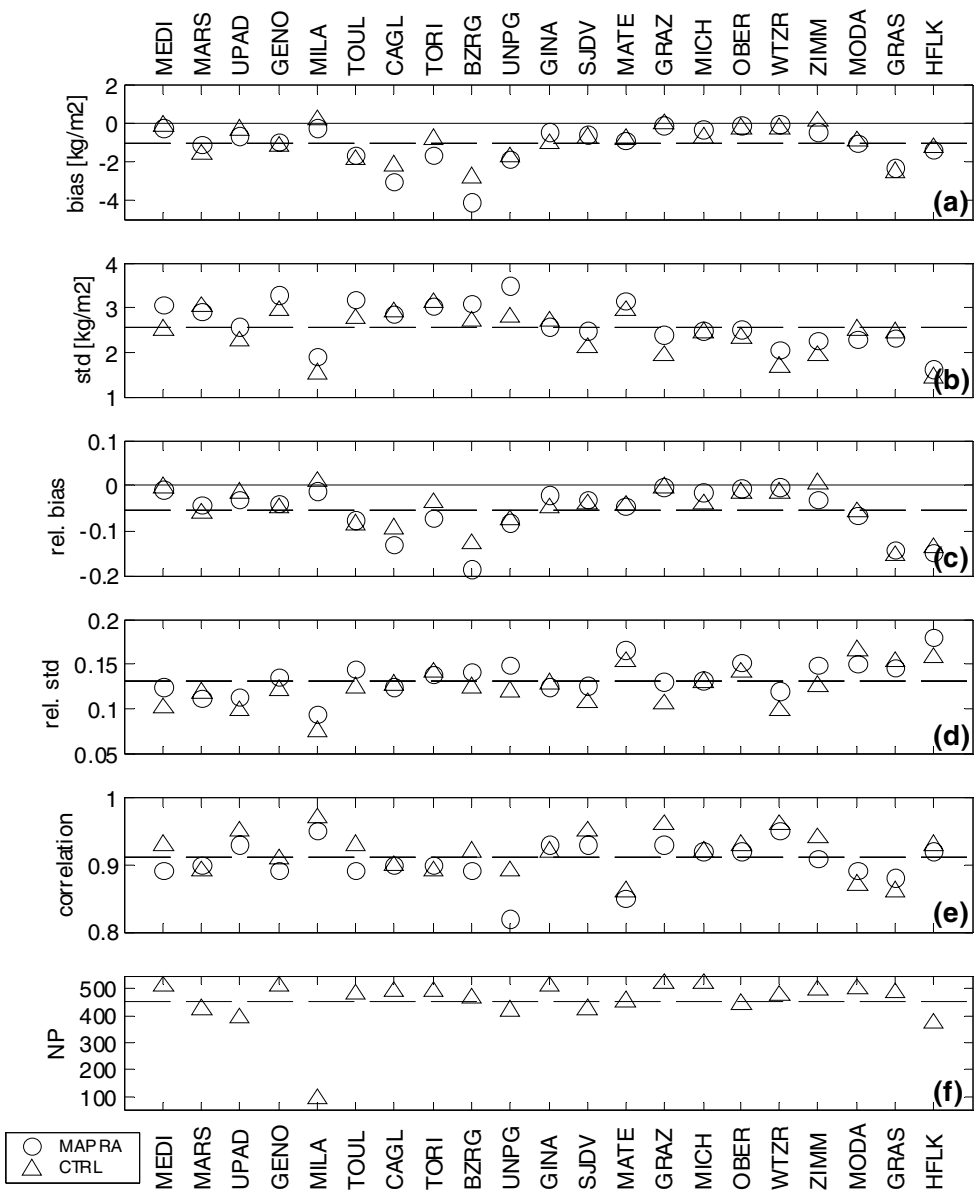

Figure 2. Comparisons of 3 h precipitable water vapour contents (PWCs) from the models MAPRA (circles) and CTRL (triangles) with PWCs based on 21 Global Positioning System (GPS) stations: (a) bias (kg m^{-2}), (b) standard deviation (kg m^{-2}), (c) relative bias (model–GPS)/GPS, (d) relative standard deviation, (e) correlation coefficient, (f) number of data pairs. These parameters are statistics over 70 days (the whole MAP SOP) and are ordered in increasing GPS site altitude from left to right. Dashed lines indicate averages over all stations. See text for further details, and Fig. 1 and Table 1 for station locations.

to another; on average, there are 450 data pairs available per site. In Milan (MILA) only 92 data pairs are available since this station was operated for 23 days only.

Figure 2 shows statistics of the differences between model and GPS PWCs for all 21 stations. It shows both absolute and relative PWC bias and standard deviation between model and GPS, as well as the correlation coefficient between both PWC time series. The statistics are calculated over the 70 days of the MAP SOP and results are ordered in

TABLE 2. MODEL PWC ANALYSES COMPARED TO GPS PWCS FOR ALL 21 GPS STATIONS AND FOR FOUR SUBSETS

Experiment	Average GPS PWC (kg m^{-2})	Mean bias (kg m^{-2})	Standard deviation of bias (kg m^{-2})	STD (kg m^{-2})	Relative bias	Relative STD	Correlation
All 21 stations							
MAPRA	19.9	−1.10	1.05	2.65	−0.057	0.14	0.90
CTRL	19.9	−1.00	0.86	2.43	−0.052	0.12	0.92
15 stations with difference in height with model <200 m							
MAPRA	20.7	−0.83	0.86	2.66	−0.037	0.13	0.91
CTRL	20.7	−0.76	0.77	2.41	−0.034	0.12	0.93
11 stations in the mountains BZRG, GRAS, GRAZ, HFLK, MICH, MODA, OBER, TORI, UNPG, WTZR, ZIMM							
MAPRA	17.6	−1.21	1.24	2.51	−0.069	0.14	0.90
CTRL	17.6	−1.03	0.98	2.31	−0.060	0.13	0.91
7 stations near the sea CAGL, GENO, GRAS, MARS, MATE, MEDI, UPAD							
MAPRA	22.3	−1.30	0.98	2.88	−0.063	0.13	0.89
CTRL	22.3	−1.24	0.91	2.72	−0.060	0.12	0.90
5 stations in the southern Alps GENO, MEDI, MILA, TORI, UPAD							
MAPRA	22.9	−0.75	0.60	2.78	−0.033	0.12	0.91
CTRL	22.9	−0.44	0.56	2.48	−0.019	0.11	0.93

The third and fourth columns indicate mean and standard deviation of bias computed over the GPS sites. The following columns indicate average values of standard deviation (STD), relative bias and relative STD of the difference between model and GPS PWCs. The last column gives the average correlation between model and GPS PWCs. See text and Table 1 for definitions and station locations.

increasing altitude of the GPS sites. The figure shows generally varying results from one site to another. However, there is an overall slight tendency for the bias (Fig. 2(a)), the relative bias (Fig. 2(c)), and the standard deviation (Fig. 2(b)) to decrease with altitude, whilst the relative standard deviation (Fig. 2(d)) increases. The correlation coefficient (Fig. 2(e)) varies from one site to another, but indicates a generally good degree of correlation at all sites (around 90%). At most sites the absolute bias lies between 0 and -2 kg m^{-2} and the relative bias is smaller than -10%. The accuracy of the bias estimate (defined at a confidence level of 68%) at a single station, is approximately $\sigma/\sqrt{N}$, where σ is the standard deviation of the differences in PWC and N is the number of data pairs. Hence, using average values for $\sigma \sim 2.6$ kg m^{-2} and $N \sim 450$ (see Table 2), an error bar of $\sim$0.12 kg m^{-2} can be put on the bias estimates. Biases (in absolute value) larger than 0.12 kg m^{-2} can, therefore, be considered as statistically significant (with a confidence level of 68%). This is the case at a number of stations. The difference in the distribution of standard deviation (Fig. 2(b)) and relative standard deviation (Fig. 2(d)) is thought to be due to the atmospheric variability that is contained in the GPS PWC, while it is more smoothed in model PWC. As the absolute PWC in the atmosphere increases, the standard deviation in PWC is also expected to increase due to the proportional increase in atmospheric variability. However, GPS PWC error is not expected to depend on the absolute PWC (Hagemann *et al.* 2003).

Figure 3 shows the spatial distribution of the relative bias and standard deviation over the central region of interest (containing 17 out of 21 GPS stations). The largest relative biases and relative standard deviations are observed at stations located in mountainous regions (such as BZRG, GRAS and HFLK) and/or near the sea (such as GRAS and CAGL, not shown). The sites in the mountains suffer from a higher uncertainty in

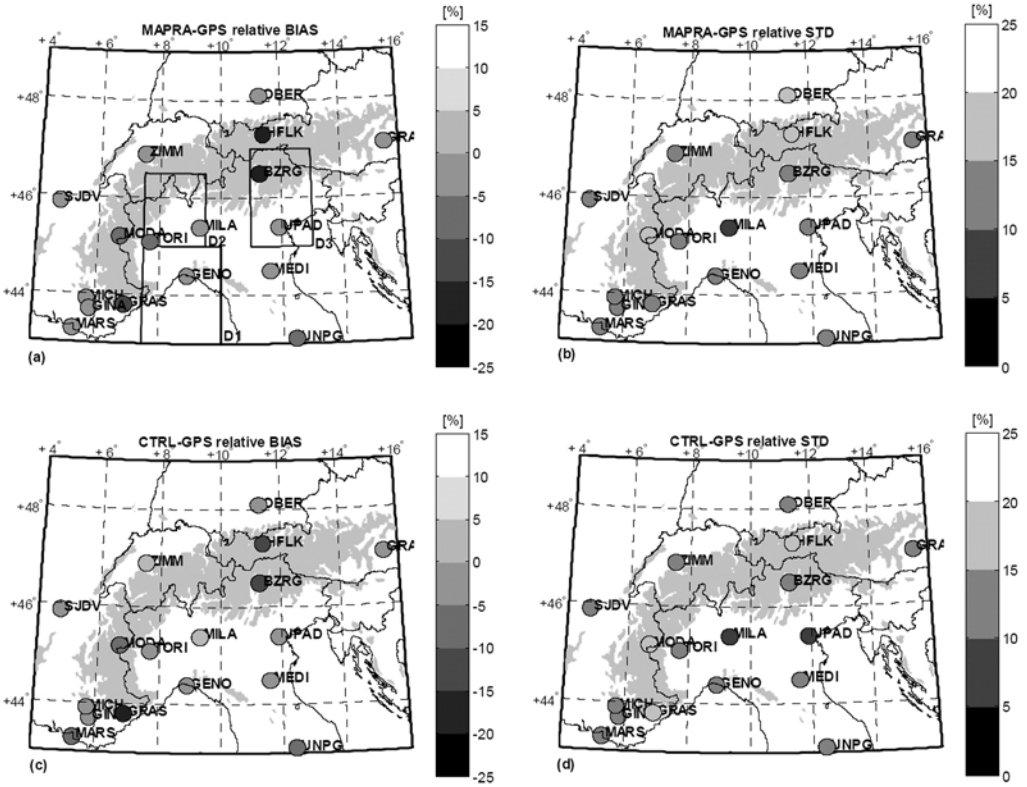

Figure 3. Spatial distributions of 3 h precipitable water vapour content (PWC) differences from the two models MAPRA and CTRL with PWCs based on Global Positioning System (GPS) stations: (a) MAPRA minus GPS PWC relative bias; (b) as (a) but for relative standard deviation; (c) and (d) as (a) and (b), respectively, but for CTRL. Data are similar to those in Figs. 2(c) and (d) but for only 17 out of 21 GPS sites. Grey background shading represents topography above the 1000 m contour. See text for further details. (Domains D1, D2, and D3 in (a) are used for the computation of area-averaged precipitation plotted in Fig. 8.)

PWC correction, due to large differences in altitude between the real orography and its representation in the model (subsection 2(c)), and from representativeness differences (with the model and GPS not representing exactly the same volume of atmosphere). The sites in the south of the Alps show, comparatively, good agreement between model and GPS.

Table 2 gives average statistics, computed over the 21 stations and four subsets of stations. Over the 21 sites, both analyses show an average bias of nearly -1 kg m^{-2} (relative bias of -5.5%), i.e. both analyses are too dry. The standard deviation is about $\sim$2.6 kg m^{-2} (relative standard deviation of $\sim$13%) and the correlation is $\sim$0.90. Differences between the model analyses are not significant at this stage, mainly because of the under-sampling induced by point comparisons at such a small number of sites. The accuracy of the average bias is now $\sigma_{\text{bias}}/(N_{\text{sta}})^{1/2}$, where σ_{bias} is the standard deviation in bias and N_{sta} the number of stations. With $\sigma_{\text{bias}} = 1.0$ kg m^{-2} and $N_{\text{sta}} = 21$, $\sigma_{\text{bias}}/(N_{\text{sta}})^{1/2} = 0.22$ kg m^{-2}. Thus any change in bias smaller than 0.22 kg m^{-2} cannot be considered as being significant. Table 2 (second subset) shows that much better agreement is found from stations with altitudes lying within 200 m of that of the nearest model grid point. This selection actually removes three of the stations where the relative bias was very large (BZRG, GRAS and HFLK, Fig. 2(c)).

Statistics for stations in the mountains (Table 2, third subset) show mainly an increase in relative bias (−6.9% for MAPRA) with respect to (w.r.t.) the average over all stations (Table 2, first subset). Stations near the sea (Table 2, fourth subset) have a larger absolute bias (−1.30 kg m^{-2} for MAPRA). Statistics over five stations in the southern Alps (Table 2, last subset) show a significant reduction in the bias (absolute, relative and spatial variability) w.r.t. the first subset of Table 2; better representativeness of model and GPS observations are likely in this region. For this small subset of stations, CTRL shows a significantly smaller bias than MAPRA.

Differences between MAPRA and CTRL, as evidenced in the southern Alps (Table 2) and at a number of other stations (Figs. 2 and 3), reveal a slightly negative impact of special observations assimilated in MAPRA. Differences in bias (Fig. 2(a)) and relative bias (Figs. 2(c), 3(a) and (c)) are observed at a number of sites (BZRG, CAGL, GINA, MARS, MILA, TORI, and UPAD), at most of which the bias in MAPRA is larger. Higher standard deviation (Fig. 2(b)) and lower correlation (Fig. 2(e)) for MAPRA are also observed at most sites, with a 20% increase in standard deviation observed at GRAZ, MEDI, MILA, UNPG, and WTRZ, and a 10% increase at additional sites: BZRG, GENO, HFLK, SJDV, TOUL, UPAD, and ZIMM. However, an improvement is also seen in MAPRA at some sites, with a slight reduction in standard deviation (GINA, MARS, and MODA). The main reason for these differences between MAPRA and CTRL are suspected to come from biases in RS data. The impact of wind profiler data was also evidenced during IOP2a (Keil and Cardinali 2004), but this might not be a systematic source of bias. Consequently the possibility of biases in RS data (from both operational and special MAP soundings) is investigated through a comparison of model and GPS with RS PWC.

(*b*) *Model analyses and GPS compared to RS*

In this subsection, RS data archived at full resolution on the MAP database are used. The comparison is performed only for those RS stations close to a GPS station, resulting in a set of 14 RS stations (see Fig. 1). Among these, only eight produced high-resolution data (see Table 1); five of these stations were operated only during the MAP SOP (Vipiteno/Sterzing, Genova, Cuneo, Verona, and Nice). For the comparison with GPS and model analyses, PWC is calculated from RS profiles for two different surface altitudes: either that of the GPS site or that of the nearest model grid point. Before use, RS data are checked for consistency with standard temperature and pressure profiles; only profiles with humidity data reported up to at least 5 km altitude are retained. Most stations in the southern Alps provided 6-hourly reports.

Average statistics for the difference between model and GPS PWC are presented in Table 3, for all stations and for subsets of stations. They are evaluated for the 70 days of the MAP SOP. From the first set of results (over all 14 stations) it is seen that GPS PWC agrees with RS PWC within ∼1.9 kg m^{-2} RMS (evaluated as the square root of mean $\mathrm{BIAS}^2 + \mathrm{STD}^2$), which is consistent with previous studies (e.g. Rocken *et al.* 1995; Emardson *et al.* 1998). It is also seen that model analyses are closer to RS than to GPS (compare the biases and standard deviations between Table 2 and Table 3). This is also consistent with previous studies (e.g. Yang *et al.* 1999; Köpken 2001). The difference in bias between the two model analyses is not statistically significant (based on similar reasoning as in subsection 3(a)). However, the differences in bias between GPS and RS, and model and RS, are significant. It is seen that the model analyses have a small dry bias, while GPS has a small wet bias. Note that, since different RS PWCs are computed for comparisons with the GPS and with the model, Table 3 does not allow us to assess the GPS–model bias directly. From the second series of results reported in Table 3,

TABLE 3. MODEL ANALYSES AND GPS OBSERVATIONS OF PWC COMPARED TO RS PWC FOR 14 RS SITES (SEE FIG. 1) AND SUBSETS

Experiment	PWC (kg m^{-2})	Mean BIAS (kg m^{-2})	Standard deviation BIAS (kg m^{-2})	STD (kg m^{-2})	Relative BIAS	Relative STD	Correlation
All soundings on 14 RS stations							
MAPRA	19.3	−0.42	0.99	2.04	−0.020	0.11	0.92
CTRL	19.3	−0.26	1.03	2.27	−0.010	0.12	0.91
GPS	19.1	0.29	1.60	1.89	0.026	0.10	0.93
Only stations with difference in height with model or GPS < 200 m (7 or 9 stations)							
MAPRA	21.0	−0.37	1.14	2.19	−0.011	0.11	0.91
CTRL	21.0	−0.03	1.11	2.55	0.005	0.12	0.89
GPS	21.9	−0.09	1.78	1.82	0.004	0.085	0.95
Only soundings with unsaturated profiles (13 stations)							
MAPRA	17.5	0.10	0.74	1.90	0.006	0.11	0.92
CTRL	17.5	0.24	0.64	1.98	0.014	0.11	0.92
GPS	17.3	0.61	1.41	1.82	0.045	0.11	0.93

Definitions of parameters are as in Table 2. Note that RS PWC is calculated above different surface heights for GPS and models. See text for definitions and algorithm for selection of unsaturated profiles, and Table 1 for station details. Station Cuneo was discarded after data selection.

it is seen that the statistics only change slightly when RS stations are selected for a better match in altitude with either the model (seven stations selected) or the GPS stations (nine stations selected).

One major reason for the different agreement between RS and model PWCs, and between RS and GPS PWCs (Table 3, first subset) is that RS data are assimilated in the model. Hence model PWC is expected to be closer to RS PWC than to GPS PWC. This is reflected in a smaller spatial variability of the bias of model–RS (0.99 kg m^{-2} for MAPRA) compared to GPS–RS (1.60 kg m^{-2}) and a smaller model–RS bias (−0.42 kg m^{-2} for MAPRA) compared to the model–GPS bias (−1.1 kg m^{-2} for MAPRA, reported in Table 2, first subset). However, from Table 3 (first subset) it can be seen that the mean bias (in absolute value) and standard deviation for MAPRA–RS are larger than for GPS–RS (mean bias of −0.42 kg m^{-2} versus 0.29 kg m^{-2}; STD of 2.04 kg m^{-2} versus 1.98 kg m^{-2}). The difference in biases can be explained from dry and wet biases in the RS data (discussed below). The difference in standard deviations has two possible explanations. The first is that not all the RS data used for the comparison were actually assimilated (confirmed in the following). This would also explain the non-zero bias in model–RS PWC. The second explanation is that the weight given to the RS observational data in the 4D-Var assimilation system might be too small (too large an observation error standard deviation or too small a background error standard deviation). This assumption is rather difficult to check. However, it is well known that the partition between background and observation error is a critical aspect in variational data assimilation, which has to be revised as the model changes and as new observations are assimilated (e.g. Järvinen and Undén 1997; Rabier *et al.* 2000).

Figure 4 shows statistics for the differences of model and GPS PWCs relative to RS PWCs, for the 14 RS sites associated with GPS stations. Figures 4(b) and (d) show that the standard deviations (absolute and relative) of PWC differences between CTRL and RS are notably larger than those between MAPRA and RS at a number of RS sites: Bologna (MEDI), Nice (GRAS), Cuneo (TORI), Verona (UPAD), and Vipiteno/Sterzing (BZRG). This is a result of the fact that all five of these RS sites were assimilated in MAPRA, whilst only the first one was assimilated in CTRL.

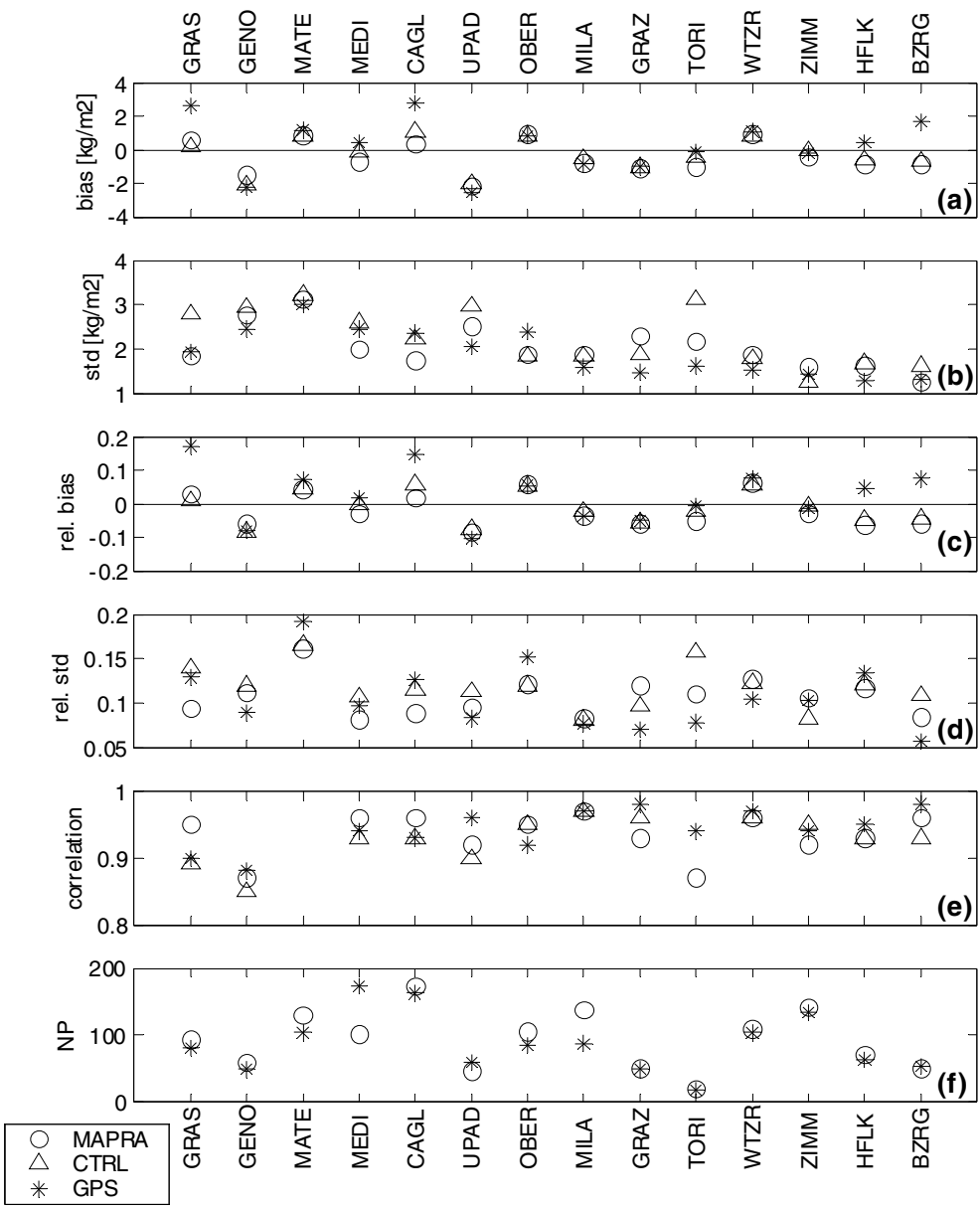

Figure 4. Comparisons of precipitable water vapour contents (PWCs) from the two models MAPRA (circles) and CTRL (triangles), and from Global Positioning System (GPS) stations (stars), with PWCs measured by radiosonde (RS) at 14 RS sites close to GPS stations, during the whole MAP SOP (70 days): (a) bias, (b) standard deviation, (c) relative bias, i.e. (model or GPS PWC minus RS PWC)/(RS PWC), (d) relative standard deviation, (e) correlation coefficient, (f) number of data pairs. Results are ordered by increasing RS site altitude from left to right. GPS four-character identifiers are used instead of RS WMO station numbers for easier comparison with Fig. 2. See text for further details, and Fig. 1 and Table 1 for station locations.

At this stage it is necessary to inspect in more detail the amount and quality of RS data assimilated in each analysis. This information is available from analysis feedback files produced at ECMWF during the assimilation cycles; it is only briefly summarized here. On average, over a domain covering the major MAP operations (43–49°N by 2–17°E) the RS humidity data usage was about 50–60% of available data for both analyses. The remainder of data was generally rejected through blacklisting (Järvinen and Undén 1997). The amount of RS data used in MAPRA was increased by a factor of 4.3 w.r.t. CTRL. Statistics of background and analysis departures computed from feedback files indicate, on average, a slightly larger bias and a smaller RMS in MAPRA compared to CTRL. This is consistent with the results reported in Table 3. In a sub-domain centred on the southern Alps (44–46°N by 7–14°E) both the bias and RMS of background and analysis departures increase for both analyses. This sub-domain contains RS stations Bologna and Milan for CTRL, and in addition, Verona, Genova and Cuneo, for MAPRA. Hence, some of the data assimilated from these stations are likely to contain biases or errors. For MAPRA, the background departure bias is increased by a factor of six and the analysis departure by a factor of four, from the larger to the smaller domain. In the latter, the absolute average vertical bias is 0.12 g kg^{-1} (0.08 g kg^{-1}), for background (analysis) departures. Inspection at individual RS sites highlights significant biases and RMS (in both background and analysis departures) at Verona and Cagliari. The sign and magnitude of these biases are consistent with those reported in Fig. 4 (UPAD and CAGL, respectively). Hence, the analysis feedback files confirm:

(i) The presence of model–RS biases at a number of RS stations;
(ii) The pertinence of results reported in Table 3 and Fig. 4, though not exactly the same RS data are used in the computation of analysis departures;
(iii) The fact that the smaller standard deviation in MAPRA–RS, compared to CTRL–RS, is due to assimilation of more RS data (and more RS stations) in MAPRA.

Points (i) and (iii) suggest that MAPRA PWC might contain biases originating in RS observations. These would explain the MAPRA–GPS biases reported in Table 2 and Figs. 2(a) and (c). GPS observations are therefore used to perform an independent assessment of errors in RS PWC.

Figures 4(a) and (c) show mean differences in PWC between model and GPS relative to RS PWC, as a function of the RS site (each RS site being associated with a GPS station; the names of the GPS stations are used as identifiers in this as in previous figures). Similar systematic biases are observed for the model and GPS at a number of sites (MATE, UPAD, OBER, MILA, GRAZ, WTZR, and ZIMM). For these sites, humidity biases in the data from the nearby RS soundings seem to be either eliminated through data screening during data assimilation or balanced by the assimilation of other data (a reduction of the analysis bias was reported in the feedback files for most of these stations). At most other stations (BZRG, CAGL, GRAS and HFLK) GPS and model exhibit very different biases w.r.t. RS. Dry biases of 1–3 kg m^{-2} in the RS data are observed at five RS stations (GRAS, MATE, CAGL, WTZR and BZRG). At some of the sites the comparisons are likely to be limited by representativeness problems (large differences in altitude between GPS and RS stations at BZRG and GRAS, a large horizontal distance at MATE, and the presence of the sea at CAGL, GENO, GRAS and the RS station in Brindisi). However, at Cagliari (CAGL) Häberli (2003, 2005) estimated the dry bias in RS data to lie in the range 6–13% (depending on the correction scheme). This is consistent with the bias of 15% reported on Fig. 4(c). It is, thus, very likely that the dry bias of -3 kg m^{-2} (-13%) in MAPRA and -2.2 kg m^{-2} (-9%) in

TABLE 4. AVERAGE PARAMETERS FROM LINEAR-FIT REGRESSION BETWEEN MODEL OR GPS PWCS, AND RS PWC OVER 14 OR 13 RS SITES

Experiment	Slope	Offset (kg m^{-2})	RMS (kg m^{-2})	Number of data pairs used
All soundings from 14 RS stations				
MAPRA	0.91	1.48	2.17	1285
CTRL	0.91	1.74	2.38	1285
GPS	0.91	2.23	2.45	1129
Only the soundings with unsaturated profiles (13 stations)				
MAPRA	0.96	1.14	2.09	807
CTRL	0.95	1.44	2.27	807
GPS	0.98	1.52	2.44	713

Model and GPS PWCs are modelled as: slope $\times$ PWC_{rs} + offset. The RMS is calculated as the standard deviation of fitted minus observed data. See text for further details.

CTRL reported for CAGL in Fig. 2 is due to the dry bias in RS data assimilated from Cagliari. Observation screening during the assimilation cycle seemed less effective for this station.

Table 4 shows results from a linear-fit regression of model analyses and GPS PWC to RS PWC. The first part of the table complements the results discussed up to now; especially, the slope and offset parameters indicate that RS data have a dry bias at low PWC values and a wet bias at large PWC. The latter is likely due to contamination from rain and clouds during RS ascents (conditions which were very common during the experiment). In order to check this assumption the RS dataset has been filtered to remove saturated profiles. Saturation was quantified through the equivalent integrated liquid water content (ILW), estimated as the integral of water vapour density (see Eq. (2)), but only over layers where relative humidity was larger than 95%. This RS data selection reduced the overall number of soundings by 38%. Stations where most soundings were rejected are Genova (79%), Bologna (73%), Verona (70%), and Milan (47%). All these stations exhibited negative biases (RS wet biases) in Fig. 4, which were properly reduced through RS data selection. At Cuneo only three soundings were retained, hence this station was discarded from the statistics calculation.

The second part of Table 4 shows linear-fit regression results from the filtered dataset, with the criterion ILW < 0.5 kg m^{-2}. The value for this threshold is not critical (similar results were obtained with ILW < 2.0 kg m^{-2}). The slope and offset parameters now indicate a much better agreement between MAPRA and GPS w.r.t. RS. Compared to both MAPRA and CTRL, RS data still have a dry bias at low PWC and a slight wet bias at large PWC. Compared to GPS, RS data now exhibit a nearly constant dry bias over all values of PWC (0–40 kg m^{-2}), with a slope parameter of 0.98. Figure 5 shows the PWC scatter-plots between MAPRA, CTRL, and GPS w.r.t. RS, and includes PWC data before and after selection. It is seen that most rejected data are for high PWCs. The change in slope and offset of the best-fit line is also clearly visible from this figure for the three datasets.

Average statistics from the selected RS data are also reported in the lower part of Table 3. It is seen that the model–RS bias is now negligible (as expected from assimilation) and that the spatial variability of bias is significantly reduced for both analyses. Again, though the model–RS agreement is very good, this does not preclude the existence of biases. Especially, the overall GPS–RS bias of 4.5% (Table 3, last subset) is consistent with the dry bias estimated by Häberli (2003, 2005). Note that the increase in GPS–RS bias from 2.6% to 4.5% is mainly due to the removal of saturated RS data which contained wet biases. Figure 6 shows the details for individual RS stations after data selection. Compared to Fig. 4, significant bias

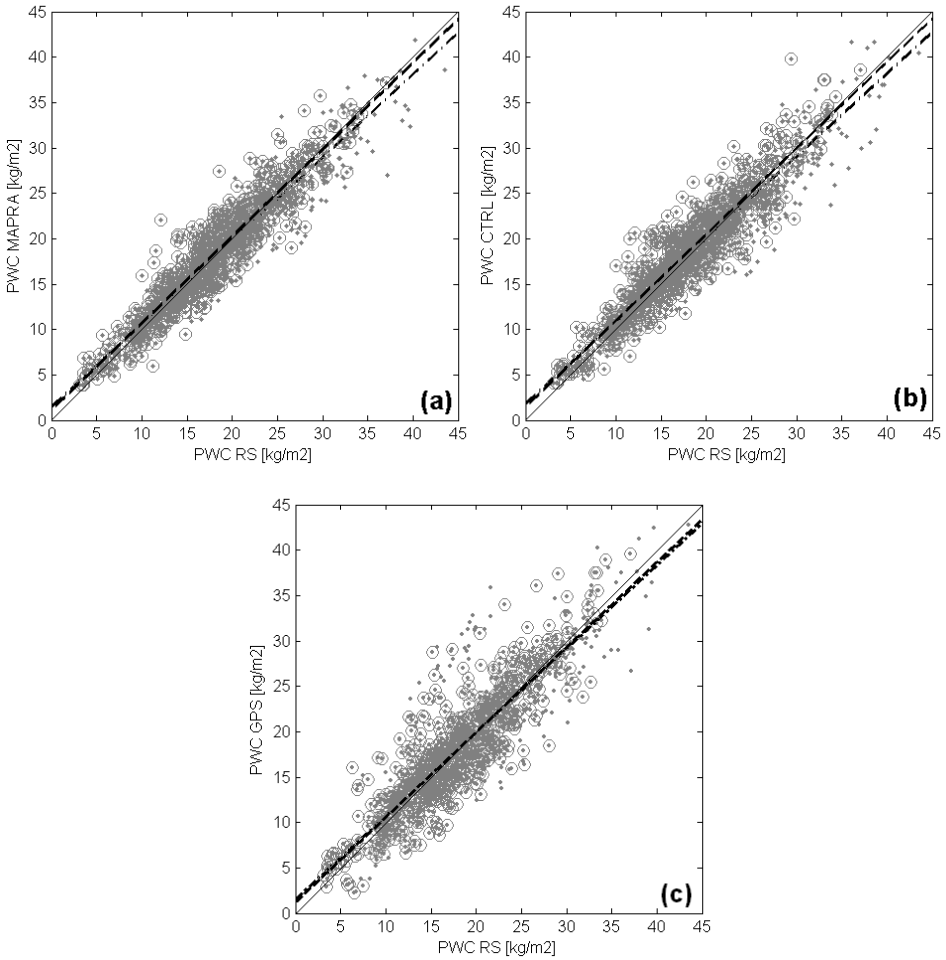

Figure 5. Scatter diagrams of precipitable water vapour contents (PWCs) from radiosonde stations (RS) compared to: (a) the MAPRA model, (b) the CTRL model, and (c) Global Positioning System (GPS) station observations. Data are plotted for all 14 RS stations, with dots for all soundings and circles for only those soundings with unsaturated profiles. Dashed (dot-dashed) line represents best linear fit for all data (unsaturated profiles). Corresponding fit parameters are given in Table 3 (first and third subsets). See text for further details.

reduction is observed at Genova (GENO), Innsbruck (HFLK), Bologna (MEDI) and Verona (UPAD). The standard deviation is either unchanged or reduced at all sites, except at Verona.

The conclusions from section 3 are that:

(i) Model analyses have an overall dry bias w.r.t. GPS observations (−5.5% of total PWC);

(ii) A similar dry bias is observed in RS PWC compared to GPS PWC (−4.5% of total PWC), which suggests that the bias in model analyses might have its source in the RS data that were assimilated;

(iii) The spatial distribution of the bias in both model analyses over the domain is quite uneven (Figs. 2 and 3; Figs. 4 and 6), this is a result of differences in quality and amount of RS data assimilated, and representativeness limitations encountered at a few stations in mountainous areas and/or coastal regions;

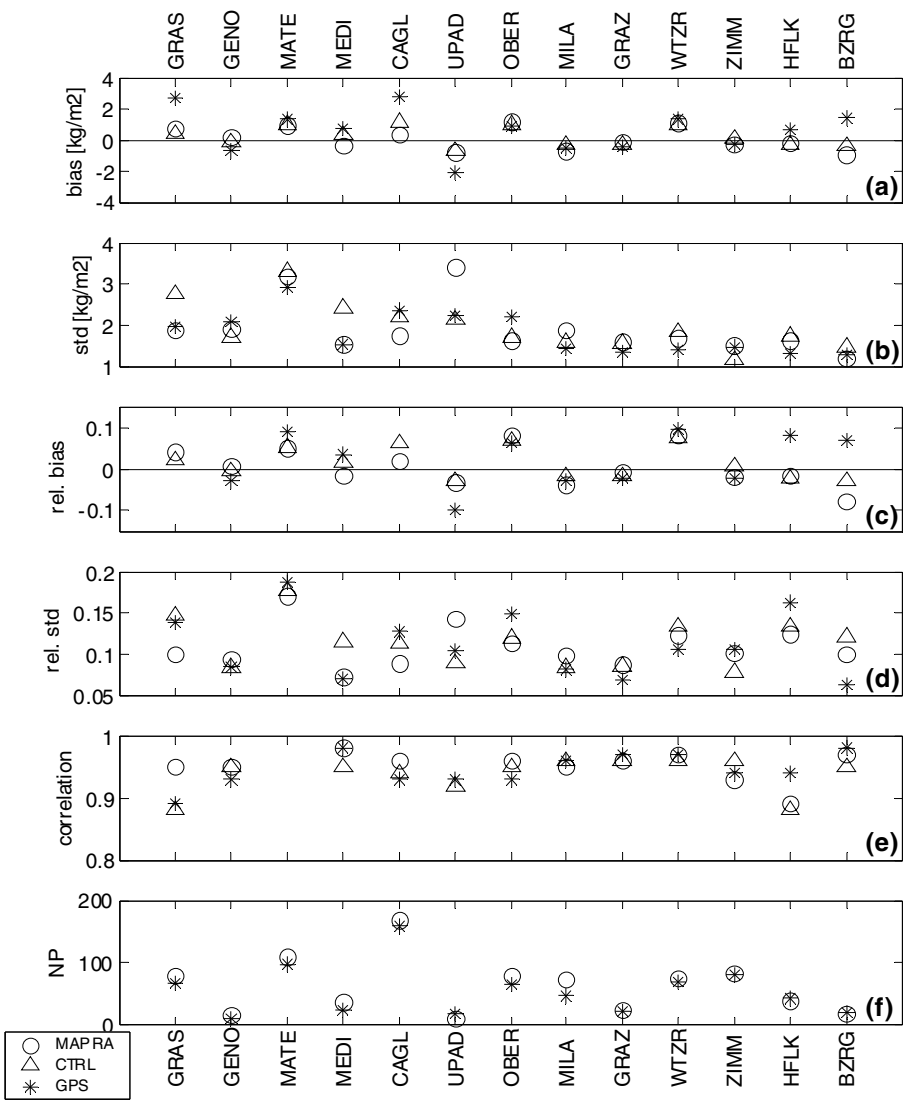

Figure 6. Similar to Fig. 4, with RS data limited to unsaturated profiles (see text for definition).

(iv) A slightly larger PWC bias in MAPRA is evidenced, and explained through the fact that more RS data are assimilated and that these data contained biases;
(v) GPS data demonstrated a high potential for the detection of humidity biases in RS data and model analyses.

4. Time evolution of PWC and PWC differences

(*a*) *Time evolution of PWC differences during the SOP*

Figure 7 shows the temporal evolution of the difference between model PWC and GPS PWC in the southern Alps (GENO, MEDI, MILA, TORI, UPAD). A large variability is seen in these differences throughout the MAP SOP. Both MAPRA and

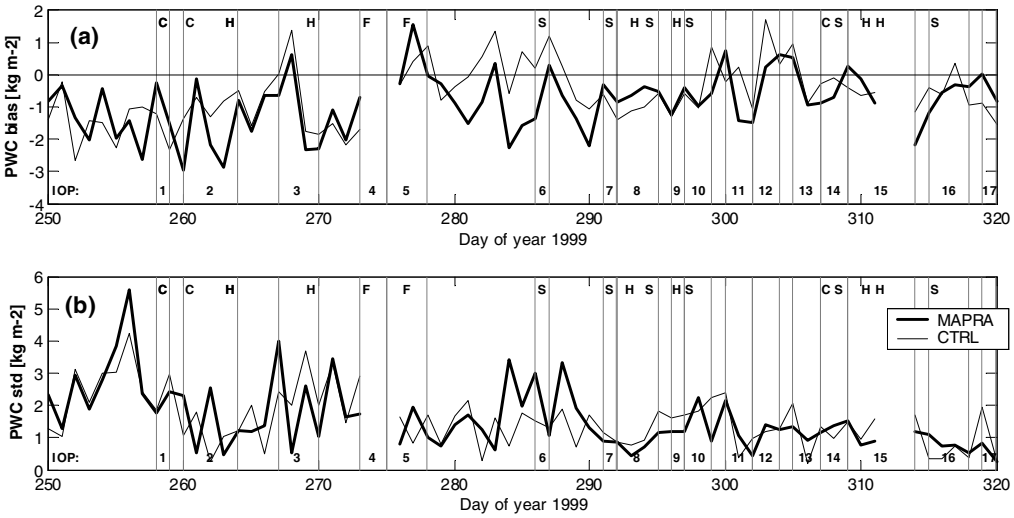

Figure 7. Time series of the difference between model precipitable water vapour contents (PWCs) and Global Positioning System (GPS) station PWC observations (model minus GPS) over five stations in the southern Alps (GENO, MEDI, MILA, TORI, UPAD). Data from each station are first averaged over 24 h. (a) Average (bias), and (b) standard deviation (spatial variability) for the daily differences. The type of event associated with precipitation is indicated in (a) as: C = convective, F = frontal, S = stratiform and H = heavy rain. MAP intensive observation period (IOP) numbers are indicated just above the abscissa of each plot. The period is from day 250 (7 September) to 320 (16 November) of 1999. See text and Fig. 1 for more details.

CTRL analyses show daily biases in the range -3 to $+1.5$ kg m^{-2}. Both the bias and spatial variability are seen to decrease with time. Especially, they seem to be much smaller during the second half of the SOP (after IOP6). This impression is partly due to the fact that both the mean PWC and its temporal variability decrease between September and November as the atmosphere becomes colder. The rapid variations in the PWC differences can have different origins:

(i) Variations in the amount of RS data being assimilated (due to missing or erroneous soundings, e.g. during severe weather);
(ii) Variations in the quality of RS data being assimilated (e.g. sensitivity to the presence of clouds or rain, variations in dry bias with solar radiation, etc.);
(iii) A dependence of the performance of the assimilation system with the weather situation;
(iv) A dependence of the GPS performance with the weather situation. A common feature in all these points is the probable dependence on weather situation.

Since five special RS stations were operated during the MAP–SOP, points (i) and (ii) are very likely to have an impact. This would explain the difference in MAPRA–GPS bias and CTRL–GPS bias, and the larger MAPRA–GPS standard deviation (Figs. 7(a) and (b)). These five RS stations were operated during most but not all IOPs. Since the forecast model is used during 4D-Var assimilation, point (iii) can result from limitations in model physics and dynamics, which are actually dependent on the weather situation. Finally, GPS observations have been shown so far to be insensitive to weather situations, though simulations indicate that in the presence of heavy rain some extra tropospheric delay might map into PWC estimates (Solheim *et al.* 1999). To check this point, differences between model PWC and GPS PWC have been plotted against precipitation recorded in the southern Alps (not shown). No correlation was apparent between peaks

in either bias or spatial variability and rain events. Hence point (iv) is not expected to contribute much to the variability observed in Figs. 7(a) and (b). It is seen in Fig. 7(a) that only for IOP2 and IOP3 is there correlation between large negative biases and convective or heavy rain events. This is mainly attributed to the lack of RS data in the southern Alps.

During IOP2 two fronts passed over the Alps, on 17 (IOP2a) and 20 September (IOP2b). On these two days, very large biases are seen in MAPRA while they are small in CTRL (Fig. 7(a), days 260 and 263). For this particular IOP, only three out of the five special RS stations reported data: Verona (UPAD), Vipiteno (BZRG), and Nice (GRAS). This reduced slightly the impact of different data usage. The agreement of MAPRA with the RS station Verona is quite bad ($+2$ to -10 kg m^{-2}) which suggest that some of the soundings were blacklisted. The latter two RS stations are actually not located in the southern Alps. Their influence on the results of Fig. 7 is thus indirect. However, the agreement of RS PWC and GPS PWC at these stations was also very bad during the period of IOP2: 0 to -7 kg m^{-2} at Nice and 0 to -5 kg m^{-2} at Vipiteno. At these two stations, MAPRA is found to follow closely the RS PWC, while CTRL behaves differently. Hence, these stations might also have an impact on the PWC analyses in the southern Alps. During IOP2, the difference in performance of MAPRA and CTRL is at least partly due to differences in the usage of RS data. Another source of difference is the impact of wind profilers discussed by Keil and Cardinali (2004).

(*b*) *PWC evolution in heavy precipitation target areas*

The study of orographic precipitation mechanisms was one of the main goals in the design of the MAP SOP (Bougeault *et al.* 2001). Orographic precipitation in the southern Alps is in most cases associated with the passage of fronts feeding strong low-level jets ahead (Buzzi and Foschini 2000). These jets transport large amounts of moisture from the Ligurian Sea (in the early stage) and Adriatic Sea (as the fronts pass over northern Italy) to the southern Alps region. Associated precipitation is in most cases stratiform; however, moderate to strong convection can appear, which is triggered in unstable or destabilized atmospheres through the interaction of atmospheric motion with orography (upslope forcing over the Alpine barrier or convergence of different flows in the Alpine arc). Heavy precipitation and flooding is frequently observed in two specific regions: the Ticino-Lago Maggiore and Friuli-Veneto (Frei and Schär 1998). These were two major target areas of the MAP SOP. Additionally, Liguria is an important source region and the Po Valley is an important channelling region for the moist inflow from the Adriatic Sea.

Figure 8 shows the evolution of the daily precipitation (from Frei and Häller 2002) and PWC from GPS observations and model analyses during the SOP. Precipitation is shown for three regions (Piedmont-Lago Maggiore, Liguria, and Friuli-Veneto) and PWC for three corresponding GPS sites (GENO, TORI, and UPAD), each one located in one of the three regions. In the three regions, 18 days, and 14 out of 17 IOPs, show significant precipitation. Heavy precipitation was recorded during IOPs 2b, 3, 5, 8, 9, and 15, with flooding during IOPs 2b, 3 and 15. Most of these events are characterized by the passage of fronts (details on individual IOPs can be found on http://www.atmos.washington.edu/gcg/MG/MAP/summ/).

The time evolution of PWC throughout the 70-day period is highly correlated between the three GPS sites (Fig. 8). This emphasizes the synoptic scale of moisture advection. Frontal passages are clearly marked by abrupt changes in PWC. A small signature of a diurnal cycle in PWC can also be seen, especially in periods of clear sky (e.g. at GENO before IOP1, which was also observed at GRAS and CAGL, not shown).

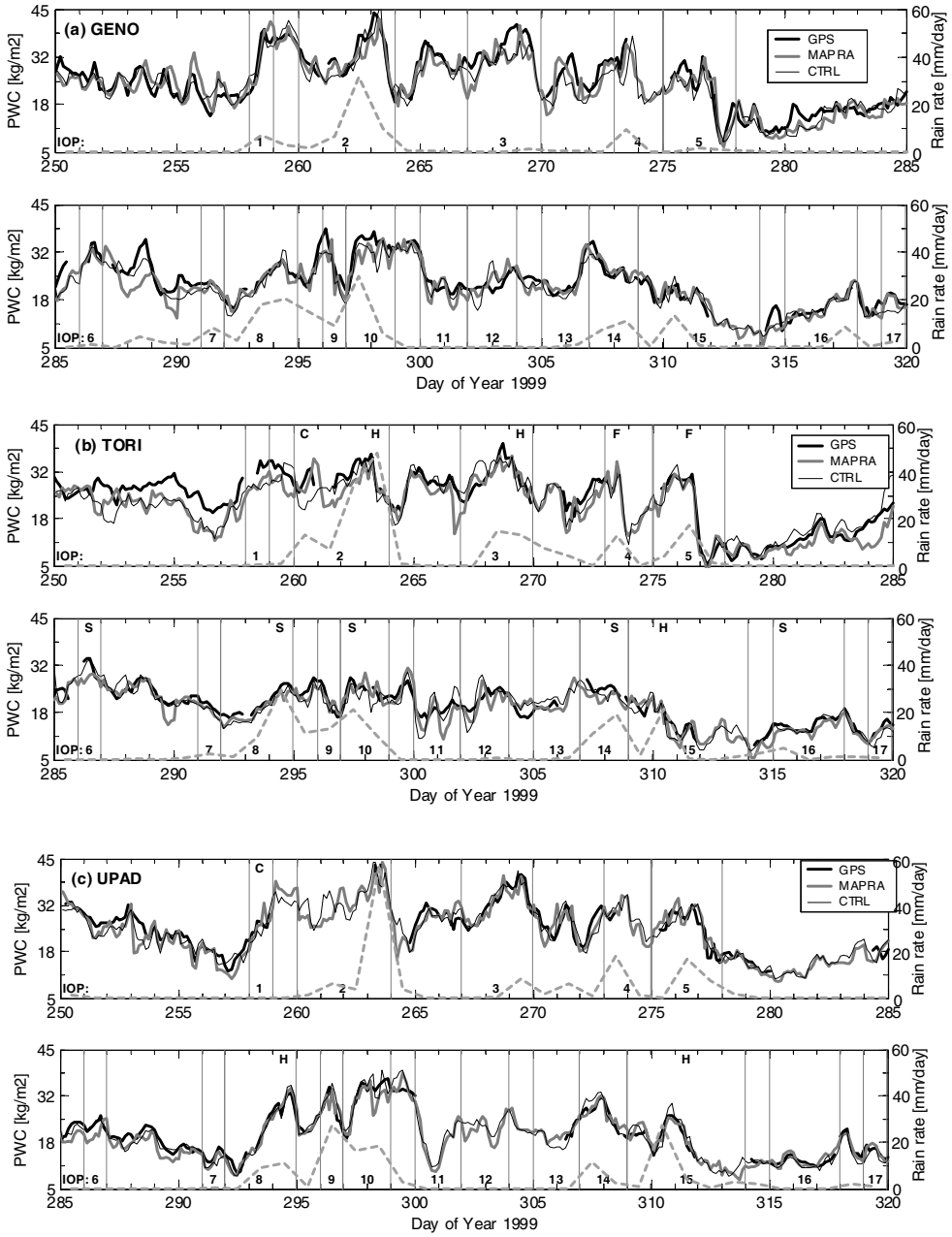

Figure 8. Time series of precipitation (grey dashed line, right-hand axis) in three regions (Liguria, Piedmont-Lago Maggiore, and Friuli-Veneto, indicated as domains D1, D2 and D3, respectively, in Fig. 3(a)), together with model precipitable water vapour contents (PWCs) from the MAPRA and CTRL model analyses, and the PWC from Global Positioning System (GPS) stations, at three sites: (a) GENO, (b) TORI, and (c) UPAD, located one in each domain. Precipitation time series were computed as area-averages from gridded data analyses of rain-gauge observations (version 2.0, produced by Frei and Häller (2002)). The type of event associated with precipitation and MAP IOP numbers are indicated in each panel (as in Fig. 7). See text for further details.

Correlation between PWC peaks and precipitation is not so obvious (e.g. the PWC peak during IOP6 is not associated with significant precipitation). On the other hand, precipitation peaks are mostly accompanied by large values or peaks in PWC. Although no systematic phase relation is observed between PWC and rain rate, PWC often drops at the end of the events.

Model analyses and GPS at the three stations in Fig. 8 exhibit overall good agreement. However, on many occasions there are differences between the two analyses, and also between the model analyses and GPS. A few events are not reproduced by the two model analyses (e.g. peaks at GENO for day 288; at TORI for days 259, 260, 266 and 283–285). The model–GPS agreement is best at UPAD (see also Fig. 2). This is linked to the fact that more RS data were assimilated from UPAD than from GENO and TORI. Significant differences between both model analyses are observed, which occasionally reach $\sim$10 kg m^{-2} (50% of total PWC). Such differences are generally due to a time lag of PWC peaks between the analyses. No systematic correlation is observed between the precipitation events and differences between the two model analyses, or between the model analyses and GPS.

5. Conclusion

PWCs from MAPRA, the final MAP re-analysis, and CTRL, the control analysis which used only conventional data, have first been compared to PWC observations from 21 GPS stations. Overall, a dry bias of -1 kg m^{-2} (-5.5% of total PWC) with a standard deviation of $\sim$2.6 kg m^{-2} (13% of total PWC) was diagnosed in both model analyses w.r.t. GPS. The bias at individual sites was quite variable: from -4 to $\sim$0 kg m^{-2}. The largest differences were observed at stations located in mountainous areas and/or near the sea. At a small number of these stations the comparison of point observations and model fields was limited by differences in representativeness. In the southern Alps both model analyses showed smaller biases, but the bias in MAPRA was larger than that in CTRL. It has been shown that this is linked to different data usage, and is very likely due to dry biases in the RS data, as more RS data were assimilated in MAPRA.

Dry biases in RS PWC were found at a larger number of sites compared to GPS PWC. A strong correlation was also observed between the GPS–RS PWC bias and model–RS PWC bias at most RS sites. Similar biases were reported in analysis feedback statistics during the assimilation experiments performed at ECMWF. This indicates that data screening during assimilation was effective for these RS stations. However, at some sites, a correlation was found between the model–GPS bias and the GPS–RS bias. There, the model bias was very likely due to a bias in the RS data (Cagliari, Nice, Vipiteno).

A linear-regression fit between GPS and RS data showed that RS data exhibit a dry bias at small PWC values and a wet bias at large PWC values. Contamination from clouds and rain is suspected to have caused the wet bias. A selection of unsaturated RS profiles significantly improves this linear-regression fit. A dry bias in RS PWC of -4.5% was found compared to GPS. This is consistent with an independent evaluation of dry biases in RS data from Häberli (2003, 2005). It was also noted that analysis departures (model–RS) reduced with the selected RS data (1% bias and 11% standard deviation for MAPRA). The -5.5% dry bias in model analysis could thus be explained from the dry bias in the RS data.

Inspection of time evolution during the MAP SOP showed occasionally up to 5–10 kg m^{-2} differences, either between the model analyses or between

analyses and GPS. They were associated with severe weather events, with lack of upper-air humidity data (e.g. at Genova and Torino), and with time lags in the PWC evolution from the two model analyses. Such large differences contributed strongly to the overall observed standard deviations.

The comparison of model analyses and GPS PWC to RS PWC showed an agreement of $\sim$1.9 kg m^{-2} RMS. Discrepancies between GPS and model analyses or GPS and RS PWCs could not be associated with limitations in GPS PWC estimates. Thus, this work gives good confidence that the assimilation of GPS data would be beneficial to the quality of humidity fields in the analysis. Positive impact of assimilating PWC data from microwave radiometers or GPS has been demonstrated with mesoscale models, especially for the simulation of precipitation events (Kuo *et al.* 1993, 1996; Guo *et al.* 2000; Falvey and Beavan 2002). These experiments used either 'nudging' or 4D-Var assimilation techniques. Assimilation of GPS ZTD rather than PWC has been discussed by e.g. De Pondeca and Zou (2001). However, the assimilation of ZTD assumes that the refractivity of dry air is properly modelled, i.e. that the model's surface pressure is correct. This might not be the case if the horizontal resolution of the model is too large (as discussed in subsection 2(c)) unless collocated surface pressure measurements are assimilated. On the other hand, using surface pressure measurements for the conversion of ZTD into PWC outside of the model makes the detection of biases in these measurements more difficult.

For future work, the assimilation of GPS data (ZTD or PWC) in a mesoscale model is considered for the simulation of the MAP IOP2a episode. For this IOP, bad results were obtained in simulations initialized with both MAPRA and CTRL analyses, as well as with the operational analysis (Lascaux *et al.* 2004). More generally, assimilation of special MAP observations in mesoscale models is considered as ongoing work in the MAP project.

On the global scale, the use of GPS data from the existing globally distributed permanent network (considered in Hagemann *et al.* (2003) for validation purposes) would probably have a positive impact on humidity analyses, which are especially poor in the Tropics (Andersson *et al.* 2005). The largest impact of such data is expected over land, in regions where upper-air humidity data are lacking. Although the quantity of humidity data from satellite observations is continuously increasing, most sensors either work only over the oceans (e.g. microwave radiometers) or are not sensitive to the lower troposphere (e.g. infrared profilers). Moreover, these operational satellites are limited to twice-daily overpasses, and thus cannot provide continuous observations locally. Comparatively, GPS stations provide a continuous observation of the total column PWC in all weather conditions. This is probably the main strength of ground-based GPS networks.

ACKNOWLEDGEMENTS

The authors would like to thank several colleagues for discussing the results presented in this paper and for providing technical support. They are especially grateful to: Carla Cardinali, Pedro Viterbo, and Dominique Lucas, from ECMWF; Frank Lefevre and Jacques Pelon, from Service d'Aéronomie, Paris, France; and Erik Doerflinger, from Laboratoire de Dynamique de la Lithosphère, Montpellier, France. The authors also thank the two anonymous reviewers for providing detailed comments, which help to improve the readability of this paper. This work was partially funded by the Programme National de Télédétection Spatiale (PNTS) of the Institut National des Sciences de l'Univers (INSU).

REFERENCES

Andersson, E., Bauer, P., Beljaars, A., Chevallier, F., Hólm, E., Janisková, M., Kållberg, P., Kelly, G., Lopez, P., McNally, A., Moreau, E., Simmons, A., Thépaut, J.-N. and Tompkins, A. 2005 Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system. *Bull. Am. Meteorol. Soc.*, **86,** 387–402

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A. and Ware, R. H. 1992 GPS meteorology: Remote sensing of the atmospheric water vapor using the Global Positioning System. *J. Geophys. Res.*, **97,** 15787–15801

Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C. and Ware, R. H. 1994 GPS Meteorology: Mapping zenith wet delay onto precipitable water. *J. Appl. Meteorol.*, **33,** 379–386

Bock, O., Flamant, C. and Duquesnoy, T. 2001 Integrated water vapor estimated by GPS compared to independent observations during MAP. *Proc. SPIE*, **4539,** 289–298

Bock, O., Flamant, C., Richard, E., Keil, C. and Bouin, M. N. 2004 'Validation of precipitable water from ECMWF model with GPS data during the MAP SOP'. In proceedings of the 11th conference on mountain meteorology, 21–25 June 2004, Bartlett, NH. American Meteorological Society, Boston, USA

Bougeault, P., Binder, P., Buzzi, A., Dirks, R., Houze, R., Kuettner, J., Smith, R. B., Steinacker, R. and Volkert, H. 2001 The MAP Special Observing Period. *Bull. Am. Meteorol. Soc.*, **82,** 433–462

Buzzi, A. and Foschini, L. 2000 Mesoscale meteorological features associated with heavy precipitation in the southern-Alpine region. *Meteorol. Atmos. Phys.*, **72,** 131–146

Buzzi, A., Davolio, S., D'Isidoro, M. and Malguzzi, P. 2004 The impact of resolution and of MAP reanalysis on the simulations of heavy precipitation during MAP cases. *Meteorol. Z.* **13,** 91–97

De Pondeca, M. S. F. V. and 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

Ducrocq, V., Ricard, D., Lafore, J.-P. and Orain, F. 2002 Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Weather and Forecasting*, **17,** 1236–1256

Emardson, T. R., Elgered, G. and Johansson, J. M. 1998 Three months of continuous monitoring of atmospheric water vapor with a network of Global Positioning System receivers. *J. Geophys. Res.*, **103**(D2), 1807–1820

Faccani, C., Ferretti, R. and Visconti, G. 2003 High-resolution weather forecasting over complex orography: Sensitivity to the assimilation of conventional data. *Mon. Weather Rev.*, **131,** 136–154

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

Fang, P., Bevis, M., Bock, Y., Gutman, S. and Wolfe, D. 1998 GPS meteorology: Reducing systematic errors in geodetic estimates for zenith delay. *Geophys. Res. Lett.*, **25,** 3583–3586

Frei, C. and Häller, E. 2002 Mesoscale precipitation analyses from MAP SOP rain-gauge data. http://www.map.ethz.ch/sop-doc/rr_sop/rr_sop.htm

Frei, C. and Schär, C. 1998 A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.* **18,** 873–900

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

Haase, J. S., Ge, M., Vedel, H. and Calais, E. 2003 Accuracy and variability of GPS tropospheric delay measurements of water vapor in the western Mediterranean. *J. Appl. Meteorol.*, 1547–1568

Häberli, Ch. 2003 'Assessment and correction of the dry bias in radiosonde humidity data during the MAP-SOP'. Pp. 221–224 in Extended abstracts of the international conference on Alpine meteorology and MAP meeting, May 19–23, 2003, Brig, Switzerland. Vol. A. Publications of MeteoSwiss, Geneva, Switzerland

Häberli, Ch. 2005 Assessment, correction and impact of the dry bias in radiosonde humidity data during the MAP SOP. *Q. J. R. Meteorol. Soc*., in press

Hagemann, S., Bengtsson, L. and Gendt, G. 2003 On the determination of atmospheric water vapor from GPS measurements. *J. Geophys. Res*., **108**(D21)**,** 4678

Hortal, M. and Simmons, A. J. 1991 Use of reduced Gaussian grids in spectral models. *Mon. Weather Rev*., **119,** 1057–1074

Järvinen, H. and Undén, P. 1997 Observation screening and background quality control in the ECMWF 3D-Var data assimilation system. Tech. Memo. 236. European Centre for Medium-Range Weather Forecasts, Shinfield, Reading, UK

Keil, C. and Cardinali, C. 2004 The ECMWF re-analysis of the Mesoscale Alpine Programme Special Observing Period. *Q. J. R. Meteorol. Soc*., **130,** 2827–2850

Klein Baltink, H., van der Marel, H. and van der Hoeven, A. G. A. 2002 Integrated atmospheric water vapor estimates from a regional GPS network. *J. Geophys. Res.*, **107**(D3)**,** 10.1029

Koch, S. E., Aksakal, A. and McQueen, J. T. 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

Köpken, C. 2001 Validation of integrated water vapor from numerical models using ground-based GPS, SSM/I, and water vapor radiometer measurements. *J. Appl. Meteorol*., **40,** 1105–1117

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

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

Lascaux, F., Richard, E., Keil, C. and Bock, O. 2004 Impact of the MAP reanalysis on the numerical simulation of the MAP IOP2a convective system, ICAM 2003. *Meteorol. Z*., **13,** 49–54

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A. 2000 The ECWMF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc*., **126,** 1143–1170

Richard, E., Cosma, S., Tabary, P., Pinty, J. P. and Hagen, M. 2003 High-resolution numerical simulations of the convective system observed in the Lago Maggiore area on 17 September 1999 (MAP IOP 2a). *Q. J. R. Meteorol. Soc*., **129,** 543–563

Rocken, C., VanHove, T., Johnson, J., Solheim, F., Ware, R. H., Bevis, M., Chiswell, S. R. and Businger, S. 1995 GPS/STORM-GPS sensing of atmospheric water vapor for meteorology. *J. Atmos. Oceanic Technol*., **12,** 468–478

Romero, R., Ramis, C., Alonso, S., Doswell, C. A. and Stensrud, D. J. 1998 Mesoscale model simulations of three heavy precipitation events in the western Mediterranean region. *Mon. Weather Rev*., **126,** 1859–1881

Solheim, F., Vivekanandan, J., Ware, R. and Rocken, C. 1999 Propagation delays induced in GPS signals by dry air, water vapor, hydrometeors, and other particulates. *J. Geophys. Res*., **104**(D8)**,** 9663–9670

Wang, J., Cole, H. L., Carlson, D. J., Miller, E. R., Beierle, K., Paukkunen, A. and Laine, T. K. 2002 Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA COARE data. *J. Atmos. Oceanic Technol.*, **19,** 981–1002

Yang, X., Sass, B. H., Elgered, G., Johansson, J. M. and Emardson, T. R. 1999 A comparison of precipitable water vapor estimates by an NWP simulation and GPS observations. *J. Appl. Meteorol*., **38,** 941–956