Assessment of small-scale integrated water vapour variability during HOPE

S. Steinke[1,2], S. Eikenberg[1], U. Löhnert[1], G. Dick[3], D. Klocke[4], P. Di Girolamo[5], and S. Crewell[1]

[1]Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany
[2]Hans-Ertel-Centre for Weather Research, Germany
[3]GeoForschungsZentrum Potsdam, Potsdam, Germany
[4]Deutscher Wetterdienst, Offenbach, Germany
[5]Scuola di Ingegneria, Università della Basilicata, Potenza, Italy

Correspondence to: S. Steinke (ssteinke@meteo.uni-koeln.de)

Abstract. The spatio-temporal variability of integrated water vapour (IWV) on small scales of less than 10 km and hours is assessed with data from the 2 months of the High Definition Clouds and Precipitation for advancing Climate Prediction (HD(CP)^2) Observational Prototype Experiment (HOPE). The statistical intercomparison of the unique set of observations during HOPE (microwave radiometer (MWR), Global Positioning System (GPS), sun photometer, radiosondes, Raman lidar, infrared and near-infrared Moderate Resolution Imaging Spectroradiometer (MODIS) on the satellites Aqua and Terra) measuring close together reveals a good agreement in terms of random differences (standard deviation $\leq 1\,\mathrm{kg\,m^{-2}}$) and correlation coefficient (≥ 0.98). The exception is MODIS, which appears to suffer from insufficient cloud filtering.

For a case study during HOPE featuring a typical boundary layer development, the IWV variability in time and space on scales of less than 10 km and less than 1 h is investigated in detail. For this purpose, the measurements are complemented by simulations with the novel ICOsahedral Non-hydrostatic modelling framework (ICON), which for this study has a horizontal resolution of 156 m. These runs show that differences in space of 3–4 km or time of 10–15 min induce IWV variabilities on the order of $0.4\,\mathrm{kg\,m^{-2}}$. This model finding is confirmed by observed time series from two MWRs approximately 3 km apart with a comparable temporal resolution of a few seconds.

Standard deviations of IWV derived from MWR measurements reveal a high variability ($> 1\,\mathrm{kg\,m^{-2}}$) even at very short time scales of a few minutes. These cannot be captured by the temporally lower-resolved instruments and by operational numerical weather prediction models such as COSMO-DE (an application of the Consortium for Small-scale Modelling covering Germany) of Deutscher Wetterdienst, which is included in the comparison. However, for time scales larger than 1 h, a sampling resolution of 15 min is sufficient to capture the mean standard deviation of IWV. The present study shows that instrument sampling plays a major role when climatological information, in particular the mean diurnal cycle of IWV, is determined.

1 Introduction

Water vapour is not only the most effective greenhouse gas (Kiehl and Trenberth, 1997) but also an important part of the hydrological cycle, so that the exact knowledge on atmospheric moisture is absolutely essential for both numerical weather prediction (NWP; e.g. Weckwerth et al., 1999) and climate modelling (e.g. Bony et al., 2006). Due to its importance, water vapour has been investigated in several field campaigns such as the HYdrological cycle in the Mediterranean EXperiment (HyMeX; Drobinski et al., 2014) and the Convective and Orographically-induced Precipitation Study (COPS; Wulfmeyer et al., 2011). However, there is still need for research about its role in various atmospheric processes. The interaction between atmospheric humidity and convec-
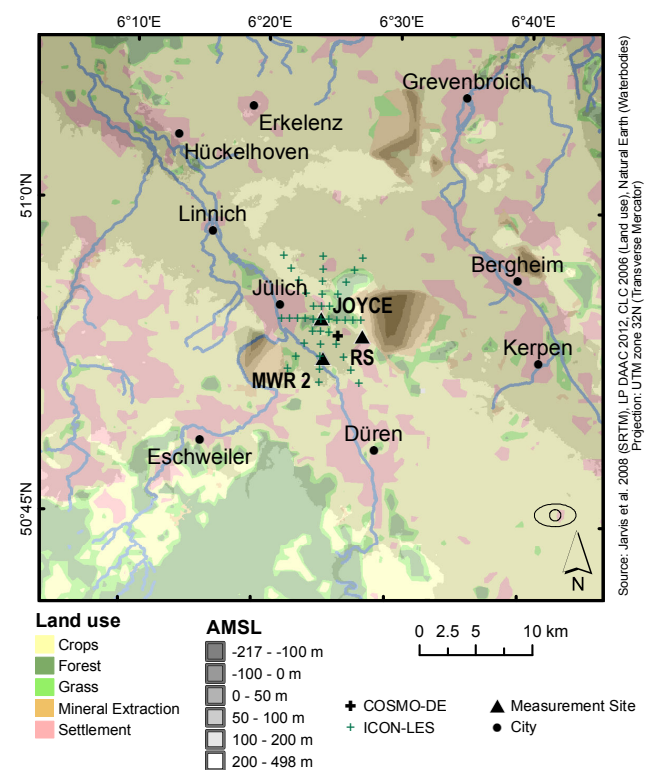
Figure 1. Map of measurement area. The measurement sites of GPS, MWR, sun photometer, BASIL (all JOYCE), radiosondes (RS), and the MWR only used in Sect. 3 (MWR 2) are marked with a black triangle. The ellipses in the lower right corner illustrate the maximum and minimum size of MODIS footprints. Black and green crosses indicate COSMO-DE and ICON grid points used in Sect. 3.

tion, for example, is still poorly understood (Sherwood et al., 2010).

The amount of water vapour in the atmosphere is influenced by both mixing and transport as well as sources and sinks, such as condensation and evaporation of clouds and precipitation and evaporation of soil moisture. The subsequent vertical transport of the atmospheric water vapour occurs by turbulent mixing on small scales (1 min and 10 m). Convective processes on different scales, such as mesoscale up- and downdrafts and eddies at convective (10–30 min, < 2 km) and smaller scales, dominate the further vertical transport of water vapour. A prominent example on the convective scale is the atmospheric boundary layer where evaporation from the heterogeneous land surface and turbulent mixing creates strong water vapour variability (Shao et al., 2013, cf. Fig. 10). In additional to these circulations, water vapour is transported by advection of air masses on large scales (> 1000 km, > 1 day). The combination of these various processes results in a high variability of atmospheric water vapour in both space and time.

Knowledge on water vapour variability is valuable for improving subgrid-scale model parametrizations, for model

evaluation, and for instrument intercomparisons (Vogelmann et al., 2011). Kahn et al. (2011) compare the integrated water vapour (IWV) variability in NWP and climate models with those directly observed by Atmospheric InfraRed Sounder (AIRS) observations and airborne measurements with a focus on stratocumulus regions over ocean. They find large differences in the magnitude of IWV variance, leading to the conclusion that in the future satellite observations are needed with a higher resolution than currently planned (10–30 km).

By moving to very high-resolution simulations, atmospheric models become less prone to uncertainties induced by parametrizations at the cost of computationally expensive simulations. The High Definition Clouds and Precipitation for advancing Climate Prediction (HD(CP)^2) initiative aims to build and use such a model with horizontal grid spacings down to 100 m based on the ICOsahedral Nonhydrostatic (ICON; Zängl et al., 2014) model. In order to provide the critical observations to evaluate this model at small scales, the HD(CP)^2 Observational Prototype Experiment (HOPE) took place from 1 April to 31 May 2013 at the Forschungszentrum Jülich, Germany (cf. Fig. 1). During this 2-month period, standard instrumentation for observing water vapour at the Jülich ObservatorY for Cloud Evolution (JOYCE; Löhnert et al., 2014), including the Global Positioning System (GPS) antenna of the GeoForschungsZentrum Potsdam (GFZ), a scanning microwave radiometer (MWR), and a sun photometer from the AErosol RObotic NETwork (AERONET), was complemented by frequent radiosoundings, four additional MWRs, and the BASILicata Raman lidar system (BASIL), all within less than 4 km distance of each other. In addition to the ground-based measurements, IWV estimates from two Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals, near-infrared (NIR) and infrared (IR), that provide information with spatial resolution of 1 and 3 km respectively are available from satellite overpasses. In contrast to other space-based instruments capable of detecting IWV, MODIS provides horizontally high-resolved IWV fields enabling a look at the horizontal gradients of IWV on smaller scales.

Different instruments sample different atmospheric conditions due to different integration times, beam widths, geometries, sampling strategies, locations, etc. For IWV, the measurement height is of particular importance as the water vapour column over the same altitude range needs to be considered and therefore corrections are necessary (cf. Böhme et al., 2011; Buehler et al., 2012). Many studies compare various IWV measurements in different geographical regions and for different time periods using different criteria for temporal and spatial matching and elevation corrections (cf. Bennouna et al., 2013; Martin et al., 2006; Morland et al., 2009; Palm et al., 2010; Schneider et al., 2010; Torres et al., 2010). Frequently these comparisons involve data sets with more than 1 h temporal and more than 20 km spatial differences as well as with different horizontal resolutions. Buehler et al. (2012) investigate the representativeness error resulting from insufficient collocation and resolution mismatch for a high-

latitude region using the Nonhydrostatic Icosahedral Atmospheric Model (NICAM; Satoh et al., 2008) with 3.5 km horizontal resolution. GPS data are used as reference and the representativeness error is calculated for ground-based slant column and satellite measurements as well as for the European Centre for Medium-Range Weather Forecasts Reanalysis ERA-Interim. They derive values of approximately 0.6–1.4 kg m^{-2} for spatial scales of several 10 km. It has to be noted that GPS does not provide true column measurements, as one observation over a 15 min interval includes the atmospheric delay measured along several links between the GPS ground station and multiple satellites.

The goal of the present study is three-fold: firstly, we aim to characterize the variability of IWV for spatial scales smaller than 10 km and temporal scales smaller than 1 h and to estimate the ability of different measurements to represent this variability. In doing so, we extend previous studies to even smaller scales, by using zenith-pointing MWR measurements which are available at a temporal resolution of approximately 2 s. To this end, a case study at the continental mid-latitude site JOYCE is presented and the unique set of instruments from HOPE is complemented by very high-resolution (156 m) simulations from the novel atmospheric model ICON. Secondly, with the goal of providing a realistic error estimate for the individual instruments observing IWV, we perform a statistical, multi-instrumental comparison covering the HOPE period. This includes the investigation of the variability of IWV on a wide range of temporal scales from a few minutes, to a couple of hours to its mean diurnal cycle. Thirdly, the ability of the novel ICON model to capture the daily IWV cycle of a realistic case is assessed.

The study is structured as follows: an overview of all instruments and the respective retrievals used in this study is given in Sect. 2.1. A first version of the ICON model is introduced together with the operational regional NWP model of Deutscher Wetterdienst (DWD) at 2.8 km horizontal resolution in Sect. 2.2. Details on how to match the various data sets are given in Sect. 2.3. Observations and model runs are analysed within a case study for a day with typical boundary layer development in order to estimate scale-dependent IWV variability (Sect. 3). The analysis is extended over the full duration of HOPE, providing statistics on the agreement between the different instruments, the relative merits of the different instruments to capture the temporal IWV variability, and the diurnal cycle (Sect. 4). The summary and conclusions are given in Sect. 5.

2 Data

2.1 Observations

In the following, the instruments used in the present study, their measurement principle, and their retrieval methods are introduced. Table 1 gives an overview of the accuracy, spatial and temporal resolution, and limitations in terms of weather conditions of the individual instruments.

2.1.1 Microwave radiometer

Two microwave radiometers, one located at JOYCE and one 3.3 km south of JOYCE (cf. Fig. 1), are used in the present study. Both MWR are Humidity and Temperature PROfilers (HATPRO; Rose et al., 2005). Here only zenith-pointing HATPRO measurements with a temporal resolution of up to 2 s are used. The antenna has a half-power beam width of 3.5° for the water-vapour-sensitive channels. Thus, the MWR measures a comparatively narrow part of the atmosphere. From this volume, it receives brightness temperatures at six frequencies along the water vapour absorption line (22.24, 23.04, 23.84, 25.44, 26.24, 27.84 GHz) and one frequency in an atmospheric window (31.40 GHz). With a low noise level of approximately 0.05 K in the measured brightness temperatures HATPRO is able to detect not only small variations in atmospheric water vapour but also cloud water whose emission increases with frequency in the microwave spectral range. The absolute accuracy of the observed brightness temperatures determined by the calibration procedure (Maschwitz et al., 2013) is 0.5 K.

IWV is derived following a statistical approach based on a least squares linear regression model (Löhnert and Crewell, 2003) from the multifrequency brightness temperatures assuming the error characteristics mentioned above.

To derive the coefficient vectors, a training data set of more than 13 000 nonprecipitating radiosoundings at De Bilt, Netherlands, is used. With this algorithm, IWV can be derived with a random error of approximately 0.5–0.8 kg m^{-2} from zenith measurements. The systematic error is assumed to be 0.5 kg m^{-2} and the noise level is 0.05 kg m^{-2}. Note that the MWR is able to measure automatically under all weather conditions with the exception of when the radome is wet. In these cases, no IWV values are provided.

2.1.2 GPS ground station

Although the main aim of GPS is precise positioning for navigation, remarkable progress in using GPS for retrieval of IWV has been achieved during the last decades (Bevis et al., 1992; Rocken et al., 1997; Fang et al., 1998).

The basic quantity estimated by any GPS receiver is the signal travel time from the GPS satellite to the receiver. From the travel times of up to 12 GPS satellites with an elevation angle larger than 7° and the satellite positions, the receiver position is estimated. The GPS signal consists of electromagnetic waves propagating through the atmosphere with frequencies of 1575.42 and 1227.60 MHz. The travel time also provides information on the atmosphere along the signal path. The signal is slightly delayed by the atmosphere and this delay, as compared to an undisturbed signal propagation in vacuum, depends on the atmospheric state. There are two

Table 1. Temporal resolution, spatial resolution or representativeness, limitations, and systematic (s), random (r) or combined error of measurements as found in literature for the instruments used in the present study.

Instrument	Temporal resolution	Spatial resolution/ representativeness	Limitations	Uncertainty kg m^{-2} or %	Reference
MWR HATPRO	≈2 s	3.5° beam width; 122 m beam width at 2 km height	no measurements during rain	0.5 (s) 0.5–0.8 (r)	Rose et al. (2005)
GPS	15 min	ca. 32 km[1]	no zenith measurement	1–2	Gendt et al. (2004)
Sun photometer	10 min	1.2° beam width	daytime/clear-sky only, direction towards sun	10 %	Alexandrov et al. (2009)
Graw DFM-09 Radiosonde	at least 1 h	drift up to 100 km	drift, measurement takes ca. 1 h	1.2 (s) 1.7 (r)	Wang and Zhang (2008)
MODIS-NIR	≤6 times per day	1 km	daytime/clear-sky only	5–10%	Gao and Kaufmann (2003)
MODIS-IR	≤6 times per day	3 km	clear-sky only	5–10%	Seemann et al. (2003)
BASIL	10 s–5 min	vertical resolution of 30 m	no measurements during rain	15 % (5%) ≤ 3 km, 40 % (20%) 3–5 km (3–10 km) daytime (night-time)	Di Girolamo et al. (2009)

major contributions to the signal delay: the ionosphere and the neutral atmosphere. The ionospheric delay can be estimated by comparing two GPS signals at different frequencies (dispersion). The remaining part of the delay is due to the neutral, moist atmosphere, which refracts incoming electromagnetic waves, increasing the travel time of GPS signals (Solheim et al., 1999).

The neutral atmosphere is nondispersive and GPS cannot provide any information to separate the impact of water vapour from the impact of the dry atmosphere. Therefore additional meteorological observations are required. Usually the pressure and temperature at the GPS receiver are observed. The signal delay due to the dry gases, that is all atmospheric gases without water vapour, can be estimated reliably using the pressure observation and certain empirical models. The remaining wet delay can be converted to the slant-integrated water vapour by using the temperature observation. In general, 40–50 observations along single paths within 15 min are combined and mapped to a representative estimate of IWV above the station.

The GeoForschungsZentrum Potsdam processes the data of approximately 300 German GPS stations operationally in near-real time and provides IWV estimates with a temporal resolution of 15 min and an accuracy of 1–2 kg m^{-2} (Dick et al., 2001; Gendt et al., 2004).

2.1.3 Sun photometer

The sun photometer (CE 318 N-EBS9, Cimel Electronique) measures the extinction of direct solar irradiance and sky radiance at nine wavelengths (340, 380, 440, 500, 675, 870, 937, 1020, and 1640 nm) fully automatically. Allowing for the extinction due to aerosols, the extinction due to the amount of water vapour in the line of sight to the sun, T_w,

can be derived from the extinction at 937 nm. The extinction can be described with the following equation:

$$T_w = \exp[-a(m \cdot \text{IWV})^b], \tag{1}$$

where a and b are constants and m is the relative optical air mass (Schmid et al., 2001). From this relationship, IWV can be derived with an accuracy of 10 % (Alexandrov et al., 2009).

The sun photometer at JOYCE is part of AERONET, meaning that data processing is performed by the National Aeronautics and Space Administration (NASA) (Dubovik et al., 2006). The data used within the present study are of quality level 1.0 and have a temporal resolution of about 10 min.

Since the sun photometer measures the direct sunlight, its IWV retrieval is limited to daytime and clear-sky conditions. Additionally, since the instrument tracks the sun, the retrieved IWV is not zenith viewing but along a slant path through the atmosphere. This implies that it samples a different atmospheric volume than the zenith-pointing instruments. An additional problem due to the changing viewing paths can occur when the sun photometer is measuring at low solar zenith angles in combination with high IWV values. This could lead to transmission approaching 0 (Ingold et al., 2000).

---

[1]The planetary boundary layer with an assumed height of 2 km contributes most to IWV. The GPS slants with the lowest angles (7°) leave the boundary layer at a distance of approximately 16 km from the GPS station, and the slants are on average azimuthally equally distributed. This leads to a spatial representativeness of 32 km.

2.1.4 Raman lidar

The BASILicata Raman lidar system (BASIL) from Scuola di Ingegneria, Università della Basilicata, is a Raman lidar operating in the ultraviolet band (Di Girolamo et al., 2009) deployed at JOYCE during HOPE. BASIL emits pulses at 355, 532, and 1064 nm simultaneously along zenith. The determination of the water vapour mixing ratio is based on the detection of the Raman backscatter signals from N_2 and H_2O molecules at 386 and 407 nm respectively. Considering the power ratio of the H_2O signal to the N_2 signal, all system-dependent parameters can be eliminated. The power ratio of the two signals has to be calibrated.

During HOPE the calibration was based on the use of clear-sky radiosoundings launched 3.9 km to the south-east (cf. Fig. 1). The comparison between the lidar power ratio and the radiosonde mixing ratio profiles for the purposes of the calibration is typically carried out in the vertical region 2.5–3.5 km. Considering this altitude region above the boundary layer minimizes the air mass differences related to the distance between the lidar and the radiosonde station and allows minimization of the effects associated with the lidar overlap function.

Due to missing overlap near the instrument, the lowest usable signal from BASIL is from a height of 150–180 m above ground. Above this height, water vapour profiles with a vertical resolution of 30 m are provided every 5 min up to a height of approximately 3–8 km depending on day or night operation (max. time resolution 10 s). Due to its limited altitude coverage no column water vapour can be provided from BASIL measurements alone. Additionally, the Raman lidar is not able to measure in and above clouds because its signal is rapidly extinguished. Due to incomplete profile information, IWV cannot be derived by BASIL measurements without the use of complementing measurements from other instruments.

2.1.5 Radiosondes

During HOPE, 226 radiosoundings were performed with Graw DFM-09 sondes. These feature a thin-film capacitance sensor in order to measure relative humidity. Together with the temperature measurements and the pressure profile derived from GPS measurements, the absolute humidity is computed. Afterwards, the absolute humidity is integrated to derive IWV from the radiosoundings.

Many studies asses the error of radiosonde measurements. They show that the error strongly depends on the type of radiosonde (Nash et al., 2010). Furthermore, the systematic and random errors of the relative humidity sensor depend on temperature and differ between day- and night-time. A comparison to IWV derived from GPS showed that the difference between Graw DFM-09 and GPS is $2 \, kg \, m^{-2}$ higher during daytime than during night-time. Other radiosonde types showed the opposite behaviour. The reason for this could be that the correction algorithm applied by the Graw software

probably overcorrects the original dry bias. In general, IWV comparisons of radiosondes with capacitance sensors to GPS measurements show a dry bias for the radiosondes of approximately $1.2 \, kg \, m^{-2}$ during daytime due to sensor exposition to solar radiation (Wang and Zhang, 2008).

Note that the drift of radiosondes during ascent is not negligible: at 850 hPa the HOPE radiosondes drift on average 5 and 8 km at their maximums, and at 200 hPa the distance is on average 39 and 106 km at their maximums. Therefore, it has to be kept in mind that a radiosonde may well be in a different air mass than the zenith-pointing ground-based instruments are sampling. However, IWV variability is low above the boundary layer because the flow is determined by large-scale advection and therefore homogeneity is high (Shao et al., 2013). Therefore, IWV from radiosondes is nevertheless included in our multi-instrumental comparison.

2.1.6 MODIS

The Moderate Resolution Imaging Spectroradiometer (MODIS) is a space-borne, passive, whisk-broom scanning radiometer which measures the radiation backscattered and emitted from Earth, clouds, and atmosphere at 36 spectral bands between 0.4 and 14.4 µm wavelength. Two MODIS instruments are currently operational in space, on board NASA's sun-synchronous near-polar-orbiting Earth Observing System Terra and Aqua platforms (http://modis.gsfc.nasa.gov/). This enables a full global coverage every 1 to 2 days. With an orbit height of approximately 705 km and a scanning pattern of 55°, the swath dimension of MODIS amounts to 2330 km across track and 10 km along track (at nadir).

Two standard IWV retrievals exist for MODIS: the infrared retrieval (MODIS-IR) and the near-infrared retrieval (MODIS-NIR). Within the present study, MODIS Level 2 MODIS-IR and MODIS-NIR products from Collection 5.1 are used, which have a grid resolution of 3 and 1 km respectively (http://modis.gsfc.nasa.gov/data/).

MODIS-NIR utilizes three channels located within the water vapour absorption wavelengths, namely 0.905, 0.936, and 0.94 µm, and two nonabsorbing channels, namely 0.865 and 1.24 µm. The ratios in reflected NIR radiation from water vapour absorption channels to window channels give the atmospheric water vapour transmittances. From these, IWV is obtained from lookup tables based on line-by-line calculations. Note that single and multiple scattering effects are assumed to be negligible. The estimated errors in retrieved IWV are typically 5–10 % and are mostly assigned to uncertainties in the spectral reflectance of the surface targets and in uncertainties in the amount of haze over dark surfaces. For details on the MODIS-NIR retrieval see Gao and Kaufmann (2003).

MODIS-IR utilizes two water vapour absorption bands which deliver information on the moisture distribution and three window bands which also have weak water vapour ab-

sorption. From the radiances measured at these bands, water vapour profiles are retrieved via a statistical regression algorithm based on previously determined relationships between radiances and water vapour profiles. Though computationally efficient, this algorithm is sometimes unphysical. Therefore, a nonlinear iterative physical algorithm is applied to the retrieved profiles, aiming to improve the solution, that is reduce the known overestimation of IWV. For details on the MODIS-IR retrieval see Seemann et al. (2003).

Being based on thermal radiation, MODIS-IR is available for both day- and night-time over ocean and land. However, it is limited to clear-sky situations. The same goes for MODIS-NIR, which is additionally restricted to daytime and highly reflective surfaces, which means land and not ocean. Both MODIS retrievals, if applied to overcast scenes, miss information from within and below clouds.

2.2 Models

2.2.1 ICON high-resolution simulation

The ICOsahedral Nonhydrostatic (Zängl et al., 2014) modelling framework is currently being developed jointly by DWD and the Max Planck Institute for Meteorology (MPI-M) as the next generation NWP and climate model. The dynamical core is formulated on an icosahedral-triangular Arakawa C grid (Arakawa and Lamb, 1977). Within the HD(CP)^2 project, ICON is extended to perform high-resolution regional simulations.

For the presented case study in Sect. 3, ICON is run in limited area mode with a horizontal resolution of 156 m on a circular domain with a diameter of 265 km centred in Cologne (50°56′33″ N, 6°57′32″ E). Fifty generalized terrain-following levels are used in the vertical with a model top at 21 km and reduced level spacings in the lower troposphere. Distances between layer midpoints range from 30 m between the lowest levels to 1170 m between the top levels. The simulation is initialized and nudged hourly on the lateral boundaries with COSMO-DE analysis. In contrast to COSMO-DE, a higher-resolution topography data set is used when generating the lower boundary conditions. High-frequency output is stored at 40 grid points arranged radially around JOYCE with 1 km spacing (cf. Fig. 1) every 135 s.

2.2.2 COSMO-DE

COSMO-DE (Baldauf et al., 2011), an application of the Consortium for Small-scale Modelling (COSMO) covering Germany and its neighbouring countries, is the operational regional NWP model of Germany's National Meteorological Service, the DWD. It is a nonhydrostatic, fully compressible model of the atmosphere. The thermohydrodynamical equations describing compressible flow in a moist atmosphere are solved using a finite-difference method on an Arakawa C grid (Arakawa and Lamb, 1977). As for the coordinates, the model uses rotated latitude/longitude coordinates in the horizontal and time-independent terrain-following coordinates in the vertical. The horizontal resolution is 2.8 km and the vertical spacing of the 50 hybrid levels ranges from approximately 20 m at the Earth's surface to 1000 m at 22 km height.

Operationally, 21 h forecasts with COSMO-DE are initialized every 3 h from new analysis and are nudged hourly on the domain boundaries with 3 h old COSMO-EU forecasts, which is a coarser-resolved application of the same model covering Europe. Latent heat nudging towards radar data is applied during the first 30 min of each forecast. COSMO-DE output is available every 15 min.

2.3 Matching the data

In the following, the spatial matching of all data sets is addressed first, and the temporal matching is addressed in the final section. All JOYCE instruments are located within a distance of 110 m to each other. The GPS receiver and sun photometer are situated on the same roof of a building at a height of 111 m above mean sea level, while the MWR and BASIL are located on the ground. The height difference to the instruments on the roof is 21 m and therefore the MWR IWV needs to be corrected. For this, the 120 m meteorological tower nearby is used to adjust the IWV of the MWR to the level of GPS and sun photometer from the water vapour density measured in heights of 2, 10, and 20 m above ground. The amount of water vapour subtracted from the MWR measurements is 0.3 kg m^{-2} at its maximum. BASIL data are not height corrected since only the profiles and not IWV are used.

The location of the radiosonde launches is at exactly the same height as JOYCE at a distance of 3.9 km to the southeast. The second MWR used in Sect. 3 is 3.3 km south of JOYCE (cf. Fig. 1). For MODIS, the horizontal and height distance to JOYCE varies with flight track. The topography of the MODIS measurements is taken from the Consultative Group on International Agricultural Research-Consortium for Spatial Information Shuttle Radar Topography Mission (CGIAR-CSI SRTM) 90 m database (http://srtm.csi.cgiar.org). The topography of the nine nearest CGIAR-CSI SRTM pixels is averaged to retrieve the height of the MODIS pixel. The nearest MODIS pixel within a distance of less than 7 km and a height difference of less than 100 m is used. To correct for the height difference the water vapour density of the meteorological tower is again used, resulting in a maximum correction of 1.5 kg m^{-2}.

The grid point of COSMO-DE used in the present study is with a distance of 1.9 km the second nearest grid point to JOYCE (cf. Fig. 1). This grid point is selected because it is only 1 m lower than the JOYCE site, whereas the nearest grid point in a distance of 1.8 km has a height difference of 10 m. Due to the small height difference, no height correction is applied to the IWV from COSMO-DE.

For ICON, no height correction is applied. The height difference between the ICON grid point used for Fig. 2 is only 4 m, so the bias introduced by this height difference is very small.

Apart from the spatial differences, the temporal differences need to be considered. If not stated otherwise, the resolution of compared IWV values is 15 min. GPS measurements are originally available in this resolution. The output of COSMO-DE is also available with a resolution of 15 min. MWR and sun photometer measurements are averaged over 15 min. IWV from the other measurements is available only with a coarser temporal resolution. MODIS measurements are matched to the corresponding 15 min period. The ascent of a radiosonde takes approximately 1 h. Since the largest amount of water vapour is in the lower atmosphere, the radiosoundings are matched to the 15 min interval during which they are started. This results in a maximum time difference of less than 15 min between two individual measurements of different instruments.

3 Case study

The capabilities and limitations of the different techniques to measure IWV are exemplarily demonstrated for a case study with fair weather conditions on 5 May 2013, when a high-pressure system dominated the synoptic situation over western Germany. The day was characterized by a classical development of the atmospheric boundary layer. Approximately 2 h after sunrise the convective mixing layer (ML) started to form and completely replaced the residual layer (e.g. Stull, 1988) of the last day around 08:00 UTC (cf. top panel in Fig. 2) as indicated by the ML height (MLH) derived from JOYCE wind lidar measurements (Schween et al., 2014). After 08:00 UTC, when the ML is fully developed, the vertically resolved BASIL measurements reveal the strong water vapour gradient between the moist ML and the dry free troposphere above. Even though the ML does not extend higher than approximately 1.5 km on this day, it contains roughly 50 % of the total IWV (estimated from radiosondes). Furthermore, the ML is characterized by a strong temporal water vapour variability as clearly seen from BASIL measurements.

Clear-sky conditions prevailed until 09:00 UTC. Later, occasional cumulus evolved which did not significantly limit the BASIL lidar observations (cf. top panel in Fig. 2).

The MODIS overflight at 10:25 UTC (cf. middle panel in Fig. 3) shows the high spatial IWV variability with values between 13 and 16 kg m^{-2} around JOYCE. In general, the north- and south-west of the domain is drier by up to 3 kg m^{-2} than the rest of the domain. Note that this MODIS map, in contrast to the MODIS data included in the statistics and the time series, is not height corrected. For this reason, the open-pit lignite mining site, which is up to 400 m lower than JOYCE (cf. Fig. 1), is recognizable on the MODIS map

by the larger IWV values next to the radiosonde station (approximately 2 kg m^{-2} higher than the rest of the domain). Note also that those areas identified as cloudy by the MODIS cloud mask are displayed in white, since the IWV would only include water vapour up to cloud top. Still, some pixels stand out for their low IWV values in comparison to the surrounding IWV values, which might indicate that some clouds may not have been detected by the MODIS cloud mask.

The time series of IWV from all instruments and the two models (cf. middle panel in Fig. 2) shows that during this day, IWV varies between 12.5 and 18 kg m^{-2}. The lowest value can be observed around 07:00 UTC when the residual layer collapses. Afterwards, IWV gradually increases during the course of the day, corresponding in part to the increase in MLH as seen from BASIL (cf. top panel in Fig. 2). Clearly, the ML development is also associated with both high fluctuations in the water vapour mixing ratio visible in BASIL measurements and high IWV fluctuations visible in the temporally highly resolved MWR observations (5 s) and to a similar degree in the ICON simulations (135 s). The amplitude of these fluctuations exceeds the noise level of the MWR (0.05 kg m^{-2}), indicating that these fluctuations are due to true atmospheric variations. The diurnal development of the standard deviation of IWV over 1 h further confirms this feature (cf. bottom panel in Fig. 2). Due to the lower temporal and/or spatial resolution, the other observations and the COSMO-DE simulation cannot reproduce these fluctuations. However, as mentioned above they are identified by BASIL to be caused by ML dynamics (cf. top panel in Fig. 2).

3.1 IWV intercomparison

Several features can be identified in the comparison of the time series of the different IWV data sets (cf. middle panel in Fig. 2). They are described in this section.

Only GPS and MWR provide continuous observations over the full day. Though they overlap within their uncertainty estimates, GPS measurements tend to lie below the MWR measurements. The GPS measurements exhibit two distinct features: firstly, they show a jump at the beginning of most full hours, which can be up to nearly 1 kg m^{-2}. These jumps are caused by the near-real-time processing routine of the GPS retrieval at GFZ (Gendt et al., 2004). Secondly, an even larger difference (ca. 5 kg m^{-2}) is seen at the end of the day from 23:45 to 24:00 UTC.

These two issues occur in nearly all cases investigated so far and are not limited to the case selected for the present study. First attempts in reprocessing the data resulted in a smoothing of the hourly jumps and a reduction of the differences at the beginning of the day. However, the bias of the reprocessed data is increased. Therefore, the reprocessing is under further investigation.

During daytime, when IWV is available from the sun photometer, its IWV agrees very well with the MWR. However,
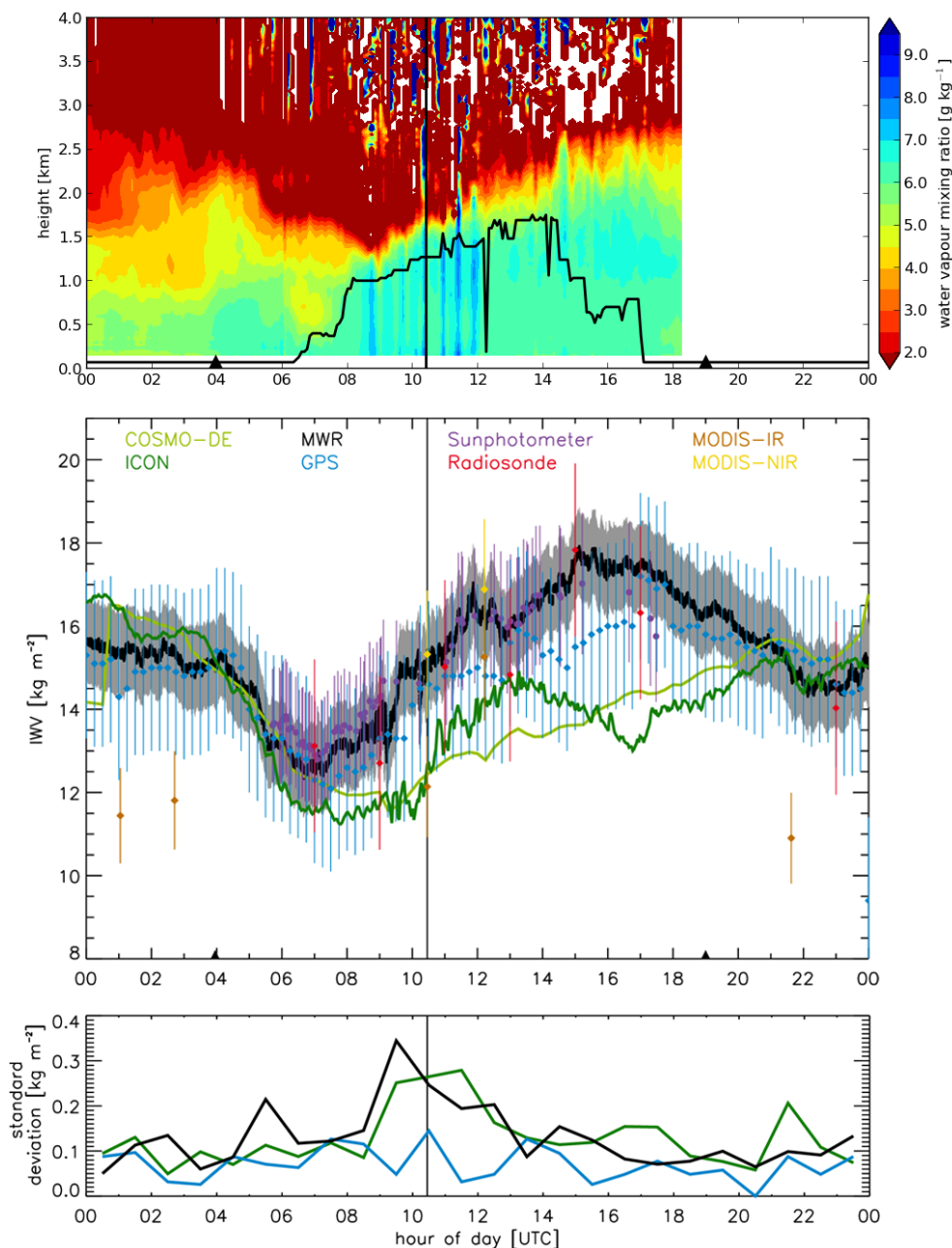
Figure 2. Time series for 5 May 2013 at JOYCE. Triangles indicate sunrise and sunset. The vertical black line indicates a MODIS overflight (cf. Fig. 3). Top panel: vertically resolved water vapour from Raman lidar BASIL for 5 May 2013 at JOYCE (colours) with ML height derived from wind lidar (black line). Middle panel: all IWV measurements and their corresponding uncertainties (cf. Table 1) together with the model simulations. Grey shading represents MWR uncertainty. Bottom panel: trend-reduced standard deviation within 1 h intervals. Line colours correspond to those in the middle panel.

the agreement is reduced during the early and late hours of daytime when the sun is at low elevation (cf. Sect. 4.3).

The MODIS-NIR estimates available for the two overpasses are perfectly within the uncertainty range of MWR and sun photometer, while MODIS-IR measurements also available during night-time are up to $4 \, \text{kg m}^{-2}$ too dry. The larger pixels of MODIS-IR (3 km) could partly be covered by

clouds which are not detected. The smaller pixels of MODIS-NIR (1 km) are less likely to be partly cloudy, which could lead to a more precise cloud detection.

The seven radiosondes which were launched during this day give IWV within the uncertainty range of the MWR, sun photometer, and/or GPS. The daytime soundings show that roughly 50 % (maximum 64 %) of the IWV is contained in
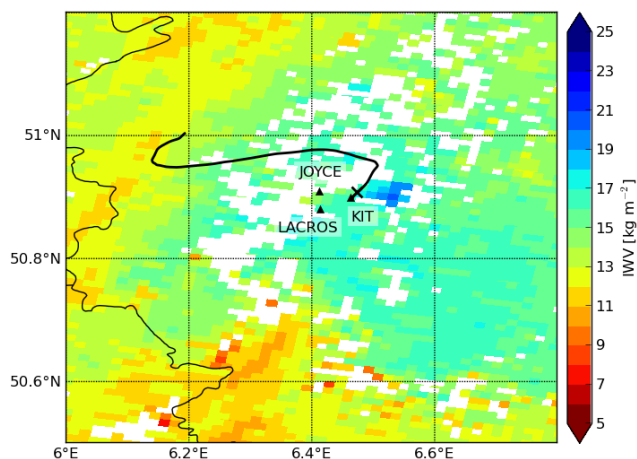
Figure 3. MODIS-NIR IWV for 5 May 2013 at 10:25 UTC. Cloudy pixels are displayed in white. The black line indicates the track of the radiosonde launched at 11:00 UTC with a cross at the location where it leaves the planetary boundary layer.

Figure 4. Correlation coefficients (left) and standard deviations (right) of IWV from ICON grid points (simulation for 5 May 2013) as a function of temporal and spatial distance. The circles represent the correlation coefficients and standard deviations from two MWRs positioned 3.3 km apart (cf. Fig.1).

the convective ML. Since the radiosonde provides point measurements along its trajectory, deviations from true-zenith measurements can occur due to sampling issues. For this case study, the horizontal drift within the ML is relatively short with approximately 4 km for the sonde launched closest to the MODIS overpass at 11:00 UTC (cf. Fig. 3). However, on this day, which does not feature a larger synoptic IWV gradient in the vicinity of JOYCE, it can be expected that differences to true-zenith estimates arise when the radiosonde is moving within dry/moist eddies in the convective ML.

The IWV simulations by the dynamic models COSMO-DE and ICON agree well with the observations until the 06:00 UTC when the increase in IWV can not be reproduced as strongly as in the observations. This might be due to problems in the forcing at the model boundaries – in particular for the ICON model, which is forced by COSMO-DE. Nevertheless, it is encouraging to see that the novel high-resolution ICON depicts a similar temporal IWV variability during the development of the convective ML as MWR and BASIL. This gives us the confidence that the model is suitable to further investigate the spatial and the temporal variability of IWV.

3.2 Assessment of representativeness

While all measurements have sampling issues, the use of a dynamic atmospheric model allows sampling of IWV nearly continuously in both time and space. Here we selected 40 ICON grid points (cf. Fig. 1) in an area of approximately 10 km around JOYCE for which IWV output was stored at high temporal resolution (2.25 min). The height above mean sea level of the sampled grid points ($dx = 156$ m) does not vary by more than 150 m.
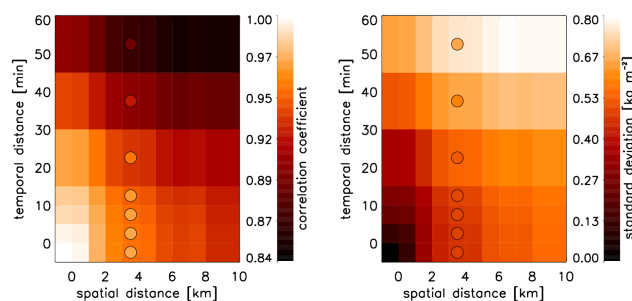
From the time series at the 40 grid points, the IWV correlations and standard deviations for distances smaller than 10 km and shorter than 1 h can be assessed (cf. Fig. 4). The correlation decreases distinctly with both temporal and spatial mismatch. For a fixed time a distance of 10 km reduces the correlation to 0.9. A similar decrease in correlation occurs when the location is fixed but a time mismatch of 30 min occurs. A mismatch of 10 km and 1 h leads to a correlation of 0.8.

A similar behaviour as for the correlation is evident in the standard deviation. Observations with a distance of 8–10 km can induce the same error as a time shift of 30–45 min ($0.6\,\mathrm{kg\,m^{-2}}$) that is around the specified uncertainty of the different observations (cf. Table 1). The combination of temporal and spatial mismatch, which is the case for radiosondes, can lead to even higher errors amounting to more than $0.8\,\mathrm{kg\,m^{-2}}$ for 10 km and 1 h difference.

In order to investigate whether the model behaviour is consistent with the observations, we use time series of 2.25 min IWV averages from two zenith-pointing MWRs located 3.3 km apart from each other. Both correlation and standard deviation decrease similarly as depicted by ICON (cf. Fig. 4). Interestingly, there are slight differences in the absolute values. Nevertheless, the comparison indicates that ICON simulations can be used to assess the small-scale variability of water vapour and help to determine to which degree instrument intercomparisons may be affected by atmospheric effects. However, it is important to note that this is a case study and the results might be rather different for different synoptic situations or geographic regions.

4 Statistical assessment

The 2 months of HOPE provide the opportunity to investigate IWV characteristics over a wide range of atmospheric conditions for a typical continental, mid-latitude site. The period was characterized by dry polar air masses at the beginning of April that transitioned into a strong synoptically
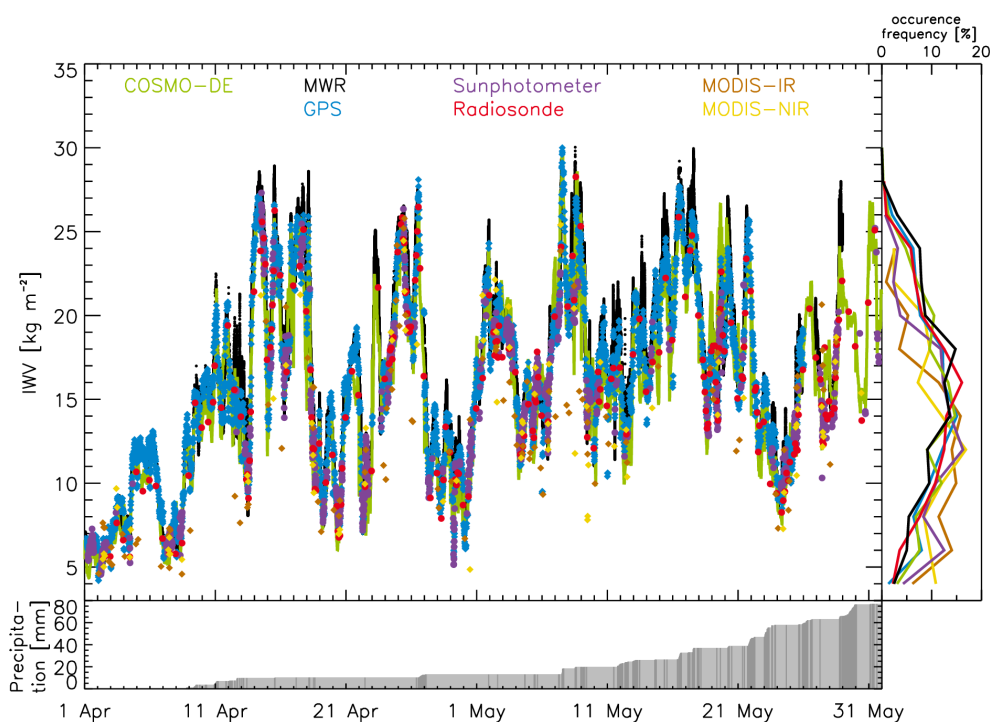
Figure 5. Time series of IWV during HOPE. Displayed are: MWR (black), GPS (blue), sun photometer (purple), radiosoundings (red), MODIS-IR (orange), MODIS-NIR (yellow), and COSMO-DE (light green). The frequency of occurrence of IWV are displayed in the right panel with corresponding colours. Accumulated precipitation is shaded in grey in the lower panel; dark grey bars indicate the time when precipitation fell.

forced regime in mid April with frequent passages of frontal systems over JOYCE during May. There were only a few rain events in April but more in May, which accumulate to 77 mm of total precipitation over the 2 months (cf. bottom panel in Fig. 5).

In this period IWV varies by $25\,\mathrm{kg\,m^{-2}}$, namely between 5 and $30\,\mathrm{kg\,m^{-2}}$ (cf. main panel in Fig. 5). IWV can increase or decrease by 10–$20\,\mathrm{kg\,m^{-2}}$ within 1 to 2 days. The different IWV data sets reveal a broad frequency distribution with a maximum around $15\,\mathrm{kg\,m^{-2}}$ (cf. right panel in Fig. 5). This distribution reveals the influence of the instrument sampling: GPS, MWR, radiosondes, and COSMO-DE show rather similar characteristics. In contrast, the distribution for the sun photometer is shifted to lower IWV values as it is restricted to daytime clear-sky measurements.

In the following, we first investigate the instrument performance during the whole period of HOPE before we analyse whether the small-scale temporal IWV variability (< 1 h) revealed in the case study is typical for the complete HOPE period.

4.1 Instrument intercomparison

Since none of the instruments can be considered as "the truth", every instrument is compared to all other instruments (cf. Fig. 6). All measurements are considered at 15 min res-

olution (see Sect. 2.3). For the following comparison, it has to be acknowledged that the maximum distance between instruments is approximately 4 km.

For the MODIS–radiosondes comparison, too few coincident measurements are available due to the infrequent satellite overflights. Excluding MODIS, the overall agreement between the instrument pairs is good. The standard deviation is not higher than $1\,\mathrm{kg\,m^{-2}}$ and the correlation coefficient is never lower than 0.98. The absolute bias varies from 0 for GPS–sun photometer to $1\,\mathrm{kg\,m^{-2}}$ for radiosondes–MWR. In the following, the individual instrument comparisons are examined in more detail.

With more than 3800 measurements, the GPS–MWR comparison includes the most cases because both instruments also measure during cloudy conditions. The bias ($0.2\,\mathrm{kg\,m^{-2}}$) is very low and the standard deviation ($0.9\,\mathrm{kg\,m^{-2}}$) is within the expected measurement uncertainty (cf. Table 1). However, there are some IWV values up to $5\,\mathrm{kg\,m^{-2}}$ lower than observed by the MWR (cf. Fig. 6). These differences occur due to problems in the processing of the GPS data at the beginning of the day, as mentioned above. Excluding the first hour of the day leads to a reduction of the bias to $0.1\,\mathrm{kg\,m^{-2}}$ and of the standard deviation to $0.8\,\mathrm{kg\,m^{-2}}$. This problem is further investigated in Sect. 4.3. Furthermore, a small dependency of the error on the IWV is found. For large IWV values the difference between GPS
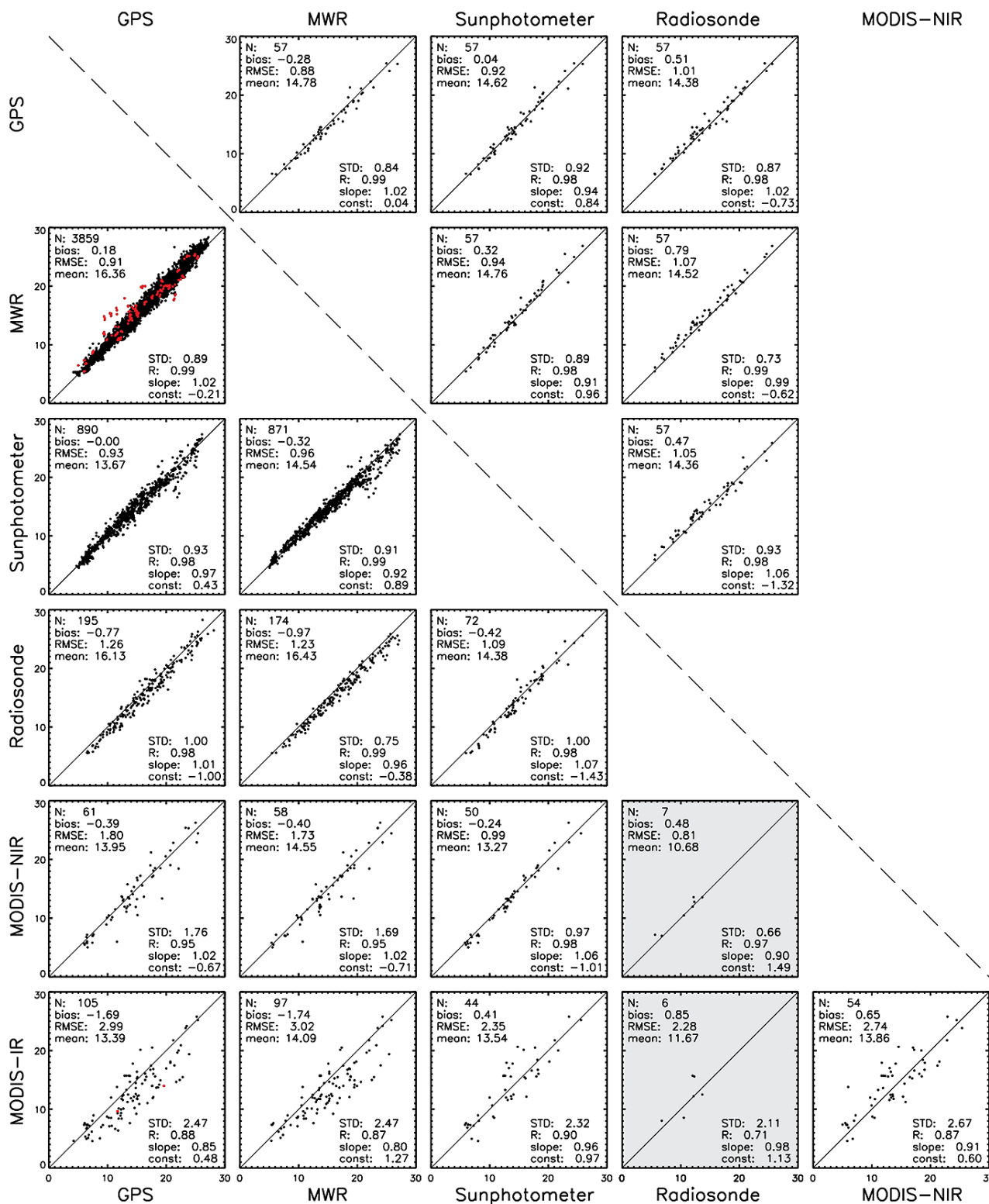
Figure 6. Scatterplots of IWV for all instruments against each other. Included are the number of measurements (N), bias (row–column in kg m^{-2}), root mean square error (RMSE in kg m^{-2}), mean (in kg m^{-2}), standard deviation (STD in kg m^{-2}), Pearson correlation coefficient (R), and slope and y intercept (const in kg m^{-2}) of linear regression. The lower left half of the figure shows comparisons when the two instruments measure. The upper right half shows comparisons when all instruments measure. MODIS is not included in the upper half due to less measurements. The GPS measurements between 00:00 and 01:00 UTC are highlighted in red.
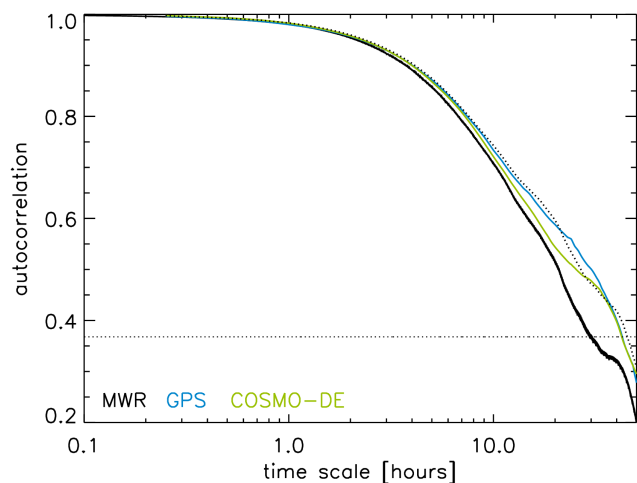
Figure 7. Autocorrelation of IWV during HOPE measured with MWR with 5 s resolution (solid black), with 15 min resolution (dotted black), with GPS (solid blue), and simulated with COSMO-DE (green). The horizontal line represents e^{-1}.

and MWR tends to be smaller than for small IWV values. Other dependencies, such as the influence of wind direction, spatial IWV gradient, temporal IWV variability, liquid water path, and distribution of GPS slants, which are used to retrieve the IWV, are tested but no significant dependency is found (not shown).

Even if the sun photometer only measures during clear-sky, daytime situations there are nearly 900 coincident measurements with GPS and MWR. To both the sun photometer shows nearly no bias (GPS < 0.01 kg m^{-2}, MWR: -0.35 kg m^{-2}) and standard deviations of 0.8 kg m^{-2}. The AERONET cloud screening and quality assurancce conducted for level 2.0 sun photometer data reduces the data set by a factor of 2, while the bias hardly changes (< 0.1 kg^{-2}) and standard deviation is reduced by 0.1–0.2 kg m^{-2} (not shown).

On average, the radiosondes are 0.8 kg m^{-2} (1.0 kg m^{-2}) drier than GPS (MWR). However, only a small difference of 0.2 kg m^{-2} between day- and night-time soundings could be identified, probably due to the correction within the Graw software (cf. Sect. 2.1.5).

The comparisons of MODIS–GPS and MODIS–MWR show that IWV measurements from both MODIS-IR and MODIS-NIR are frequently too low. However, these MODIS measurements are not included in the MODIS–sun photometer comparisons, since there are no sun photometer measurements at these times. The reason for this is probably that cloudy cases are not reliably detected by the MODIS cloud-identifying algorithm. Clouds lead to a lower IWV because the amount of IWV below and inside the cloud is not detected by MODIS. A clear difference can be seen in the standard deviation in the comparisons involving MODIS-NIR and MODIS-IR; the latter has more than double the standard

deviation of the first, which could be due to the coarser resolution or due to poorer physical constraints in the algorithm.

Since each instrument intercomparison is carried out during different atmospheric conditions (a consequence of the varying instrument limitations), the mean IWV of the measurements included in each comparison differs by approximately 3 kg m^{-2}. To allow for a better comparison of the errors of different instrument combinations, 57 simultaneous measurements of all instruments with the exception of MODIS are also investigated separately. The mean of these comparison then only differs by 0.4 kg m^{-2} (cf. Fig. 6) and the standard deviation is reduced for all instrument combinations to be lower than 1 kg m^{-2}. This results likely from sampling more homogeneous conditions. By including only measurements when the sun photometer is measuring, night-time measurements and most importantly all rainy cases and cases with clouds in the direction of the sun are excluded.

In summary, the agreement of the IWV measurements on the 15 min basis is very good with standard deviations of around 1 kg m^{-2} with the exception of MODIS. However, it has to be kept in mind that the representative error of IWV at 4 km spatial distance is only 0.4 kg m^{-2}. The representativeness analysis for 5 May 2013 estimated the effect of atmospheric variation to be approximately 0.4 kg m^{-2} (cf. Sect. 3.2). As expected, a reduction of the compared data sets by only including coincident measurements simultaneously excluding all night-time, rainy, and cloudy cases, leads to an improvement in the overall agreement. However, the mean values over the HOPE period range from around 16 kg m^{-2} (GPS, MWR) to lower than 14 kg m^{-2} (sun photometer, MODIS). This difference, which is distinctly higher than the bias of most of the instrument comparisons, implies significant errors when climatologies are constructed from data sets with a poor sampling.

4.2 Temporal variability

Having assured the good general agreement between the different instruments during HOPE, the temporal variability of IWV is investigated in more detail in the following. For this, the autocorrelation of the continuous data sets, namely MWR, GPS, and COSMO-DE, is computed (cf. Fig. 7). All three data sets with a temporal resolution of 15 min show a similar behaviour: their autocorrelation function decreases monotonically with increasing lag time and they have a similar e folding time of roughly 13 h. This result is not surprising considering the large IWV changes associated with the synoptic variability (cf. Fig. 5), but it gives important limitations on the influence of temporal matching in IWV comparisons and on generations of climate data records. Interestingly, the e folding time decreases to 12 h when MWR measurements with higher resolution, that is 5 s, are used, indicating the importance of small-scale processes.

For a closer look at the variations due to small-scale processes, the IWV standard deviation during HOPE is com-
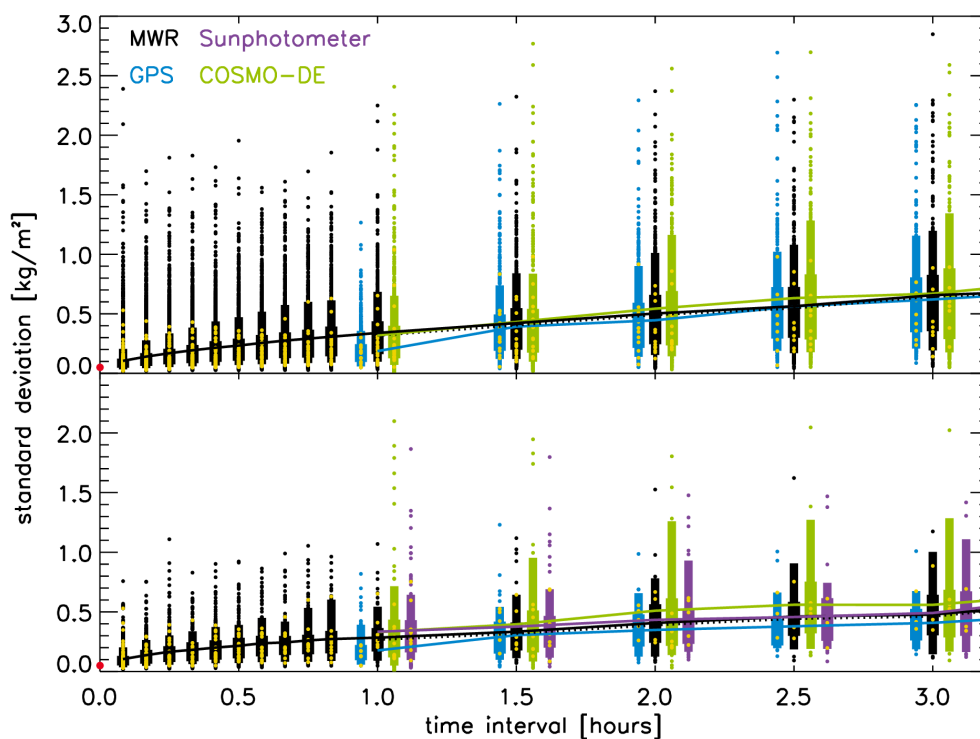
Figure 8. Lines: mean standard deviation of IWV during HOPE computed for varying intervals. Displayed are: MWR with 15 min resolution (dotted black), MWR with 5 s resolution (solid black), GPS (blue), and COSMO-DE (green). For the 5 s MWR measurements, the GPS measurements, and the COSMO-DE simulation the vertical bars indicate the 10, 25, 75, and 90 %-percentiles of the standard deviation. The single dots indicate the outliers. The data points from the case study (cf. Fig. 2) are given in yellow. The bottom panel additionally includes sun photometer data (purple) and is limited to coincident measurements during daytime clear-sky conditions. The red dot on the y axis represents the noise level of the MWR.

puted over varying time intervals from 5 min to 3 h (cf. top panel in Fig. 8). Note that only coincident measurements and simulations are used and only the MWR can provide estimates below 1 h. Generally, the mean standard deviation increases from $0.1\,\mathrm{kg\,m^{-2}}$ at 5 min to $0.4\,\mathrm{kg\,m^{-2}}$ at 1.5 h, showing some saturation with $0.6\,\mathrm{kg\,m^{-2}}$ at 3 h intervals.

For time intervals of 1.5 h and longer, MWR, GPS, and COSMO-DE again show a similar behaviour, as seen in the autocorrelation. In fact, they lie within their 25 and 75 % percentiles. However, extreme values reach standard deviation of $2.0\,\mathrm{kg\,m^{-2}}$ and higher at time intervals > 1 h. Interestingly, none of these points are evident during the day of the case study (cf. Sect. 3) because the highest standard deviations stem from cloudy situations (see discussion below).

The GPS measurements show an offset for the 1 h interval. This is caused by the processing method. As seen in the middle panel of Fig. 2, GPS measurements within 1 h are relatively smooth. However, the mean standard deviation of the 15 min MWR averages is overall only slightly smaller than the mean standard deviation of the 5 s averages. This indicates, firstly, that for time scales of a few hours, the coarser resolution of 15 min is sufficient enough for resolving the mean IWV variability. Secondly, for these time intervals GPS

is well suited as a reference instrument for model evaluation since it captures the same variability as the MWR. Thirdly, operational NWP model COSMO-DE is capable of reproducing the observed mean variability of IWV.

For time intervals shorter than 1 h, only the 5 s MWR data can partially resolve the small-scale, turbulence-induced variability of IWV. The minimum detected average standard deviation at 5 min averaging time of $0.1\,\mathrm{kg\,m^{-2}}$ is twice as high as the MWR noise level and thus represents a lower boundary for the evaluation MWR measurement. As for the variability on intervals greater than 1 h, the standard deviation increases with increasing time interval; however, the slope is steeper on the shorter time scales. At the shortest time scales, the variability is dominated by a cascade of turbulence elements in the inertial subrange, whereas at increasing time scales the variability is probably dominated by the variability of subsequent updraught and downdraught regions. Noteworthy are also standard deviation values larger than $1\,\mathrm{kg\,m^{-2}}$ even at the shortest time scales, which are predominantly caused by clouds.

Focusing on clear-sky, daytime cases allows to include the sun photometer (cf. bottom panel in Fig. 8). When only coincident data from MWR, GPS, sun photometer, and COSMO-
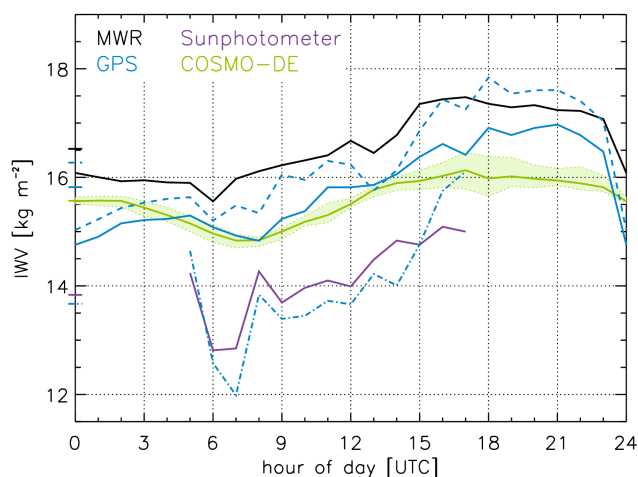
Figure 9. Mean daily cycle of IWV during HOPE measured with MWR with 15 min resolution (black), GPS (solid blue), GPS for coincident measurements with MWR (dashed blue), GPS for coincident measurements with sun photometer (dash-dotted blue), sun photometer (purple), and simulated with COSMO-DE (green). The shaded green area represents the spread of differently aged forecasts of COSMO-DE. The ticks on the y axis represent the respective 2-month means.

DE are used, the mean standard deviations are lower by approximately $0.25\,\mathrm{kg\,m^{-2}}$ compared to the full time series (cf. bottom panel in Fig. 8). This is caused by the exclusion of cloudy cases that lead to the disappearance of high standard deviations, that means hardly any standard deviations higher than $1\,\mathrm{kg\,m^{-2}}$ occur once (partially) cloudy scenes are filtered out. The IWV standard deviation observed during the case study seems to be representative of the whole HOPE campaign on time scales shorter than 1 h.

In summary, the change of the mean standard deviation with different time intervals, over which it is computed, shows that the variability of IWV is high even for time periods shorter than 1 h, which is mostly due to clouds, and that this variability cannot be resolved by more coarsely resolved data. High-resolution time series from MWR are therefore well suited to high-resolution atmospheric models, like ICON, aiming to derive better subgrid parametrizations for climate models. However, for more synoptic scale comparisons, a resolution of 15 min is sufficient to resolve the mean standard deviation and therewith variability of IWV.

4.3 Diurnal cycle

The previous sections show the importance of the IWV variability associated with atmospheric turbulence and convection. In this section we focus on the mean diurnal cycle of IWV over the HOPE campaign as this is strongly influenced by combined effects of land-surface processes and boundary layer dynamics. It represents an aggregated quantity that tests to which degree different instruments and/or models can pro-

vide a consistent answer. Only those measurements that are available on a near-continuous basis, that is MWR, GPS, and sun photometer, and COSMO-DE output are included in this comparison with 1 h means. Note that it is ensured (by manual checking) that this daily cycle is not due to a few singular synoptic-induced events but is rather a true mean behaviour of IWV.

Figure 9 reveals a clear mean daily IWV cycle over the HOPE period with lowest values in the morning and maximum in the afternoon/evening hours. The daily IWV range varies from 1 to $2\,\mathrm{kg\,m^{-2}}$ depending on the data set. This is significantly higher than the daily IWV range reported by Morland et al. (2009) for a 5-year data set from Bern, Switzerland ($0.6\,\mathrm{kg\,m^{-2}}$), and can be attributed to the comparably high surface fluxes during springtime.

As mentioned before, the mean IWV is instrument-dependent because of sampling issues, which leads to differences in the absolute values in the mean diurnal cycle and also to differences in the amplitude of the mean diurnal IWV cycle. The amplitude is smallest for the COSMO-DE forecasts ($1.3\,\mathrm{kg\,m^{-2}}$) that are here represented by the ensemble of differently aged forecasts (cf. Sect. 2.2.2). Interestingly, the spread between the different ensemble members is highest around the time of maximum IWV ($\sim 17{:}00$ UTC). Since there is interaction between humidity, time, and strength of convection, and resulting precipitation (Schlemmer et al., 2011) this might be associated with difficulties of the forecast model with convective precipitation.

The GPS is the only instrument that provides data under all weather conditions and can directly be compared to the COSMO-DE output. With $2.2\,\mathrm{kg\,m^{-2}}$ GPS shows a stronger diurnal cycle than COSMO-DE, with the maximum IWV occurring also 4 h later around 21:00 UTC. The later maximum of IWV in the GPS might be due to the use of surface temperatures in the GPS retrievals as these are not representative of the atmospheric temperature as found by Morland et al. (2009). They apply a dampened mean atmospheric temperature to compensate for this surface effect, which leads to a better agreement of the diurnal cycle with coincident MWR measurements.

The high IWV range of GPS measurements might partly be caused by a dry offset of approximately $1\,\mathrm{kg\,m^{-2}}$ in the beginning of the day compared to the end of day. This is a known characteristic of the near-real-time processing of GPS data, which is also seen in the investigation of the daily cycle at stations in North America by Dai (2002). The exact reason for this feature is not finally clarified yet and subject of ongoing investigation.

The MWR IWV exhibits a similar shape of the diurnal cycle as GPS and COSMO-DE though the time of the maximum IWV is earliest in the MWR around 15:00 UTC and its IWV range ($1.9\,\mathrm{kg\,m^{-2}}$) is between COSMO-DE and GPS. However, it needs to be considered that the outdoor MWR HATPRO cannot measure during rain and therefore a fair comparison can only be guaranteed if GPS data are filtered

accordingly. While such a filtering gives a similar bias as the analysis shown in Fig. 6, the mean diurnal variation of the GPS ($2.8\,\mathrm{kg\,m^{-2}}$) is clearly $1\,\mathrm{kg\,m^{-2}}$ larger than that of the MWR.

Due to its measurement principle, the sun photometer (cf. Sect. 2.1.3) is limited to clear-sky conditions from 5 to 17:00 UTC, resulting in the lowest IWV values of all data sets. Nevertheless, an increase in IWV during daytime with an even stronger slope than for the other data sets can be seen. These measurements show the same trend of smaller IWV values in the morning than in the afternoon. The diurnal cycle of coincident GPS measurements shows a good agreement with the sun photometer measurements. For the difference between the sun photometer and MWR, a dependency on the position of the sun is found (not shown). In the morning and afternoon, IWV from the sun photometer is smaller than from the MWR because the sun photometer measures under lower-elevation angles. At noon it is the other way around. This could be due to an inaccurate relative air mass (Eq. 1) used by the retrieval or the transmission approaching 0 at low elevation angles.

In summary, the accurate description of the mean diurnal cycle is strongly limited by instrumental and sampling effects requiring an accurate matching when different data sets are compared. Longer time series are desirable. Nevertheless, the results indicate that the operational COSMO-DE model underestimates the amplitude of the diurnal cycle.

5 Summary and conclusions

The present study uses multi-instrumental observations and model simulations of IWV at the mid-latitude site JOYCE (Löhnert et al., 2014) to investigate its spatial–temporal variability. The – to our knowledge – unprecedented set of instruments (MWR, GPS, sun photometer, radiosondes, Raman lidar, MODIS-IR, MODIS-NIR) located in close proximity during the 2 months of the HOPE campaign (http://hdcp2.zmaw.de/HOPE.2306.0.html) is complemented by a well-established operational NWP model (COSMO-DE) and – in the frame of a case study – the novel high-resolution atmospheric model ICON.

The different instruments have different sampling characteristics, uncertainties, and limitations (cf. Table 1) that are important to consider when assessing IWV variability. Most importantly, a height correction is necessary as an elevation difference of only 20 (100) m can introduce errors of 0.3 (1.5) $\mathrm{kg\,m^{-2}}$. Pairwise comparison of the IWV-measuring instruments with 15 min temporal resolution shows a generally good agreement over the whole HOPE period with a small standard deviation ($\leq 1\,\mathrm{kg\,m^{-2}}$) and a high correlation coefficient (≥ 0.98), with the exception of MODIS. The absolute bias varies from 0 to $0.97\,\mathrm{kg\,m^{-2}}$. IWV from MODIS is often lower than from the other instruments because cloudy pixels are most probably not always identified by the MODIS

cloud-detection algorithm. Nevertheless, MODIS is the only instrument capable of assessing the small-scale spatial structure of IWV – once corrected for elevation and filtered for clouds – over the whole globe.

The multi-instrumental intercomparison reveals a number of aspects for the individual instruments:

– sun photometer measurements show a good agreement with the other measurements but can only be conducted during clear skies in daytime and seem to suffer from problems when the sun is low

– IWV from MWR and GPS differs only slightly (bias: $0.2\,\mathrm{kg\,m^{-2}}$ (1 %), standard deviation: $0.9\,\mathrm{kg\,m^{-2}}$ (6 %), cf. Fig. 6), taking the specified instrument uncertainties into account

– near-real-time processed GPS data exhibit inconsistencies at the beginning of each day and each hour due to the processing procedure that might also lead to a shift in the diurnal cycle of IWV. Further work on the processing might increase the performance of the GPS measurements.

Despite the characteristics of the measurements themselves, other aspects have to be taken into account to judge the instruments. For example, a comprehensive GPS network exists, thus making GPS better suited to evaluate models over their whole domain.

The analysis of the temporal variability of IWV reveals three distinct sources:

– synoptic influence is mainly responsible for the fact that the e folding time of the autocorrelation is approximately half a day

– clouds and broken cloud fields can cause standard deviations of IWV of over $1.5\,\mathrm{kg\,m^{-2}}$ within time intervals of a few hours

– atmospheric turbulence determines IWV variability also in cloud-free conditions on scales below 1 h.

The high standard deviations during cloudy time periods do not occur when only daytime clear-sky IWV estimates are considered (cf. Fig. 8). Therefore, instrument intercomparisons under cloud-free conditions are advantageous to assure more homogeneous conditions. The high resolution (a few seconds) of the MWR enables to observe standard deviations higher than $0.5\,\mathrm{kg\,m^{-2}}$ for time intervals less than 30 min. This information is interesting for the development of subgrid parametrizations for atmospheric models but also implies that instrument intercomparisons should make use of suitable measures to identify atmospheric conditions with low variability in order to isolate instrument errors.

The standard deviation derived from high-resolution MWR time series is able to identify turbulent mixing within the growing ML, as demonstrated for a case study with

the help of vertically resolved water vapour and wind lidar data. For the same day, simulations at 156 m grid resolution with the novel ICON model were used to assess the spatio-temporal IWV correlation and standard deviation for time differences smaller than 1 h and shorter than 10 km. It is shown that a temporal mismatch of 30–45 min or a spatial mismatch of 8–10 km can already lead to a random error of $0.6\,\mathrm{kg\,m^{-2}}$. A combination of temporal and spatial mismatch introduces even higher errors. The results are confirmed from observations of two MWR operated 3.3 km apart.

An important aspect for climatological studies is that mean IWV over HOPE, as derived from the different sources, differs by up to $3\,\mathrm{kg\,m^{-2}}$ because different time periods are included in the measurements. These differences occur due to limitations of the measurement principles and measurement gaps of instruments. These differences introduce deviations in the statistics of the different instruments or models. Therefore, as done in this study, only coincident data should be compared. This is particularly true for the mean diurnal cycle over the whole campaign where our study reveals an underestimation of the amplitude by the operational COSMO-DE model. In the future, longer simulations with the novel ICON model, which are yet not possible due to limited computing power, will be performed to investigate whether the encouraging results from the case study presented here can be confirmed in more general terms.

References

Alexandrov, M. D., Schmid, B., Turner, D. D., Cairns, B., Oinas, V., Lacis, A. A., Gutman, S. I., Westwater, E. R., Smirnov, A., and Eilers, J.: Columnar water vapor retrievals from multifilter rotating shadowband radiometer data, J. Geophys. Res., 114, D02306, doi:10.1029/2008JD010543, 2009.

Arakawa, A. and Lamb, V. R.: Computational design of the basic dynamical processes of the UCLA general circulation model, in: Methods of Computational Physics 17, Academic Press, New York, 173–265, 1977.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T.: Operational convective-scale numerical weather prediction with the COSMO Model: description and sensitivities, Mon. Weather Rev., 139, 3887–3905, doi:10.1175/MWR-D-10-05013.1, 2011.

Bennouna, Y. S., Torres, B., Cachorro, V. E., Ortiz de Galisteo, J. P., and Toledano, C.: The evaluation of the integrated water vapour annual cycle over the Iberian Peninsula from EOS-MODIS against different ground-based techniques, Q. J. Roy. Meteor. Soc., 139, 1935–1956, doi:10.1002/qj.2080, 2013.

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H.: GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system, J. Geophys. Res., 97, 15787–15801, doi:10.1029/92JD01517, 1992.

Böhme, T., Stapelberg, S., Akkermans, T., Crewell, S., Fischer, J., Reinhardt, T., Seifert, A., Selbach, C., and van Lipzig, N.: Long-term evaluation of COSMO forecasting using combined observational data of the GOP period, Meteorol. Z., 20, 119–132, doi:10.1127/0941-2948/2011/0225, 2011.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G., and Webb, M. J.: How well do we understand and evaluate climate change feedback processes?, J. Climate, 19, 3445–3482, doi:10.1175/JCLI3819.1, 2006.

Buehler, S. A., Östman, S., Melsheimer, C., Holl, G., Eliasson, S., John, V. O., Blumenstock, T., Hase, F., Elgered, G., Raffalski, U., Nasuno, T., Satoh, M., Milz, M., and Mendrok, J.: A multi-instrument comparison of integrated water vapour measurements at a high latitude site, Atmos. Chem. Phys., 12, 10925–10943, doi:10.5194/acp-12-10925-2012, 2012.

Dai, A.: Diurnal variation in water vapor over North America and its implications for sampling errors in radiosonde humidity, J. Geophys. Res., 107, ACL 11-1–ACL 11-14, doi:10.1029/2001JD000642, 2002.

Dick, G., Gendt, G., and Reigber, C.: First experience with near real-time water vapor estimation in a German GPS network, J. Atmos. Sol.-Terr. Phy., 63, 1295–1304, doi:10.1016/S1364-6826(00)00248-0, 2001.

Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N.F., Font, J., Grubišić, V., Gualdi, S., Homar, V., Ivančan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P., Llasat, M.C., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero, R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., Wernli, H., HyMeX: A 10-Year Multidisciplinary Program on the Mediterranean Water Cycle, Bull. Amer. Meteor. Soc., 95, 1063–1082, doi:10.1175/BAMS-D-12-00242.1, 2014.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W. J., Leon, J.-F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, J. Geophys. Res., 111, D11208, doi:10.1029/2005JD006619, 2006.

Fang, P., Bevis, M., Bock, Y., Gutman, S., and Wolfe, D.: GPS and meteorology: reducing systematic errors in geodetic estimates for zenith delay, Geophys. Res. Lett., 25, 3583–3586, doi:10.1029/98GL02755, 1998.

Gao, B.-C. and Kaufman, Y. J.: Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels, J. Geophys. Res., 108, 4389, doi:10.1029/2002JD003023, 2003.

Gendt, G., Dick, G., Reigber, C., Tomassini, M., Liu, Y., and Ramatschi, M.: Near real time GPS water vapor monitoring for numerical weather prediction in Germany, J. Meteorol. Soc. Jpn., 82, 361–370, 2004.

Di Girolamo, P., Summa, D., and Ferretti, R.: Multiparameter raman lidar measurements for the characterization of a dry stratospheric intrusion event, J. Atmos. Ocean. Tech., 26, 1742–1762, doi:10.1175/2009JTECHA1253.1, 2009.

Ingold, T., Schmid, B., Mätzler, C., Demoulin, P., and Kämpfer, N.: Modeled and Empirical Approaches for Retrieving Columnar Water Vapor from Solar Transmittance Measurements in the 0.72, 0.82, and 0.94 μm Absorption Bands, J. Geophys. Res.-Atmos., 105, 24327–24343, doi:10.1029/2000JD900392, 2000.

Kahn, B. H., Teixeira, J., Fetzer, E. J., Gettelman, A., Hristova-Veleva, S. M., Huang, X., Kochanski, A. K., Köhler, M., Krueger, S. K., Wood, R., and Zhao, M.: Temperature and water vapor variance scaling in global models: comparison to satellite and aircraft data, J. Atmos. Sci., 68, 2156–2168, doi:10.1175/2011JAS3737.1, 2011.

Kiehl, J. T. and Trenberth, K. E.: Earth's annual global mean energy budget, B. Am. Meteorol. Soc., 78, 197–208, doi:10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2, 1997.

Löhnert, U. and Crewell, S.: Accuracy of cloud liquid water path from ground-based microwave radiometry 1. Dependency on cloud model statistics, Radio Sci., 38, 8041, doi:10.1029/2002RS002654, 2003.

Löhnert, U., Schween, J. H., Acquistapace, C., Ebell, K., Maahn, M., Barrera-Verdejo, M., Hirsikko, A., Bohn, B., Knaps, A., O'Connor, E., Simmer, C., Wahner, A., and Crewell, S.: JOYCE: Jülich observatory for cloud evolution, B. Am. Meteorol. Soc., doi: 0.1175/BAMS-D-14-00105.1, 2014.

Martin, L., Mätzler, C., Hewison, T. J., and Ruffieux, D.: Intercomparison of integrated water vapour measurements, Meteorol. Z., 15, 57–64, doi:10.1127/0941-2948/2006/0098, 2006.

Maschwitz, G., Löhnert, U., Crewell, S., Rose, T., and Turner, D. D.: Investigation of ground-based microwave radiometer calibration techniques at 530 hPa, Atmos. Meas. Tech., 6, 2641–2658, doi:10.5194/amt-6-2641-2013, 2013.

Morland, J., Collaud Coen, M., Hocke, K., Jeannet, P., and Mätzler, C.: Tropospheric water vapour above Switzerland over the last 12 years, Atmos. Chem. Phys., 9, 5975–5988, doi:10.5194/acp-9-5975-2009, 2009.

Nash, J., Oakley, T., Vömel, H., and Wei, L.: WMO intercomparison of high quality radiosonde systems, Yangjiang, China, 12 July–

3 August 2010, Tech. Rep. 107, World Meteorological Organization, available at: http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-107_Yangjiang/IOM-107_Yangjiang.zip (last access: 1 September 2014), WMO/TD-No. 1580, 2011.

Pałm, M., Melsheimer, C., Noël, S., Heise, S., Notholt, J., Burrows, J., and Schrems, O.: Integrated water vapor above Ny Ålesund, Spitsbergen: a multi-sensor intercomparison, Atmos. Chem. Phys., 10, 1215–1226, doi:10.5194/acp-10-1215-2010, 2010.

Rocken, C., Van Hove, T., and Ware, R.: Near real-time GPS sensing of atmospheric water vapor, Geophys. Res. Lett., 24, 3221–3224, doi:10.1029/97GL03312, 1997.

Rose, T., Crewell, S., Löhnert, U., and Simmer, C.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, Atmos. Res., 75, 183–200, doi:10.1016/j.atmosres.2004.12.005, 2005.

Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., and Iga, S.: Nonhydrostatic Icosahedral Atmospheric Model (NICAM) for global cloud resolving simulations, J. Comput. Phys., 227, 3486–3514, doi:10.1016/j.jcp.2007.02.006, 2008.

Schlemmer, L., Hohenegger, C., Schmidli, J., Bretherton, C. S., Schär, C.: An Idealized Cloud-Resolving Framework for the Study of Midlatitude Diurnal Convection over Land, J. Atmos. Sci., 68, 1041–1057, doi:10.1175/2010JAS3640.1, 2011.

Schmid, B., Michalsky, J. J., Slater, D. W., Barnard, J. C., Halthore, R. N., Liljegren, J. C., Holben, B. N., Eck, T. F., Livingston, J. M., Russell, P. B., Ingold, T., and Slutsker, I.: Comparison of columnar water-vapor measurements from solar transmittance methods, Appl. Optics, 40, 1886–1896, 2001.

Schneider, M., Romero, P. M., Hase, F., Blumenstock, T., Cuevas, E., and Ramos, R.: Continuous quality assessment of atmospheric water vapour measurement techniques: FTIR, Cimel, MFRSR, GPS, and Vaisala RS92, Atmos. Meas. Tech., 3, 323–338, doi:10.5194/amt-3-323-2010, 2010.

Schween, J. H., Hirsikko, A., Löhnert, U., and Crewell, S.: Mixing layer height retrieval with ceilometer and Doppler lidar: from case studies to long-term assessment, Atmos. Meas. Tech. Discuss., 7, 4275–4319, doi:10.5194/amtd-7-4275-2014, 2014.

Seemann, S. W., Li, J., Menzel, W. P., and Gumley, L. E.: Operational retrieval of atmospheric temperature, moisture, and ozone from MODIS infrared radiances, J. Appl. Meteorol., 42, 1072–1091, doi:10.1175/1520-0450(2003)042<1072:OROATM>2.0.CO;2, 2003.

Shao, Y., Liu, S., Schween, J. H., and Crewell, S.: Large-Eddy atmosphere–land-surface modelling over heterogeneous surfaces: model development and comparison with measurements, Bound.-Lay. Meteorol., 148, 333–356, doi:10.1007/s10546-013-9823-0, 2013.

Sherwood, S. C., Roca, R., Weckwerth, T. M., and Andronova, N. G.: Tropospheric water vapor, convection, and climate, Rev. Geophys., 48, RG2001, doi:10.1029/2009RG000301, 2010.

Solheim, F. S., Vivekanandan, R., Ware, R. H., and Rocken, C.: Propagation Delays Induced in GPS Signals by Dry Air, Water Vapor, Hydrometeors and other Atmospheric Particulates, J. Geophys. Res., 104, 9663–9670, 1999.

Stull, R.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, 1988.

Torres, B., Cachorro, V. E., Toledano, C., Ortiz de Galisteo, J. P., Berjón, A., de Frutos, A. M., Bennouna, Y., and Laulainen, N.: Precipitable water vapor characterization in the Gulf of Cadiz region (southwestern Spain) based on Sun photometer, GPS, and radiosonde data, J. Geophys. Res., 115, D18103, doi:10.1029/2009JD012724, 2010.

Vogelmann, H., Sussmann, R., Trickl, T., and Borsdorff, T.: Intercomparison of atmospheric water vapor soundings from the differential absorption lidar (DIAL) and the solar FTIR system on Mt. Zugspitze, Atmos. Meas. Tech., 4, 835–841, doi:10.5194/amt-4-835-2011, 2011.

Wang, J. and Zhang, L.: Systematic errors in global radiosonde precipitable water data from comparisons with ground-based GPS measurements, J. Climate, 21, 2218–2238, doi:10.1175/2007JCLI1944.1, 2008.

Weckwerth, T. M., Wulfmeyer, V., Wakimoto, R. M., Hardesty, R. M., Wilson, J. W., and Banta, R. M.: NCAR–NOAA lower-tropospheric water vapor workshop, B. Am. Meteorol. Soc., 80, 2339–2357, doi:10.1175/1520-0477(1999)080<2339:NNLTWV>2.0.CO;2, 1999.

Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G.C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C. N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G., Richard, E., Rotach, M. W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D. D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A., Wirth, M., The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights, Quart. J. Roy. Meteor. Soc., 137, 3–30, doi:10.1002/qj.752, 2011.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: description of the non-hydrostatic dynamical core, Q. J. Roy. Meteor. Soc., doi:10.1002/qj.2378, 2014.