Precipitable water vapor over oceans from the Maritime Aerosol Network: Evaluation of global models and satellite products under clear sky conditions

Daniel Pérez-Ramírez[1,2], Alexander Smirnov[3,4], Rachel T. Pinker[5], Maksym Petrenko[3,6], Roberto Román[7], W. Chen[5], Charles Ichoku[3], Stefan Noël[8], Gonzalo Gonzalez Abad[9], Hassan Lyamani[1,2], and Brent N. Holben[3]

[1] Applied Physics Department, University of Granada, Granada, Spain

[2] Andalusian Institute for Earth System Research (IISTA), Granada, Spain

[3] NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

[4] Science Systems and Applications, Inc., Lanham, Maryland, USA.

[5] Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA

[6] ADNET Systems Inc., Bethesda, Maryland, USA.

[7] Atmospheric Optics Group, University of Valladolid, Valladolid, Spain

[8] Institute of Environmental Physics, University of Bremen, Bremen, Germany

[9] Harvard Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA

Correspondence to: Daniel Perez-Ramirez; E-mail: dperez@ugr.es

ABSTRACT: We present results from an evaluation of precipitable water vapor (*W*) over remote oceanic areas as derived from global reanalysis models and from satellites against observations from the Maritime Aerosol Network (MAN) for cloudless skies during the period of 2004–2017. They cover polar, mid latitude and tropical oceanic regions and represent a first effort to use MAN observations for such evaluation. The global reanalysis model products evaluated in this study are from the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2), the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA I), and the Climate Forecast System Reanalysis (CFSR) model. The satellite products evaluated are from the Moderate Resolution Imaging Spectroradiometer (MODIS), the Polarization and Directionality of the Earth's Reflectances (POLDER), the Global Ozone Monitoring Experiment (GOME-2), the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), and the Atmospheric Infra-red Sounder (AIRS). Satellite retrievals of *W* are based on the attenuation of solar reflected light by water vapor absorption bands, except those from AIRS that rely on brightness temperature measurements. A very good agreement is observed between the model estimates and MAN, with mean differences of ~5% and standard deviations of ~15 %. These results are within the uncertainties associated with the models and the measurements, indicating the skill of the reanalysis models to estimate *W* over oceans under clear sky conditions. Mean differences of *W* between the satellite and MAN products are ~ 11, 6.7, 12, -7, and 3% for MODIS, POLDER, GOME-2, SCIAMACHY and AIRS respectively, while their standard deviations are 31, 29, 28, 20 and 17 %. These differences reveal the need to address inconsistencies among different satellite sensors and ground-based measurements to reduce the uncertainties associated with the retrievals.

1. Introduction

Water vapor is one of the most important components of the Earth's atmosphere that affects both weather and climate. It dominates tropospheric diabatic heating by condensation of water into liquid in the lower troposphere [*Trenberth and Stepaniak, 2003*], and is the most important gaseous constituent for infrared opacity in the atmosphere [*Trenberth et al., 2007*]. Information on water vapor is essential for understanding mesoscale meteorological systems and cloud formation [*Wulfemeyer et al., 2015*]. Water vapor also contributes indirectly to radiative forcing, influencing the microphysical processes leading to the formation of clouds, and affecting the size, shape and the chemical composition of aerosols [*Reichard et al., 1996*]. Information on water vapor over oceans is especially important because more than three quarters of the total exchange of water between the atmosphere and the Earth's surface occurs through ocean evaporation and precipitation [*Schmitt, 2008*].

The Compendium of Meteorology of the American Meteorological Society defines the precipitable water vapor (*W*) as "the total atmospheric water vapor contained in a vertical column of unit cross-section, extending in terms of the height to which that water substance would stand if completely condensed and collected in a vessel of the same unit cross section" [*AMS, 2000*]. Measurements of *W* are available from different ground-based remote sensing instruments, such as sun-photometers [e.g. *Alexandrov et al., 2009*], moon/star photometers [e.g. *Barreto et al., 2013*], Fourier-Transform spectrometers [e.g. *Leblanc et al., 2011*], microwave radiometers [e.g. *Cadeddu et al., 2013*], and global positioning system (GPS) receivers [e.g. *Bevis et al., 1992*]. Precipitable water vapor is also obtained by integrating water vapor vertical profiles from radiosondes [e.g. *Durre et al., 2006*] and Raman lidar systems [e.g. *Whiteman et al., 2010, 2012*]. However, most of these instruments are deployed over land.

Recent versions of global reanalysis models assimilate many meteorological variables, including moisture profiles from radiosondes, and are capable of simulating *W* over the entire globe. Satellite sensors provide a global coverage of *W* using space-borne instruments that utilize different physical concepts for remote sensing of *W*. MODIS [*King et al., 1992*] and POLDER [*Deschamps et al., 1994*] are based on Earth´s reflectance of water vapor absorption channels in the infrared and near-infrared; GOME-2 [*Munro et al., 2006, 2016*] and SCIAMACHY [*Bovensmann et al., 1999; Gottwald and Bovensmann, 2011*] use Differential Optical Absorption Spectroscopy (DOAS) with the absorption bands of O_2 and H_2O; other space-borne sensors such as AIRS [*Aumann et al., 2003*] rely on microwave radiometry. However, in spite of the wide-ranging data sources, it is still a great challenge to evaluate water vapor estimates over oceans due to lack of surface-based measurements over remote oceanic areas.

Measurements from ships are essential to augment the low rate of *W* measurements over oceans; several field campaigns have been organized [e.g. *Nalli et al., 2011*] to address this shortcoming. The Maritime Aerosol Network (MAN) is a component of the Aerosol Robotic Network (AERONET) [*Holben et al*., 2001] and aims to primarily improve our knowledge of aerosol properties over oceans using sun photometry. MAN has been operating since October 2004, with over 450 cruises completed and more than 6000 measurement days recorded, and the data are stored in a web-based public data archive (https://aeronet.gsfc.nasa.gov/new_web/maritime_aerosol_network.html). Consequently, MAN has had a great success in providing ground truth for evaluating satellite-derived aerosol optical properties over oceans [e.g. *Smirnov et al., 2011, 2017*].

Currently, most of the MAN campaigns operate sun-photometers with filters centered around 940 nm wavelength, which is one of the main atmospheric water vapor absorption bands [e.g. *Reagan et al., 1986*; *Halthore et al., 1997*] and, therefore, it is possible to retrieve *W*. MAN follows the same processing protocol as AERONET, making MAN an excellent data source for evaluating *W* data over oceans under clear sky conditions. MAN data are only available when the sun is not obstructed by clouds, yet, they can provide information on *W* during the precursory stages of extreme weather [*Ye et al., 2014; Fujita and Sato, 2016*] or for studying aerosol hygroscopic growth [e.g. *Veselovskii et al., 2009*].

In section 2 we describe the instrumentations and methodologies used. Section 3 is devoted for the main results while in section 4 we provide the main conclusions.

2. Instrumentation and Methodology

2.1. Maritime Aerosol Network (MAN)

The standard instrument used in the MAN is the Microtops II sun photometer [*Smirnov et al., 2009*]. Microtops II is a portable and hand-held manually operated instrument that measures direct solar irradiance. Microtops II has five spectral channels and can accommodate several filter configurations within the spectral range of 340-1020 nm. The bandwidths of the interference filters vary from 2 to 4 nm for UV channels, to 10 nm for visible and near-infrared channels. Microtops II provides information allowing estimating aerosol optical depth (AOD) and also precipitable water vapor (W) if the filter centered at 940 nm is used. MAN instruments follow the calibration criteria and data processing of AERONET. Each Microtops II instrument is calibrated against an AERONET master-CIMEL Sun/sky radiometer at NASA Goddard Space Flight Center (GSFC), traceable to Langley plot measurements at Mauna Loa. These Microtops II calibrations are done under clear sky and stable atmospheric conditions to ensure accurate and stable results. Filters are replaced when drastically degraded. Microtops II sun photometers have demonstrated good calibration stability over the years [*Ichoku et al., 2002*].

The measurement protocol of MAN is described in detail in *Smirnov et al. (2009*), briefly summarized here. Measurements are taken as 6-10 scans when the solar disk is free of clouds. Each scan takes about 7-8 seconds; each measurement sequence takes over a minute plus some time for a GPS to lock ship's position. If the interval between two consecutive scans is more than two minutes, then these points are placed into a different time series. A series is considered a single data point and can have one or more measurement points (typically five).

Sun is considered not obstructed by clouds based on visual assessment; depending on sky conditions, measurements should be repeated several times during the day. MAN instruments follow data processing of AERONET and here we use MAN Level 2.0 results that guarantee acceptable cloud-screening and data quality (e.g. Smirnov et al., 2009). Briefly, within a series of observations, the minimum aerosol optical depth (AOD_{min}) is computed at each wavelength. For the rest of points if the absolute difference AOD_i – AOD_{min} for each spectral channel is less than the maximum of {AOD_{min}*0.05, 0.02}, that point within a series is considered cloud-free and pointing error free. We note that the criterion is applied to *AOD*, but if the point does not pass the test, then all spectral channels for these measurements are removed, including the W channel. Finally, after this test using AOD, if only one point remains after this evaluation, an additional criterion consisting of evaluating Angstrom parameter is used: if it is greater than 0.1 then the point is considered cloud-screened and with accurate pointing.

For our purposes of studying W, the direct solar irradiance at 940 nm measured by Microtops II instrument allows direct estimation of water vapor transmittance (T_w (940 nm)) using a simplified expression of T_w (940 nm), as given by [e.g. *Schmid et al., 2001*]:

$$T_w(940nm) = \exp\left(-a(m_w W)^b\right) \quad (1)$$

where m_w is the relative optical water vapor air mass and '*a'* and '*b'* are coefficients that depend on the wavelength position, width and shape of the sun-photometer filter function, and the atmospheric condition [*Halthore et al., 1997*]. Each Microtops II instrument has its own unique set of '*a*' and '*b*' values depending on its specific filter configuration. These coefficients are considered fixed until the filter is changed. More information about the computation of coefficients '*a*' and '*b'* can be found in *Smirnov et al.,* [2004].

The good agreement between Microtops II and AERONET values of W was demonstrated by *Ichoku et al*., [2002] for correlative measurements with both instruments. Therefore, we assume that MAN values of W (W_{MAN}) have similar uncertainties to AERONET values as discussed in *Pérez-Ramírez et al.,* [2014] who reported uncertainties below 10 %.

2.2. Global reanalysis Models and Satellite Sensors

Table 1 summarizes the main characteristics of the global W products from reanalysis models and the satellite sensors that were evaluated in this study, including their spatial resolutions and data availability periods. The reanalysis models whose W data have been selected for evaluation are the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) from the NASA Global Modeling and Assimilation Office (GMAO) - *Gelaro et al. [2017]*], the Climate Forecast System Reanalysis (CFSR) –

Saha et al. [2010] from The National Centers for Environmental Prediction (NCEP), and the ERA Interim Reanalysis model (ERA-I) - *Berrisford et al*., [2011] from The European Center for Medium-Range Weather Forecast (ECMWF). All of these global reanalysis models assimilate meteorological parameters measured from different space-borne sensors (e.g. radiances, surface wind speeds and vectors, temperature and ozone profiles). Global reanalysis models must be evaluated against independent and accurate ground-based measurements.

[Insert Table 1 here]

The satellite products evaluated in this study include those of the Moderate Resolution Imaging Spectroradiometer (MODIS) [*King et al., 1992*] and the Polarization and Directionality of the Earth's Reflectances (POLDER) [*Deschamps et al., 1994*] that obtain *W* from the ratio of reflected radiances at water vapor absorption channels and non-absorbing bands in the infrared and near infrared regions of the spectrum. All MODIS and POLDER data used are cloud-screened and they are based on passive remote sensing techniques (low power supply, continuous operation). For MODIS, we use the infrared algorithm (5 x 5 km pixel resolution) that employs ratios of water vapor absorbing channels at 0.905, 0.936, and 0.940 µm with atmospheric window channels at 0.865 and 1.24 µm [*Kaufman and Gao, 1992; Gao and Goetz, 1990*], while POLDER is based on the ratio of reflected radiances at 910 nm and 865 nm [*Vesperini et al., 1999*]. The ratios partially remove the effects of variation of surface reflectance with wavelengths and provide water vapor transmittances, although can be affected by spectral dependences of aerosol attenuation. In MODIS, *W* is derived from water vapor transmittances using look-up table procedures and we are using the current Level 2 Collection 6 data (https://modis.gsfc.nasa.gov/data/dataprod/mod05.php), while for POLDER, an approximate empirical equation is used for estimating *W* [*Vesperini et al., 1999*], and we are using the Level 2 data (http://www.icare.univ-lille1.fr/).

Other sensors whose *W* retrievals are evaluated are the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) [*Bovensmann et al., 1999; Gottwald and Bovensmann, 2011*] and the Global Ozone Monitoring Experiment (GOME-2) [*Munro et al., 2006, 2016*]. The *W* retrieval technique for these instruments is based on the Differential Optical Absorption Spectroscopy (DOAS) approach. Again, these two instruments are based on passive remote sensing and the data used are cloud-screened. SCIAMACHY data are provided by the University of Bremen (http://www.iup.uni-bremen.de/amcdoas/), and their method involves fitting the differential structures of the measured spectral reflectance [B*urrows et al., 1999*], where upon the water vapor is retrieved using an approach similar to the simplified T_w *(940 nm)* of equation 1, but spectrally resolved for wavelengths close to 700 nm. Furthermore, an additional correction based on simultaneous O_2 measurements is performed [*Noël et al.,* 1999, 2004, 2008]. The GOME-2 data are provided by the Earth Observation Center of the German Aerospace Center (http://atmos.eoc.dlr.de/) and their retrieval algorithm consists of fitting water vapor absorption bands in the range 614-683 nm and also uses simultaneous O_2 measurements [*Wagner et al., 2003, 2006*].

The additional satellite sensor whose *W* data have been used is the Atmospheric Infrared Sounder (AIRS) [*Aumann et al., 2003*], which is a hyperspectral, scanning infrared sounder. AIRS measures the infrared brightness from Earth's surface and from atmospheric constituents. By having multiple infrared detectors, each sensing a particular wavelength, temperature and water vapor profiles can be estimated. AIRS has 2378 detectors while previous sensors had only 15. Such instrument is well suited for climate studies allowing high accuracy of temperature and water vapor. Particularly, AIRS water vapor retrieval algorithm uses 66 spectral channels that are generally selected to cover a range of wavelengths on and off water vapor absorption bands [*Susskind et al., 2003*]. The use of several detectors in the infrared regions minimizes sources of errors associated with surface reflectance or with aerosols. The AIRS *W* data used are version 6 Level 2 (https://airs.jpl.nasa.gov/data/). Although AIRS can provide *W* estimates under cloudy conditions, we utilized only the clear-sky observations.

The different satellite sensors used for *W* estimates over oceans are affected by additional systematic and random errors such as errors of calibration of the channels used, errors in the radiative transfer in the forward models or errors associated with the viewing angles (viewing geometry). These issues have been addressed by previous studies and were included in the final error uncertainties for each satellite product (*Ichoku et al*. [2005] for MODIS and POLDER, Noël et al. [2008] for GOME-2 and SCIAMACHY and *Susskind et* *al.* [2003, 2006] for AIRS). Other sources of errors in the estimates of W by satellite sensors are the inaccurate surface reflectance characterization and the different hypothesis assumed in the retrievals by each sensor.

2.3. Matchups between Maritime Aerosol Network and Global Reanalysis Models/ Satellite Sensors

To compare with model data, MAN 'series' are first timely averaged around the standard times when models provide information, namely, 00, 03, 06, 09, 12, 15, 18 and 21 UTC for MERRA-2 and 00, 06, 12 and 18 UTC for ERA-I and CFSR. Temporal windows are of ±1.5 hours for averaging for MERRA-2 and ±3 hours for ERA-I and CFSR. Mean W, latitude and longitude are therefore determined for the data within each temporal window. For models, a sampling area of 1°x1° around mean latitude and longitude by MAN is selected and the corresponding model value of W is a weighted mean using the distances to the averaged coordinates of the corresponding MAN observations.

For the match-ups with satellite observations, we use the Multi-Sensor Aerosol Products Sampling System (MAPSS) [*Petrenko et al., 2012*] adapted for MAN [*Smirnov et al., 2017*]. For each MAN series measurement and each satellite sensor, MAPPS check if there is an overpass that contains pixels retrieved within ±30 minutes and ±50 km (±27.5 km for MODIS) of ship-based measurements. These selected MAN data are subsequently averaged including W, latitude and longitude, and identified as a single 'central' ship-based measurement. MAAPS samples coincident space-borne pixels within ±50 km (±25.5 km for MODIS) of this central ship-based location and corresponding space-borne value of W is a weighted mean using the distances to the central ship-based location. Note that MAN measurements in a one-hour time window coincides with at most a single overpass for a given sensor due to the low speed of the ships.

In our analysis, the calculation of deviations between the MAN measurements and model-assimilated or satellite datasets are based on the mean differences or relative differences that represent the systematic errors, while their standard deviations, which represent the variability of these differences, are denoted as the uncertainty measures of these datasets.

3.0 Results

3.1 Precipitable Water Vapor over Oceans by Maritime Aerosol Network

Figure 1 shows daily averages of W for all MAN cruises. There are more than 36000 measurements for the period 2004-2017 covering several oceanic regions, although the most frequently sampled places are the areas close to the continents and in the mid-Atlantic region. Also, the Red, Black, North, Mediterranean, Caribbean, Baltic, and Chinese seas are very well sampled. Other places with numerous measurements are the Gulf of Bengal and of Mexico, the high latitude oceanic regions with cruises in the Arctic Ocean and near Antarctica. The Pacific Ocean has many measurements, but because of its large size it is not considered well sampled. The situation is similar for the Indian Ocean.

Figure 1 illustrates regional variability of W under clear sky conditions. The highest values of W are found in the tropics with 75 % of W values between 2-4 cm and maximum values above 6 cm. Values of W below 1 cm in the tropics are rare, with only 1 % of occurrence. Mid latitudes present lower values of W with 75% of the data between 1-3 cm. Mid-latitudes also present the largest variability in W with 18% of the data below 1 cm and 6% of the data above 4 cm. High latitudes present the lowest values with 80 % of the data below 1 cm. Values of W above 2 cm for these latitudes are uncommon with only 1% of occurrence.

Statistics for latitudes above 30º and below -30º reveal mean values of 0.99 ± 0.77 cm for the southern hemisphere and of 1.57 ± 0.81 for the northern hemisphere. But due to the limitations of the sun-photometry (measurements are only available when solar disk is cloud-free) and to the differences on ship tracks in different latitudes, no additional hemispheric dependence can be investigated.

[Insert Figure 1 here]

3.2 Evaluation of W using global reanalysis models

Figure 2 (a)-(c) shows differences in W between models and MAN data as a function W as measured by MAN (W_{MAN}). The dashed lines in the plot represent ±10% difference versus measured MAN values while the dot lines represent ±20% differences. Figure 2 (d)-(f) shows W from models as a function of W as measured by MAN (W_{MAN}), where red lines are the least-square fits and the dashed lines the 1:1 line (reference for a perfect agreement). For clarity, we use number density plots in Figure 2. They divide the plot into different pairs of ‘x_i’ and ‘y_i’ values. In Figures 2 (a)-(c) ‘x_i’ are the W_{MAN} values below 7.0 cm, while ‘y_i‘ are the differences between global reanalysis and MAN data varying between -2.0 and 2.0 cm (there are some outliers with larger deviations omitted for clarity). In Figures 2 (d)-(f) ‘x_i’ are again W_{MAN} while ‘y_i‘ are the W global reanalysis models estimates, being now both ‘x_i’ and ‘y_i’ below 7.0 cm. Later, we compute the number of occurrences for every pair (x_i,y_i) and finally, results are plotted on a map, where the scale goes from zero to the maximum number of occurrences.

Table 2 summarizes the main statistics of these evaluations, particularly, the mean, median and standard deviations values of the differences W_i - W_{MAN} and of the relative difference $(W_i - W_{MAN})/W_{MAN}$. Given are also parameters of the classical least-squares linear fit $y_M = Ax+B$, where the coefficient A is the slope of the linear fit and the coefficient B is the ordinate intercept. Table 2 also includes the total number of comparisons for each model and sensor, and we note that the differences in number of data are explained by the different periods of measurements available and the different spatial resolutions.

[Insert Figure 2 here]

[Insert Figure 3 here]

[Insert Table 2 here]

Estimates of precipitable water vapor from global reanalysis models are for all sky conditions, while W_{MAN} is only for clear-skies. Global reanalysis models assimilate many atmospheric parameters including satellite radiances. Assimilated radiances from the visible and near infrared regions are for clear sky conditions; only radiances from the microwave regions under cloudy conditions are useful for assimilations, which are critical for improving model forecast capabilities in these conditions [e.g. *Reale et al.,* 2008].

Models and MAN data are highly correlated (R^2 above 0.87) with slopes of the linear fit very close to unity and abscissas cut-off very close to zero. The models show a very good agreement with MAN, with only a small overestimation that is below 5%. The standard deviations of ~ 15 % between models and MAN implies 5% uncertainties in model estimates of W when considering 10 % uncertainties for sun photometry [*Perez-Ramirez et* *al., 2014*]. The 5% uncertainty for models is supported by Figure 2 (a)-(c) where most of the data fall within the region of ±10% difference. Deviations from these uncertainties are observed but are assumed as outliers between models and MAN, and probably associated with incorrect MAN data (e.g. possible cloud contamination) or issues with models.

Figure 3 shows the differences in between models and MAN W data as a function of latitude and reflects differences between Tropical, Mid-latitudes and Polar regions: the largest and smallest differences in W are found in the Tropical and Polar Regions, respectively. However, when relative differences $(W-W_{MAN})/W_{MAN}$ are evaluated, no significant differences with latitude are observed that can be explained by the dependences of W on latitude (Figure 1).

Figure 4 shows the frequency histograms of the differences between model and MAN values of W. The frequency histograms are normal and centered close to zero (they are exactly centered at the mean values of Table 2 and the full width at half maximum (FWHM) are the standard deviations). Therefore, from the results presented here, models based on reanalysis reproduce well-observed values of W over oceans with an approximate accuracy of 10 - 15 %, which reflect the robustness and feasibility of W estimates over oceans under clear sky conditions by global reanalysis models..

[Insert Figure 4 here]

3.3 Evaluation of W using satellite observations

Accurate retrievals of W from visible and near infrared satellite observations require a-priori cloud-filtering. For MODIS and POLDER cloud-filtering algorithms are applied [e.g. Martins et al., 2002; Levy et al., 2013]; in GOME-2 and SCIAMACHY clouds are removed because bias are introduced in the retrievals of W depending on clouds heights [see Figure 5 of du *Piesanie et al*. 2013]. For AIRS, clouds still affect the microwave radiation and for accurate information clouds need to be removed [*Susskind et al*. 2003]. Refined algorithms for cloud clearing in AIRS measurements and correct analysis of water vapor retrievals are found in *Susskind et al*. [2006, 2011, 2014].

Figure 5 (a)-(e) shows differences in W between satellite retrievals and MAN with dashed and doted lines representing ±10% and ±20% relative difference versus MAN measured values, while Figure 5 (f)-(j) shows W from satellite sensors versus W_{MAN}, where the red lines represent the least-square fits and the dashed lines the 1:1 line. Figure 6 shows the same differences as a function of latitude. Table 2 summarizes again all statistical parameters. For satellite sensors, differences in the number of points (*N*) available for comparison are explained by the frequency of correlative measurements and by the different spatial resolutions, e.g., MODIS presents a larger data set because there are two MODIS instruments on different platforms and it has a higher spatial resolution than the other satellite sensors involved in this study. The period of measurements also has an influence on data availability (SCIAMACHY and POLDER present the lowest number of data because of their shorter operation time).

[Insert Figure 5 here]

[Insert Figure 6 here]

Figure 5 (a) – (e) reveals departures from the zero line that are consistent with the non-unity slopes obtained for the regressions. From these linear fits we also observe slope departures from unity for all the satellite sensors, ranging between 0.88 and 1.06. For the ordinate intercept *B* there is also variability ranging from ~ 0.01cm to ~ 0.23 cm. Satellite and MAN data are again highly correlated (R^2 above 0.87) although their relative differences are larger than those obtained when comparing with global reanalysis models. All comparisons show outliers with large underestimation/overestimation above ±2.0 cm that may be associated with incorrect satellite *W* values (e.g. possibly affected by cloud contamination) or the natural variability of water vapor during the matchup process. Percentages of data within ±20% of relative differences are 59.6, 69.0, 67.1, 60.4 and 85.5% for MODIS, POLDER, SCIAMACHY, GOME-2 and AIRS, respectively; the percentages within ±10 % relative differences are of 33.5, 45.7, 38.4, 30.5 and 55.3 %. However, there are differences in the analysis for each satellite sensor. The differences in measurement techniques, retrieval methodologies and effects of spatial resolutions and viewing geometries of each sensor can cause differences among satellites and MAN.

For the MODIS sensor we show only data for the infrared algorithm. The MODIS infrared retrievals of *W* overestimate MAN data by ~ 11%, although the median difference is (~ 5 %), indicating that outliers with very high *W* from MODIS can contaminate the statistics. Over land, estimates of *W* from MODIS observations have been reported to be about ~ 10-15 % [A*lbert et al., 2005; Román et al., 20014; Liu et al., 2015; Alradadawi et* *al., 2018],* larger than the ~ 5 % found in this study. But the standard deviations of the differences over oceans and seas showed here of ~ 30% suggest that assuming 10 % uncertainty in MAN yields 20% uncertainty in MODIS retrievals in the best case when errors are correlated. No statistically significant differences were found between instruments on Terra and Aqua platforms (relative differences of 8% for Terra and 12% for Aqua, and both had 30% standard deviation). Departures from the ±20% relative differences (Figure 5a) are observed for all the ranges of *W*. The detailed analyses revealed that approximately 30 % of the data are above 20% relative difference and 11% of data are below -20% relative differences. The analysis was repeated for Tropical, mid latitudes and Polar Regions; no latitudinal dependence of the relative differences was found (Figure 6a). A possible reason for systematic discrepancies between MODIS and MAN could be the assumptions in MODIS retrievals that the ratio between signals inside/outside the absorption band does not depend on surface reflectance. A revision of the radiative transfer code might improve the results presented here.

POLDER had low differences between satellite estimates and MAN observations, with mean deviations of about 6.7% and standard deviations of about 30%. Over land estimates from POLDER were found to be ~ 15-20 % [*Vesperini et al., 1999*]. The better agreement over oceans can be associated with the more homogenous surface reflectance that affects the retrievals. But the dependence of the differences with *W* revealed important features for low values of *W* (Figure 5b). Actually, for W < 1 cm 49% of the data present relative difference above 20% and 7% of the data with present relative difference below -20%. This dependence of the relative differences with *W* explains the dependence on latitude seen in Figure 6b, with mean values of the differences of -5.6 ± 13 % for the Tropics, -6.0 ± 15 % for mid latitudes and 23.4 ± 34.5 % for Polar Regions. Because POLDER uses a similar measurement strategy to MODIS, differences between instruments and between regions can be explained by differences in the retrieval technique, namely, correction for surface reflectance or the assumption of the constant surface reflectance for all oceanic areas that can be important in Polar regions due to effects of ice and snow.

The satellite retrievals based on the DOAS technique present different biases. GOME-2 (Figure 5c) overestimates MAN data, with a mean relative difference of ~12.5 % and fairly similar difference between the instrument placed in MetOp-A (~ 13.2 %) and MetOp-B (~

9.1 %), while SCIAMACHY (Figure 5d) shows an underestimation of MAN data with a mean relative difference of ~ -7.2 %. The results obtained here are similar to those obtained from GOME-2 over land [e.g. *Antón et al., 2015; Román et al., 2015; Vaquero-Martínez et al., 2018*]. The standard deviations of W evaluation over oceans are ~30 % and ~20 % for GOME-2 and SCIAMACHY, respectively, which implies uncertainties in W of 20% and 10%, assuming a 10% uncertainty in MAN data. GOME-2 shows departures from W_{MAN} data for the entire range of W. For $W < 1$ cm, 38% of the data present relative difference above 20 % while 9% of the data shows relative difference below -20%. Very similar percentages are found for $W > 1$ cm. These dependence of GOME-2 relative differences explain the dependencies of W with latitudes (Figure 6c), being mean relative differences of 7.8 ± 18.9, 15.9 ± 29.2 and 21.4 ± 39.6 % for Tropical, mid-Latitude and Polar Regions, respectively, clearly indicating that they are larger for lower values of W. Outliers are observed everywhere, but particularly, for low values of W in the polar regions for GOME-2 with differences of up to 2 cm, which is more than 200 % and can influence the statistics. These large differences in W between GOME-2 and ground-based measurements are also found over land at these latitudes, with systematic underestimations of W by GOME-2 *[e.g. Palm et al., 2010*]. Other studies found systematic overestimation of W by GOME-2 for very low values of W [*Vaquero-Martínez et al., 2018*], typically below 1.0 cm and most frequently found at polar regions. We believe that the variability of surface reflectance in Polar Regions can affects W retrievals. However, SCIAMACHY presents a very similar pattern of the relative differences with W, most of relative differences (~70%) being within the ±20% uncertainty (Figure 5d). These dependencies of the relative differences also justify the low regional dependences (Figure 6d) which are of -8.9 ± 13.2, -16.1 ± 18.6 and -0.04 ± 19.8 % for Tropical, mid-Latitude and Polar Regions, respectively. Note the lack of outliers in the Polar Regions, which explains the very good agreement with MAN, and also the better estimation of W over oceans by SCIAMACHY when compared with ground-based measurements over land at these latitudes [*e.g. Palm et al., 2010*]. As for MODIS and POLDER, differences between sensors can be explained by the different assumptions in the retrieval algorithms. We note the large difference in the number of matchups between GOME-2 and SCIAMACHY that can affect the statistics (see Table 2).

The best agreement between satellite and MAN data is observed for the AIRS system (Figure 5e), showing an overestimation of 3.1 %. The standard deviation of 17.3 % is within the uncertainties associated with each instrument, e.g., 10 % for both sun photometry and microwave radiometry, respectively. But important relative differences are found with *W* (Figure 6e): for *W* > 1 cm, 88% of the data are within the ±20% uncertainty while for *W* < 1cm this percentage is reduced to 44%. These dependencies with *W* explain the regional dependences observed (Figure 6e), with mean relative differences in *W* of 3.3 ± 13.4, 3.4 ± 14.9 and 16.8 ± 39.8 % for Tropical, mid latitudes and Polar Regions. Larger relative differences for low values of *W* are consistent with the literature in the comparisons between sun-photometry and microwave radiometry and needs for further studies using the same spectral database for both instruments [*Perez-Ramirez et al., 2014*]. Similar results are found from comparison of AIRS with ground-based measurements over land areas [*e.g. Qin et al., 2012; Roman et al., 2016*]; with larger values of *W* from AIRS for land areas close to the Artic [*Alradadawi et al., 2018*]. The better results from AIRS indicate that this instrument is possibly less sensitive to the presence of clouds.

Frequency histograms of the differences between satellite sensors and MAN data are given in Figure 7. Both MODIS and POLDER show normal distributions slightly skewed towards positive values, which explains the mean differences of approximately 6 -10 % (Table 2). Similar skewness is observed for GOME-2, while SCIAMACHY is skewed towards negative values. Differences among space sensors can be explained by the different assumptions in the retrieval methodologies, by the wavelength-dependence in surface reflectance and by the different data sample sizes used due to the different number of collocations. Also, the natural variability of water vapor can influence these findings when comparing measurements of different temporal and sampling resolution and when comparing the optical air mass from the ground and the path of reflected radiance to space sensors. Another important reason for the discrepancies is the assumption of the simplified water vapor transmittance $T_w = a(m_w W)^b$ used in satellite and sun photometry retrievals, as the constants '*a*' and '*b*' are filter-dependent functions and their calculation depends on the radiative transfer code used. Furthermore, the differences in the retrieved *W* between using lookup tables and simplified T_w equation depend on *W,* and vary between 9% for *W* > 1 cm and up to 25 % for lower values *[Pérez-Ramírez et al., 2012]*. This dependence on methodology is supported by the lower relative differences found in MODIS, which uses look-up tables, and with POLDER that shows the largest discrepancies for Polar regions with low *W*.

Finally, the frequency histogram for AIRS reveals a unimodal size distribution centered at 3.1% and with 17.3 % standard deviation which suggests that AIRS data over oceans present an uncertainty below 10 %. The 3.1 % overestimation found agrees with the general comparison between microwave radiometry and AERONET sun photometry [e.g. *Pérez-Ramírez et al., 2014*], although overestimation increases with low values of *W*. The low agreement for low values of *W* is independent of the satellite sensor. Actually, for very low values (W < 0.1 cm) the differences can reach up to 50 % frequently because absolute difference can be of ±0.04 cm. This is similar for global reanalysis models (Figure 2). These results imply the need for a minimum accuracy of ±0.02 cm for all sensors and methodologies.

[Insert Figure 7 here]

4.0 Conclusions

In this study we have described the use of the Maritime Aerosol Network (MAN) observations to evaluate precipitable water vapor (*W*) estimates over oceans as derived by global reanalysis models and satellite sensors. The Maritime Aerosol Network is a very unique observational network and covers a large portion of the oceans (tropics, mid-latitudes and polar regions) with the potential of providing information both on aerosols and water vapor. It complements the well-established and credible AERONET network (operating over land) and follows the same operating protocol. MAN measurements started in 2007 and are based on sun-photometry which implies clear-sky conditions. The study presented here has enhanced MAN capabilities for the evaluation of satellite products on remote oceanic areas.

The relative differences between global reanalysis models and MAN are below 15 %, which implies uncertainties in *W* estimates below 5%, and therefore, points to the usefulness of *W* estimates by global reanalysis models for atmospheric research and for climate monitoring. On the other hand, for satellite sensor estimates of W, generally differences between MAN and MODIS, POLDER, GOME-2 and SCIAMACHY were below 30% which is significantly larger compared to global models. Differences with latitude have been also observed being the largest for Polar Regions where the lowest values of *W* were observed; this can be explained because of the different hypothesis in the retrievals, e.g., differences in the assumptions on surface reflectance due to changes in ice areas. AIRS instrument is unique in deriving W and in this study we have demonstrated the best agreement with MAN compared to other satellite sensors, having uncertainties below 10%. Our results indicate that there is a need for a joint effort to comprehensively address the inconsistencies among the remote sensing techniques used with different satellite sensors and ground-based instruments in order to reduce uncertainties associated with the retrievals.

The results of this study are unique since they provided information on *W* for clear sky conditions over a large portion of the oceans. For cloudy conditions, different types of observations are needed (e.g., radiosondes). Measurements by active remote sensing such as Raman lidar or radars would also allow advances in the understanding of water vapor over oceans during extreme weather conditions. Such measurements should be of great interest for further advances in modeling reliable estimates of *W* and also in the evaluation of future estimates of *W* by space-borne sensors under cloudy conditions.

ACKNOWLEDGEMENTS

This work was supported by the Marie Skłodowska-Curie Individual Fellowships (IF) ACE_GFAT (grant agreement No 659398). The authors thank Dr. Hal Maring from NASA Headquarters and Dr. Steven Platnick from EOS Project Science Office for their support of AERONET. The work of RTP and WC benefited from support under NASA grant NNX13AC12G, the Energy and Water Cycle Study (NEWS) program. The authors would like to acknowledge managerial and operational support from M. Sorokin, J. Kraft, A. Scully at NASA GSFC, and MAN PIs for the creation and stewardship of the Sun photometer data records We thank the science and support teams of MODIS, POLDER, GOME-2, SCHIAMACHY and AIRS for retrieving and making available their respective products, and also to the teams of MERRA-2, ERA-Interim and CFSR for providing their modeled values. We also thanks the two anonymous referees for their suggestions to improve the manuscript

REFERENCES

American Meteorological Society AMS (2000), Glossary of Meteorology, 2nd ed., Boston, Mass.,

Albert, P., Bennartz, R., Preusker, R., Leinweber, R., and Fischer, J. (2005) Remote sensing of atmospheric water vapor using the Moderate Resolution Imaging Spectroradiometer, Journal of Atmospheric and Oceanic Technology, 22, 309-314.

Alraddawi, D., Sarkissian, A., Keckhut, P., Bock, O., Noël, S., Bekki, S., Irbah, A., Meftah, M., and Claud, C. (2018) Comparison of total water vapour content in the Arctic derived from GNSS, AIRS, MODIS and SCIAMCHY, *Atmospheric Measurement Techniques*, 11, 2949-2965.

Alexandrov, M. D., B. Schmid, D. D. Turner, B. Cairns, V. Oinas, A. A. Lacis, S. I. Gutman, E. R. Westwater, A. Smirnov, and J. Eilers (2009), Columnar water vapour retrievals from multifilter rotating shadowband radiometer data, *Journal of Geophysical Research*, 114, doi:10.1029/2008.jd010543.

Antón, M., Loyola, D., Román, R., and Vömel, H. (2015) Validation of GOME-2/MetOp-A total water vapour column using reference radiosonde data from the GRUAN network, Atmospheric Measurement Techniques, 8, 1135-1145.

Aumann, H. H., et al. (2003), AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems, *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 253-264

Barreto, A. E. Cuevas, B. Damiri, P. M. Romero, and F. Almansa (2013) Columnar water vapour determination in night period with a lunar photometer prototype, *Atmospheric Measurement Techniques*, 6, 2159-2167.

Berrisford, P., P. Kållberg, S. Kobayashi, D. Dee, S. Uppala, A. J. Simmons, P. Poli, and H. Sato (2011), Atmospheric conservation properties in ERA-Interim, *Quarterly Journal of the Royal Meteorological Society*, *137*(659), 1381-1399.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware (1992), GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System, *Journal of Geophysical Research*, 97, 15787-15801.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. H. P. (1999), SCIAMACHY - mission objectives and measurement modes. *Journal of Atmospheric Sciences,* 56, 127-150.

Burrows, J. P., et al. (1999), The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, *Journal of the Atmospheric Sciences*, *56*, 151-175.

Cadeddu, M. P., Liljegren, J. C. and D. D. Turner (2013), The Atmospheric Radiation Measurement (ARM) program network of microwave radiometers: instrumentation, data and retrievals, *Atmospheric Measurement Techniques*, 6, 2359-2372.

Deschamps, P.-Y., F.-M. Breon, M. Leroy, A. Podaire, A. Bricaud, J.-C. Buriez, and G. Seze (1994), The POLDER mission: instrument characteristics and scientific objectives, *IEEE Transactions on Geoscience and Remote Sensing*, *32*, 598-615.

Durre, I., R. S. Vose, and D. B. Wuertz (2006), Overview of the integrated global radiosonde archive, *J. Clim.,* 19, 53–68.

Fujita, M., and Sato, T. (2017) Observed behaviors of precipitable water vapour and precipitation intensity in response to upper air profiles estimated from surface air temperature, *Nature,* DOI:10.1038/s41598-017-04443-9

Gao, B.-C., and A. F. H. Goetz (1990), Column atmospheric water vapor and vegetation liquid water retrievals from airborne Imaging Spectrometer Data, *Journal of Geophysical Research*, *95*, 3549-3564.

Gao, B., et al., 2015. MODIS Atmosphere L2 Water Vapor Product. NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA: http://dx.doi.org/10.5067/MODIS/MOD05_L2.006

Gelaro, R., W. McCarty, M. J. Suarez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. Da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S. D. Schubert, M. Sienkiewicz, and B. Zhao (2017) The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *Journal of Climate*, 30, 5419-5454.

Gottwald, M., and Bovensmann, H. (2011), SCIAMACHY - Exploring the Changing Earth's Atmosphere. Springer, Dordrecht, Heidelberg, London, New York, doi: 10.1007/978-90-481-9896-2.

Halthore, R. N., T. F. Eck, N. B. Holben, and B. L. Markham (1997), Sun Photometric measurements of atmospheric water vapor, *Journal of Geophysical Research*, 102, 4343-4352.

Holben, B. N., D Tanre, A Smirnov, T. F Eck, I Slutsker, N Abuhassan, W. W Newcomb, J. S. Schafer, B. Chatenet, F. Lavenu, Y. J. Kaufman, J. Vande Castle, . Setzer, B Markham, D. Clark, R. Frouin, R. Halthore, A Karneli, N. T O'Neill, C Pietras, R. T. Pinker, K. Voss, G. Zibordi, 2001. An emerging ground-based climatology: Aerosol optical depth from AERONET. *Journal of Geophysical Research: Atmospheres*, 106, D11, 12067-12097.

Ichoku, C., R. Levy, Y. J. Kaufman, L. A. Remer, R.-R. Li, V. J. Martins, B. N. Holben, N. Abuhassan, I. Slutsker, T. F. Eck and C. Pietras (2002), Analysis of the performance characteristics of the five-channel Microtops II Sun photometer for measuring aerosol optical thickness and precipitable water vapour, *Journal of Geophysical Research*, 107, 4179.

Kaufman, Y. J., and B.-C. Gao (1992), Remote sensing of water vapor in the near IR from EOS/MODIS, *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 871-884.

King, M. D., Y. J. Kaufman, W. P. Menzel, and D. Tanre (1992), Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS), *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 2-27.

Kneizys, F. X., E. P. Shettle, L. W. Abreu, J. H. Chetwynd, G. P. Anderson, W. O. Gallery, J. E. A. Selby, and S. A. Clough, Users guide to LOWTRAN 7, AFGL–TR–8–0177, Air Force Geophys. Lab., Bedford, Mass., 1988.

Ichoku, C., Remer, L.A., Eck, T.F. (2005) Quantitative evaluation and intercomparison of morning and afternoon Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol measurements from Terra and Aqua, *Journal of Geophysical Research*, 110, D10S03

Leblanc, T., T. D. Walsh, I. S. McDermid, G. C. Toon, J.-F. Blavier, B. Haines, W. G. Read, B. Herman, E. Fetzer, S. Sander, T. Pongetti, D. N. Whiteman, T. G. McGee, L. Twigg, G. Sumnicht, D. Venable, M. Calhoun, A. Dirisu, D. Hurst, A. Jordan, E. Hall, L. Miloshevich, H. Vömel, C. Straub, N. Kampfer, G. E. Nedoluha, R. M. Gomez, K. Holub, S. Gutman, J. Braun, T. Vanhove, G. Stiller, and A. Hauchecorne (2011), Measurements of Humidity in the Atmosphere and Validation Experiments (MOHAVE)-2009: overview of campaign operations and results, *Atmospheric Measurement Techniques*, 4, 2579–2605.

Levy, R.C., Mattoo, S., Munchack, L.A., Remer, L.A., Sayer, A.M., Patadia, F., and Hsu, C. (2013), The Collection 6 MODIS aerosol products over land and ocean, *Atmospheric Measurements Techniques*, 6, 2989-3034.

Liu, H., Tang, S., Zhang, S., and Hu, J. (2015) Evaluation of MODIS water vapour products over China using radiosonde data, International Journal of Remote Sensing, 36, 680-690.

Martins, J.V., Tanré, D., Remer, L., Kaufman, Y., Mattoo, S., and Levy, R. (2002), MODIS cloud screening for remote sensing of aerosols over oceans using spatial variability, *Geophysical Research Letters*, 29, 1619.

Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A. P. (2006), GOME-2 on MetOp, in: Proc. of the 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, 12–16 June 2006, EUMETSAT P.48.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M. (2016) The

GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, *Atmos. Meas. Tech.,* 9, 1279-1301,

Nalli, N. R., J. Everette, V. N. Morris, C.D. Barnet, W. W. Wolf, D. Wolfe, P. J. Minnett, M. Szczodrak, M. A. Izaguille, R. Lumpkin, H. Xie, A. Smirnov, T. S. King and J. Wei (2011) Multiyear observations of the tropical Atlantic atmosphere, *Bulletin of the American Meterological Society*, 765-789.

Noël, S., M. Buchwitz, and J. P. Burrows (2004), First retrieval of global water vapour column amounts from SCIAMACHY measurements, *Atmospheric Chemistry and Physics*, *4*, 111-125.

Noël, S., S. Mieruch, H. Bovensmann, and J. P. Burrows (2008), Preliminary results of GOME-2 water vapour retrievals and first applications in polar regions, *Atmospheric Chemistry and Physics*, *8*, 1519-1529.

Noël, S., M. Buchwitz, H. Bovensmann, R. Hoogen, and J. P. Burrows (1999), Atmospheric water vapor amounts retrieved from GOME satellite data, *Geophysical Research Letters*, *26*, 1841-1844.

Palm, M., Melsheimer, C., Noël, S., Heise, S., Notholt,, J., Burrows, J., and Schrems, O. (2010) Integrated water vapor above Ny Alesund, Spitsberg: a multi-sensor intercomparison, Atmospheric Chemistry and Physics, 10, 1215-1226.

Pérez-Ramírez, D., F. Návas-Guzmán, H. Lyamani, J. Fernández-Gálvez, F. J. Olmo, and L. Alados-Arboledas (2012), Retrievals of precipitable water vapor using star photometry: assessment with Raman lidar and link to sun photometry, *Journal of Geophysical Research*, 117, D05202, doi:10.1029/2011JD016450.

Pérez-Ramírez, D., D. N. Whiteman, A. Smirnov, H. Lyamani, B. N. Holben, R. Pinker, M. Andrade, and L. Alados- Arboledas (2014), Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites, *J. Geophys. Res. Atmos.,* 119, 9596–9613.

Petrenko, M., C. Ichoku, and G. Leptoukh, (2012), Multi-sensor Aerosol Products Sampling System (MAPSS), *Atmos. Meas. Tech.,* 5, 913-926.

du Piesanie, A., Piters, A.J.M., Aben, I., Schrijver, H., Wang, P., and Noël, S. (2013), Validation of two independent retrievals of SCIAMACHY water vapour columns using radiosonde data. *Atmospheric Measurements Techniques*, 6, 2925-2640.

Qin, J., Yang, K., Koike, T., Lu, H., and Xu, X. (2012) Evaluation of AIRS precipitable water vapor against ground-based GPS measurements over the Tibetan Plateau and its surroundings, *Journal of the Meteorological Society of Japan*, 90, 87-98.

Reale, O., Susskind, J., Rosenberg, R., Brin, E., Liu, E., Riishojgaard, L.P., Terry, J., and Jusem, J.C. (2009), Improving forecast skill by assimilation of quality-controlled AIRS temperature retrievals under partially cloudy conditions, *Geophysical Research Letters,* 35, L08809.

Reagan, J. A., L. W. Thomason, B. M. Herman, and J. M. Palmer, (1986). Assessment of atmospheric limitations on the determination of the solar spectral constant from ground based spectroradiometer measurements. *IEEE Transaction on Geosciences and Remote Sensing*, GE-24, 258-265.

Reichard, J., U. Wandinger, M. Serwazi, and C. Weitkamp (1996), Combined Raman lidar for aerosol, ozone and moisture measurements, *Opt. Eng.,* 35, 1457–1465.

Román, R., Bilbao, J., de Miguel, A (2014) Uncertainty and variability in satellite-based water vapor column, aerosol optical depth and Angström exponent, and its effect on radiative transfer simulations in the Iberian Peninsula, *Atmospheric Environment*, 89, 556-569.

Román, R., Antón, M., Cachorro, V.E., Loyola, D., Ortiz de Galisteo, J.P., de Frutos, A., Romero-Campos, P.M. (2015) Comparison of total water vapor column from GOME-2 on MetOp-A against ground-based GPS measurements at the Iberian Peninsula, Science of the Total Environment, 533, 317-328.

Roman, J., Knuteson, R., August, T., Hultberg, T., Ackerman, S., and Revercomb, H., (2016) A global assessment of NASA AIRS v6 and EUMESAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations, Journal of Geophysical Research: Atmospheres, 121, 8925-8948.

Saha, S., et al. (2010), The NCEP climate forecast system reanalysis, *Bulletin of the American Meteorological Society*, 1015-1057

Schmid, B., J. J. Michalsky, D. W. Slater, J. C. Barnard, R. N. Halthore, J. C. Liljegren, B. N. Holben, T. F. Eck, J. M. Livingston, P. B. Russel, T. Ingold, and I. Slutsker (2001), Comparison of columnar water-vapor measurements from solar transmittance methods, *Applied Optics*, 40(12), 1886-1895.

Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker, (2000), Cloud-screening and quality control algorithms for the AERONET database, *Remote Sensing of Environment*, 73(3), 337–349, doi: 10.1016/S0034-4257(00)00109-7.

Smirnov, A., B. N. Holben, A. Lyapustin, I. Slutsker, and T. F. Eck (2004), AERONET processing algorithms refinement, Proceedings of AERONET workshop, El Arenosillo, Spain, NASA/GSFC Aeronet project.

Smirnov, A., M. Petrenko, C. Ichoku, and B. N. Holben (2017) Maritime Aerosol Network optical depth measurements and comparison with satellite retrievals from various different sensors. *Proc. of SPIE*, Vol. 10424, 1042402-1, In Remote Sensing of Clouds and the Atmosphere XXII, edited by Adolfo Comerón, Evgueni I. Kassianov, Klaus Schäfer, Richard H. Picard, Konradin Weber.

Smirnov, A., B. N. Holben, I. Slutsker, D. M. Giles, C. R. McClain, T. F. Eck, S. M. Sakerin, A. Macke, P. Croot, G. Zibordi, P. K. Quinn, J. Sciare, S. Kinne, M. Harvey, T. J. Smyth, S. Piketh, T. Zielinski, A. Proshutinsky, J. I. Goes, N. B. Nelson, P. Larouche, V. F. Radionov, P. Goloub, K. Krishna Moorthy, R. Matarrese, E. J. Robertson, and F. Jourdin (2011) Marine Aerosol Network as a component of Aerosol Robotic Network, *Journal of Geophysical Research: Atmospheres*, 114, D06204.

Smirnov, A., B. N. Holben, D. M. Giles, I. Slutsker, N. T. O'Neill, T. F. Eck, A. Macke, P. Croot, Y. Courcoux, S. M. Sakerin, T. J. Smyth, T. Zielinski, G. Zibordi, J. I. Goes, M. J. Harvey, P. K. Quinn, N. B. Nelson, V. F. Radionov, C. M. Duarte, R. Losno, J. Sciare, K. J. Voss, S. Kinne, N. R. Nalli, E. Joseph, K. Krishna Moorthy, D. S. Covert, S. K. Gulev, G. Milinevsky, P. Larouche, S. Belanger, E. Horne, M. Chin, L. A. Remer, R. A. Kahn, J. S. Reid, M. Schulz, C. L. Heald, J. Zhang, K. Lapina, R. G. Kleidman, J. Griesfeller, B. J. Gaitley, Q. Tan, and T. L. Diehl (2011) Maritime aerosol network as a component of AERONET- first results and comparison with global aerosol models and satellite retrievals, *Atmospheric Measurement Techniques*, 4, 583-597.

Schmitt, R. W., (2008) Salinity and the global water cycle. *Oceanography*, **21**, 12–19.

Susskind, J., C. D. Barnet, and J. M. Blaisdell (2003), Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 390-409.

Susskind, J., Barnet, C., Blaisdell, J., Iredell, L., Keita, F., Kouvaris, L., Molnar, G., and Chahinnes, M. (2006), Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fraction cloud cover, *Journal of Geophysical Research*, 111, D09S17.

Susskind, J., Blaisdell, J.M., Iredell, L., and Keita, F. (2011), Improved temperature sounding and quality control methodology using AIRS/AMSU data: the AIRS Science Team Version 5 retrieval algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 49, 883-907.

Susskind, J., Blaisdell, J.M., and Iredell, L. (2014), Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm, *Journal of Applied Remote Sensing*, 8(1), 084994.

Trenberth, K. E., and D. P. Stepaniak (2003a), Seamless poleward atmospheric energy transports and implications for the Hadley circulation, *Journal of Climate*, 16, 3705-3721.

Trenberth, K. E., P. D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J. A. Renwick, M. Rusticucci, B. Soden, and P. Zhai (2007), Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S.,D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Vaquero-Martínez, J., Antón, M., Ortiz de Galisteo, J.P., Cachorro, V.E., Álvarez-Zapatero, P., Román, R., Loyola, D., Costa, M.J., Wang, H., González-Abad, G., and Noël, S. (2018) Inter-comparison of integrated water vapor from satellite instruments using reference GPS data at the Iberian Peninsula, *Remote Sensing of Environment*, 204, 729-740.

Veselovskii, I., Whiteman, D.N., Kolgotin, A., Andrews, E., and Korenskii, M. (2009) Demonstration of aerosol property profiling by multiwavelength lidar under varying relative humidity conditions, Journal of Atmospheric and Oceanic Technology, 26 1543-1557.

Vesperini, M., F.-M. Breon, and D. Tanre (1999), Atmospheric water vapor content from spaceborne POLDER measurements, *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 1613-1619.

Whiteman, D., K. Rush, S. Rabenhorst, W. Welch, M. Cadirola, G. McIntire, F. Russo, M. Adam, D. Venable, R. Connell, I. Veselovskii, R. Forno, B. Mielke, B. Stein, T. Leblanc, S. McDermid and H. Vömel (2010), Airborne and ground-based measurements using a high-performance Raman Lidar, *J. Atmos. Ocean. Technol*., 27, 1781-1801. doi: 10.1175/2010JTECH1391.1.

Whiteman, D. N., M. Cadirola, D. Venable, M. Calhoun, L. Miloshevich, K. Vermeesch, L. Twigg, A. Dirisu, D. Hurst, E. Hall, A. Jordan, and H. Vömel (2012), Correction technique for Raman water vapor lidar signal-dependent bias and suitability for water vapor trend monitoring in the upper troposphere, *Atmospheric Measurement Techniques,* 5, 2893–2916, doi: 10.5194/amt-5-2893-2012.

Wagner, T., Heland, J., Zöger, M., and Platt, U. (2003), A fast H2O total column density product from GOME – Validation with in situ aircraft measurements, *Atmos. Chem. Phys., 3, 651–663,*

Wagner, T., Beirle, S., Grzegorski, M., and Platt, U. (2006) Global trends (1996–2003) of total column precipitable water observed by Global Ozone Monitoring Experiment (GOME) on ERS-2 and their relation to near-surface temperature, *J. Geophys. Res*. 111,

Wulfmeyer, V., R. M. Hardesty,D. D. Turner, A. Behrendt, M. P.Cadeddu, P. Di Girolamo, P. Schlüssel, J. Van Baelen, and F. Zus (2015), A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles, *Rev. Geophys*., 53, 819–895.

Ye, H., Fetzer, E.J., Wong, S., Behrangi, A., Olsen, E.T., Cohen, J., Lambrigtsen, B., and Chen, L. (2016), Impact of increased water vapor on precipitation efficiency over northern Eurasia, *Geophysical Research Letters,* 41, 2941-2947.

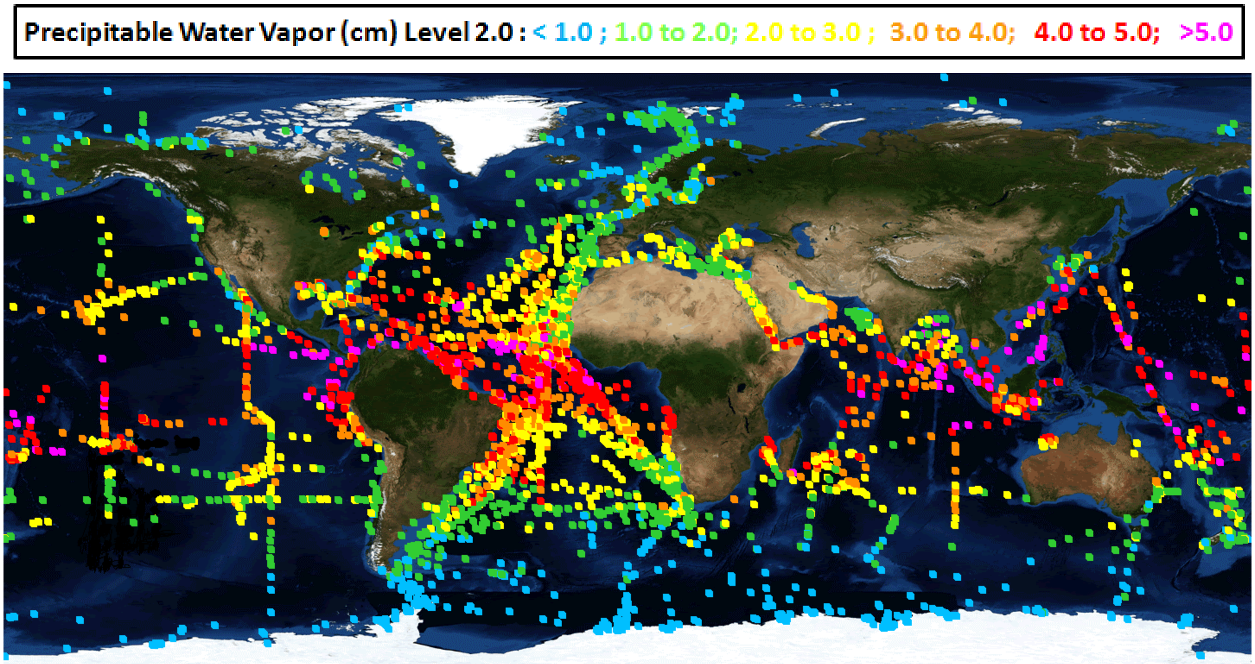
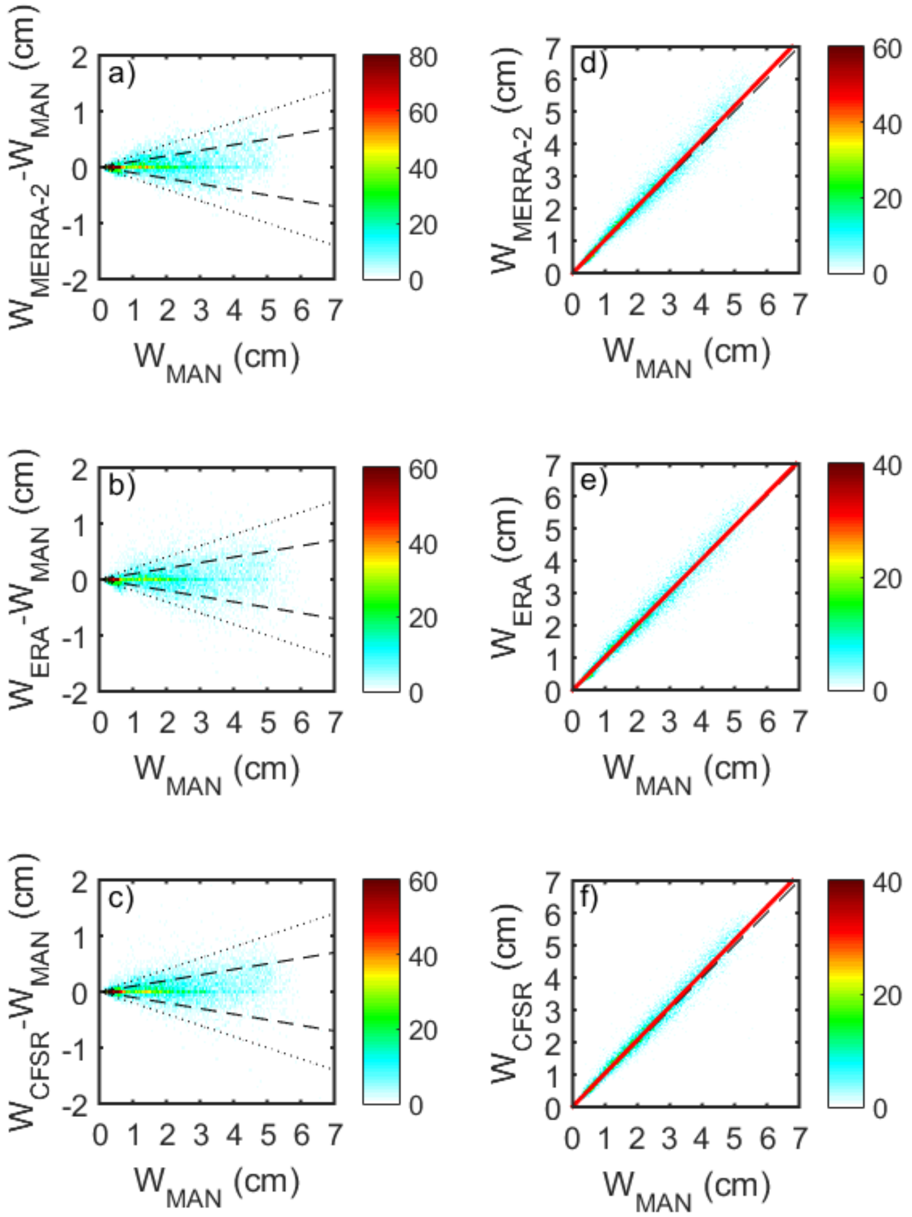
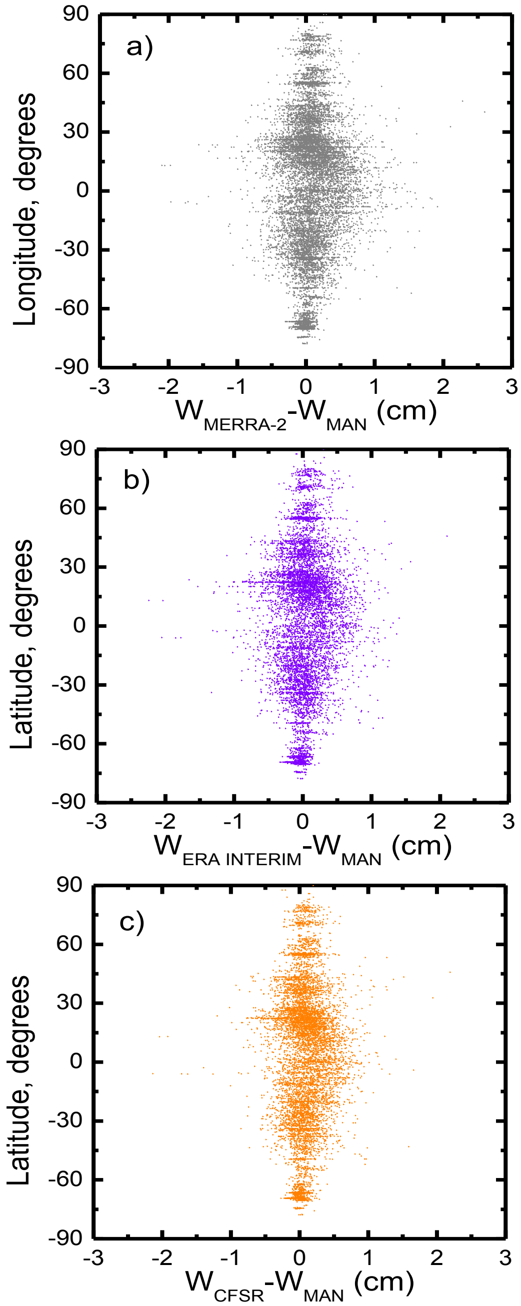


Figure 1: Level 2.0 Marine Aerosol Network global coverage, showing the cruise tracks and corresponding daily averages of precipitable water vapor (W). Squares representing the average daily sampling locations are color-coded with respect to W values, i.e. blue for W < 1.0 cm, green for 1.0 ≤ W < 2.0 cm, yellow for 2.0 ≤ W < 3.0 cm, orange for 3.0 ≤ W < 4.0 cm, red for 4.0 ≤ W < 5.0 cm, and purple for W ≥ 5.0 cm.



<u>Figure 2:</u> Number density plots of (a)-(c) differences in W between models and MAN data as a function W measured by MAN (W_{MAN}). Dashed lines represent ±10% difference versus measured W_{MAN}while the dot lines represent ±20% differences. (d)-(f) W by global reanalysis models versus W_{MAN}. Red lines are the results of the least-square fits and dashed lines are the 1:1 line



<u>Figure 3:</u> Differences with latitude in precipitable water vapor (*W*) between global reanalysis models and Marine Aerosol Network (MAN) data

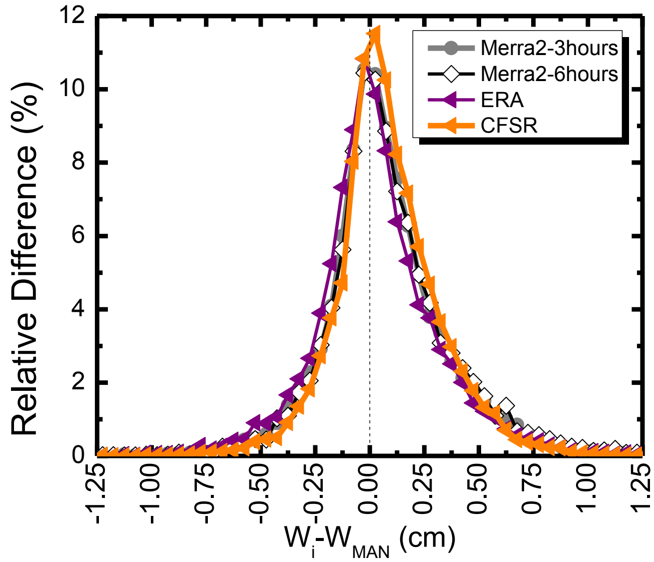
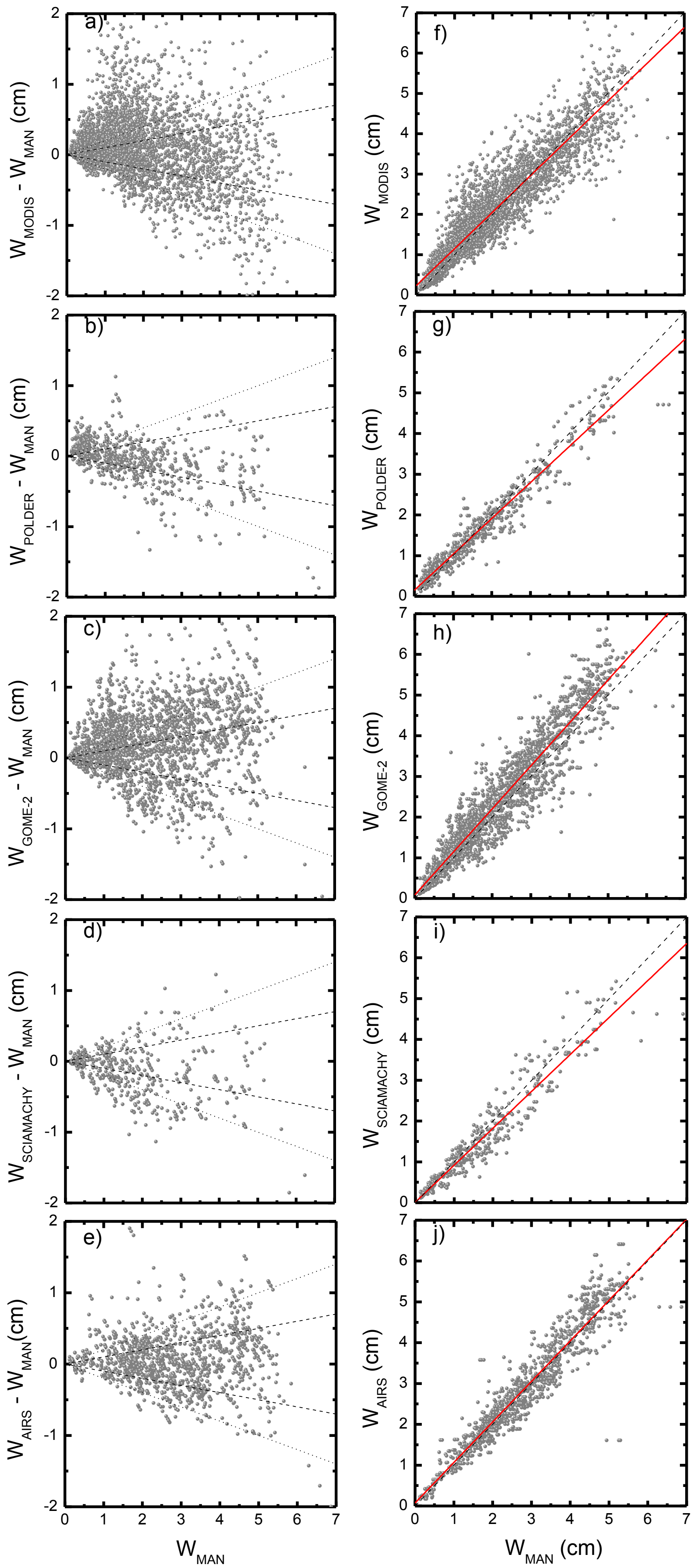


Figure 4: Frequency histograms of the differences in precipitable water vapor (W) between Marine Aerosol Network (MAN) and global models. Total number of data for each dataset are given in Table 2.



a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
W_MODIS - W_MAN (cm)
W_POLDER - W_MAN (cm)
W_GOME-2 - W_MAN (cm)
W_SCIAMACHY - W_MAN (cm)
W_AIRS - W_MAN(cm)
W_MODIS (cm)
W_POLDER (cm)
W_GOME-2 (cm)
W_SCIAMACHY (cm)
W_AIRS (cm)
W_MAN
W_MAN (cm)

Figure 5: (a) – (e) Differences in precipitable water vapor between satellite sensors and Marine aerosol Network data (W_{MAN}) as function of W_{MAN}. Dash lines represents 10% uncertainties and dotted lines represents 20% uncertainties. (f)-(j) Precipitable water vapor from satellite sensors as function of W_{MAN}. Dashed lines represent 1:1 line, while red lines are the values from the least-squares fits.

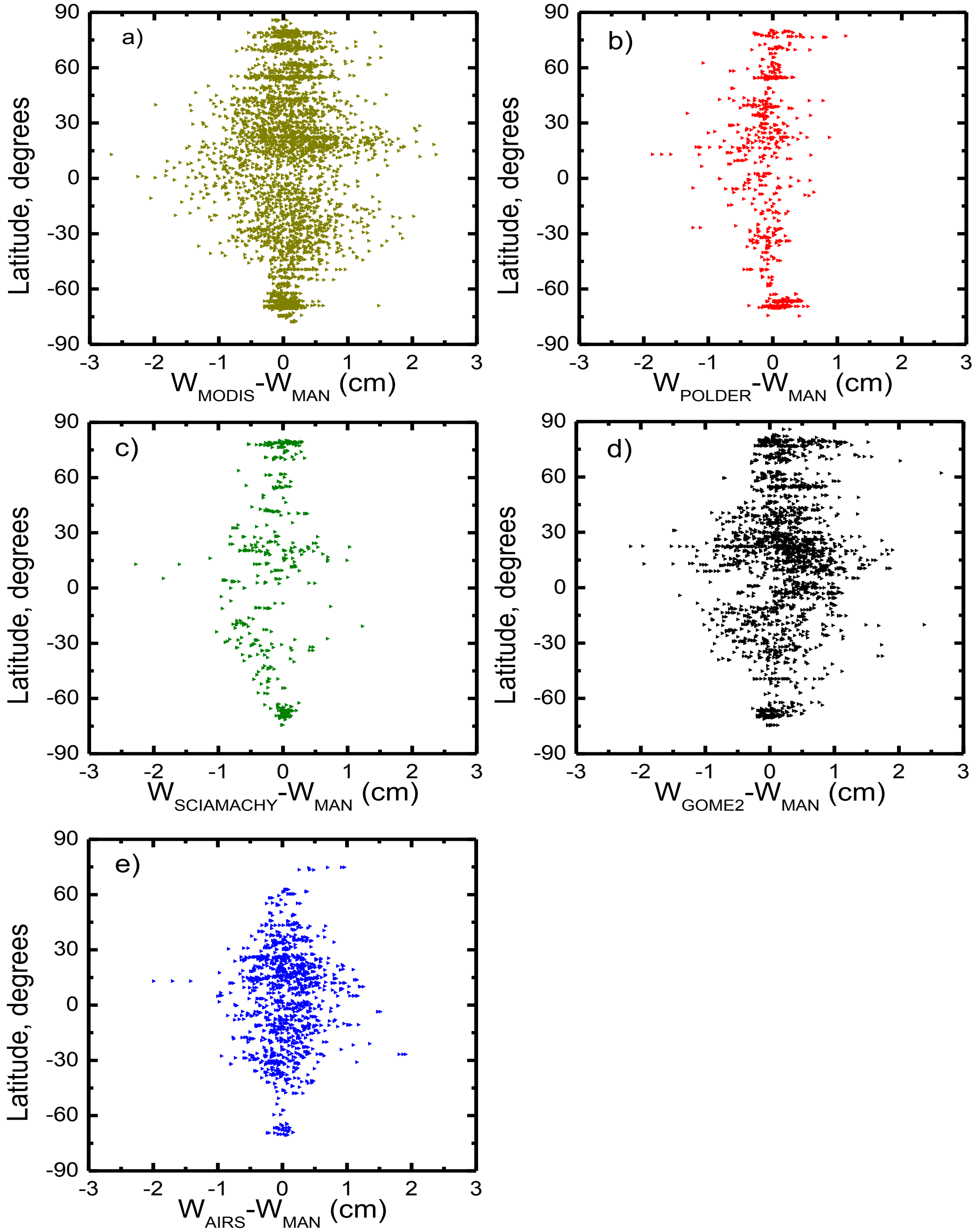


Figure 6: Differences with latitude in precipitable water vapor (W) between satellite sensors and Marine Aerosol Network (MAN) data.

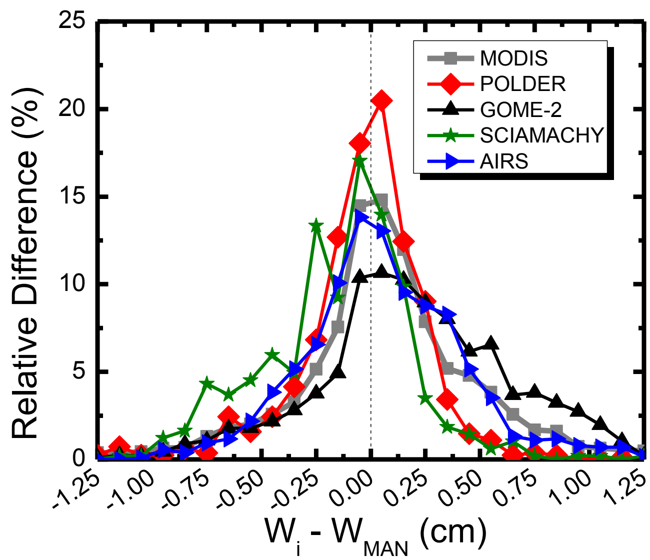


Figure 7: Frequency histograms of the differences in precipitable water vapor W between Marine Aerosol Network (MAN) and satellite sensors. Total number of data for each dataset are given in Table 2.

	Name	Institute/ Platform	Spatial Resolution	Data Period	W Estimation Approach	References
Global Model	MERRA-2	NASA GMAO	0.50º x 0.625º – 72 level heights	1980 - present	Reanalysis based on the assimilation of meteorological data obtained from different satellite sensors.	*Gelaro et al., 2017*
	ERA Interim	ECMWF	0.75º x 0.75° - 40 levels heights			*Berrisford et al.*, 2011
	CFSR	NCEP	0.50º x 0.50º - 60 levels height			*Saha et al., 2010*
Satellite Sensor	MODIS	Terra and Aqua	Infrared Approach 5 x 5 km^2	1999 - present	Ratio of signals in the infrared (absorption and no absorption water vapor bands)	*Kaufman and Gao, 1992*
	POLDER	PARASOL	50 x 50 km^2	2004 - 2013		*Vesperini et al.*, 1999
	GOME-2	MetOp-A and MetOp-B	80 x 40 km^2	2006 - present	DOAS technique that fits differential structures of the measured spectral reflectance	*Nöel et al.*, 1999, 2004, 2008
	SCIAMACHY	ENVISAT	60 x 30 km^2	2002-2012		
	AIRS	Aqua	50 x 50 km^2	2002 - present	Microwave Radiometry	*Susskind et al.*, 2003

Table 1: Summary of global models and satellite sensors whose precipitable water vapor (*W*) datasets are evaluated in this paper using MAN measurements.

	Global Model / Satellite Sensor	N	W_i - W_{MAN} (cm)			(W_i - W_{MAN})/ W_{MAN} (%)			$W_i = AW_{MAN} + B$		
			Mean	**STD**	**Median**	**Mean**	**STD**	**Median**	**A**	**B (cm)**	**R2**
Model	Merra-2	12523	0.07	0.30	0.04	2.8	14.1	2.6	1.03	-0.001	0.957
	ERA Interim	8520	0.03	0.29	0.01	0.9	14.7	0.7	1.01	-0.001	0.956
	CFSR	8760	0.08	0.26	0.06	3.9	13.0	3.5	1.03	0.014	0.967
Satellite Sensor	MODIS	3920	0.08	0.48	0.05	10.8	30.9	5.0	0.92	0.23	0.874
	POLDER	820	-0.04	0.31	-0.01	6.7	29.0	-0.3	0.88	0.15	0.945
	GOME-2	1706	0.21	0.49	0.18	12.4	28.3	10.3	1.06	0.09	0.897
	SCIAMACHY	487	-0.16	0.36	-0.10	-7.2	19.7	-7.3	0.91	0.01	0.920
	AIRS	1280	0.05	0.42	0.03	3.1	17.3	1.5	0.99	0.07	0.899

Table 2: Statistical parameters for the evaluations of precipitable water vapor (W) of different global models and satellite sensors versus the Marine Aerosol Network (MAN). The total number of points 'N' for the intercomparisons of model/satellite sensor with MAN is given. Mean, median and standard deviations (STD) are included. Also the parameters of the linear fits are provided, being the coefficient 'A' the slope of the linear fit and 'B' is the ordinate intercept.