Impact of humidity biases on light precipitation occurrence : observations versus simulations

Sophie Bastin[1], Philippe Drobinski[2], Marjolaine Chiriaco[1], Olivier Bock[3], Romain Roehrig[4], Clemente Gallardo[5], Dario Conte[6], Marta Domínguez Alonso[5,6], Laurent Li[2], Piero Lionello[7,8], Ana C. Parracho[1,3]

[1] LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, France
[2] LMD/IPSL, Ecole Polytechnique, Sorbonne Université, ENS, CNRS, Palaiseau, France
[3] IPGP, IGN, ENSG, Université Paris Diderot, Sorbonne Paris Cité, UMR 7154 CNRS, Paris, France
[4] CNRM, Université de Toulouse, Meteo-France, CNRS, Toulouse, France
[5] UCLM, Instituto de Ciencias Ambientales, Toledo, Spain
[6] AEMET, Agencia Estatal de Meteorología, Madrid, Spain
[7] CMCC, Euro Mediterranean Center on Climage Change, 73100 Lecce, Italy
[8] DiSTeBA, University of Salento, 73100 Lecce, Italy

Abstract. This work uses a network of GPS stations over Europe from which a homogenised integrated water vapor (IWV) dataset has been retrieved, completed with colocated temperature and precipitation measurements over specific stations to i) estimate the biases of six regional climate models over Europe in terms of humidity ; ii) understand their origins ; iii) and finally assess the impact of these biases on the frequency of occurrence of precipitation. The evaluated simulations have been performed in the framework of HYMEX/Med-CORDEX programs and cover the Mediterranean area and part of Europe at horizontal resolutions of 50 to 12 km.

The analysis shows that models tend to overestimate the low values of IWV and the use of the nudging technique reduces the differences between GPS and simulated IWV. Results suggest that physics of models mostly explain the mean biases, while dynamics affects the variability. The land surface/atmosphere exchanges affect the estimation of IWV over most part of Europe, especially in summer. The limitations of the models to represent these processes explain part of their biases in IWV. However, models correctly simulate the dependance between IWV and temperature, and specifically the deviation that this relationship experiences regarding the Clausius-Clapeyron law after a critical value of temperature (T_{break}). The high spatial variability of T_{break} indicates that it has a strong dependence on local processes which drive the local humidity sources. This explains why the maximum values of IWV are not necessarely observed over warmer area, that are often dry area.

Finally, it is shown over SIRTA observatory (near Paris) that the frequency of occurrence of light precipitation is strongly conditioned by the biases in IWV and by the precision of the models to reproduce the distribution of IWV as a function of the temperature. The results of the models indicate that a similar dependence occurs in other areas of Europe, especially where precipitation has a predominantly convective character. According to the observations, for each range of temperature, there is a critical value of IWV from which precipitation starts to increase. The critical values and the probability to exceed them are simulated with a bias that depends on the

model. Those models which present too often light precipitation generally show lower critical values and higher probability to exceed them.

1 Introduction

Humidity plays a major role in the water and energy cycles due to its strong radiative effect associated with a positive feedback on climate (Randall et al., 2007) and its importance to control precipitation and particularly extreme ones (Held and Soden, 2006 ; Neelin et al., 2009 ; Sahani et al., 2012). Trends and variability of humidity and precipitation are strongly correlated (Trenberth et al., 2003 ; Zhang et al., 2013) and several studies have revealed that the rate of increase in daily extreme precipitation is highly connected with the warming following the Clausius-Clapeyron (C-C) relation (Allen and Ingram, 2002; Pall et al., 2007; Kharin et al., 2007). This rate of precipitation is indeed affected by the humidity content of the atmosphere (Integrated water Vapor (IWV)), which rises as climate warms (e.g Trenberth, 2011). Nevertheless, dynamical processes (O'Gorman and Schneider, 2009; Sugiyama et al., 2010; Singleton and Toumi, 2013; Muller, 2013 ; Drobinski et al., 2016a), lack of humidity sources leading to a decrease of relative humidity (Drobinski et al., 2016b), or low/high precipitation efficiency (Drobinski et al., 2016a ; Trenberth et al., 2003) can explain the deviation from C-C rate locally. Humidity variability at regional scale -and not only at the surface - thus needs to be assessed to better anticipate the precipitation change, and more specifically the rate of heavy precipitation, that are not well estimated by global models (e.g Allan and Soden, 2008).

Another aspect that links IWV and precipitation concerns the triggering of precipitation and thus the frequency of occurrence of precipitation: Holloway and Neelin (2008) showed that precipitation over the tropical oceans is strongly sensitive to free-tropospheric humidity even more than surface humidity, and Neelin et al. (2009) and Sahany et al. (2012) further conclude that there exists a threshold of IWV, which depends on the mean tropospheric temperature, over which precipitation starts to increase significantly. They also showed that this critical value of IWV does not correspond to the saturation value when temperature increases, i.e that at higher temperature, deep convection occurs at a lower value of relative humidity. This means that IWV is a relevant parameter to measure over long-term periods, at high temporal resolution and at the regional scale in order to establish the relationship between IWV-precipitation and temperature and monitor its possible evolution. Models still have strong difficulties to simulate adequately the water cycle (Trenberth et al., 2003, Flato et al., 2013), and often presents the « too often too light precipitation » problem (e.g Sun et al., 2006 ; Panthou et al., 2016). A better knowledge of the IWV-precipitation relationship would be a help to better constrain models.

Up to now, very few long-term (> 15 years) and homogeneous datasets of water vapor measurements exist, even less at sub-daily time scales. These datasets are necessary to understand the humidity variability at regional scales at different time scales. Besides, the co-location of such measurements with independent measurements of precipitation and vertical profiles of temperature provide a strong added-value for better climate understanding. Reanalyses are of course a good tool to have these three parameters co-located over long-term and at sub-daily time scales, however precipitation mostly relies on the model physics. Moreover, Flato et al. (2013) have shown that even in reanalyses, the relationship between the IWV trend and the temperature trend presents differences between reanalyses and deviates from C-C over tropical oceans.

In this study, we make use of the Global Positioning Sysem (GPS) IWV dataset that has been processed as done by Parracho et al., (2018) and which provides IWV measurements over a hundred of European sites covering a period of 5 years or more. The GPS technique accurately measures IWV (accuracy around 1-2 $kg.m^{-2}$ according to Bock et al., 2005 ; 2013 and Ning et al., 2016) in all weather conditions including rainy situations, which is an

important aspect for our study (e.g Wang et al. 2007). GPS measurements have been successfully used to better understand atmopsheric processes at high resolution (Bastin et al., 2005, 2007 ; Bock et al., 2008 ; Champollion et al., 2009). Here we use a GPS IWV dataset to analyse the humidity biases in regional climate models over Europe at interannueal, seasonal and daily time scales and to better understand the source of errors of models. We also use GPS IWV measurements co-located with precipitation and tropospheric temperature measurements from the SIRTA observatory in France to consider the relationships between these parameters. Note they have not yet been considered outside the tropics. We compare observations and regional models output at the site level, and extend the analysis of models to other locations over Europe.

The paper is organised as follow: section 2 presents the observational datasets and the different simulations used in this study. Section 3 describes the methodology to compare observations and models. In section 4, the ability of models to reproduce the mean value of humidity and its variability over Europe at different time scales is evaluated. The influence of dynamical and physical processes is discussed, and a special focus on the scaling of IWV with temperature is developed. In section 5, the issue of how much a bias in IWV can enhance the problem of 'too often too light precipitation' behavior of models is raised by considering the relationship between mean tropospheric temperature, IWV and precipitation in the different models and observations over the SIRTA supersite in France. Then, the generalization of this relationship, by considering other stations over Europe is assessed. Finally, a conclusion is given in section 6.

2 Material

2.1 GPS IWV data

The GPS dataset used in this study is based on homogeneously reprocessed GPS delay data produced by the NASA's Jet Propulsion Laboratory in the framework of the first International GNSS Service (IGS) reprocessing campaign. The data cover the period from January 1995 to May 2011 and include more than 400 stations globally. For the present study, the delay data were screened and converted into 6-hourly IWV estimates as described in Parracho et al. (2018). Then, daily values are computed using this 6-hourly dataset. The dataset was restricted to the period from January 1995 to December 2008 and includes 95 GPS stations over Europe as shown in Fig. 1. It represents much more stations than in the study of Parracho et al. (2018) because for the purpose of their study which estimates IWV trends, they restricted their selection to stations with only small gaps over the 15-year period from 1995 to 2010. Here, only stations with less than 5 years of observations are not considered in the evaluation of models humidity bias.

2.2 Observations at SIRTA

This study also uses observations collected at the SIRTA atmospheric observatory, located 20-km South West of Paris (2.2°E/48.7°N 160 m of altitude; black triangle on Fig. 1), from 2003 (Haeffelin et al. 2005). This observatory has collected many observations, which are now synthesized into the so-called "SIRTA-ReOBS dataset" described in Chiriaco et al. (2018) and used in Chiriaco et al. (2014) and Bastin et al. (2016). After

many steps of data quality control and harmonization, the "SIRTA-ReOBS" file contains hourly averages of more than 50 variables at this site. The sample of data varies from one variable to another. Among these variables, the IWV retrieved from GPS measurement since 2008 and the precipitation rate from a single raingauge over 2003-now are available. A regional-scale precipitation estimate, deduced from the measurements of other Météo-France raingauges located around Paris area, is also provided.

The Météo-France COMEPHORE (« COmbinaison en vue de la Meilleure Estimation de la Precipitation HOraiRE ») product is used to allow a fairer intercomparison between models and observations than the single raingauge (Chen and Knutson, 2008): it is a hourly reanalysis of precipitation by merging radar data and raingauges over France at 1km x1km resolution (more details in Fumiere et al., in revision ; see also Laurantin et al., ERAD 2012). From this product, we can have a better knowledge of the average precipitation rate over a model grid of 50 km x 50 km or higher resolution. However, this product only covers the period 1997-2007 (at the time of this study) and is not concomitent with the IWV dataset over SIRTA. Despite this, a comparison has been made between the statistics of the different datasets when possible (see section 3.2).

The Météo-France radiosoundings launched twice a day from Trappes (near 00 UTC and 12 UTC), 15 km to the west of SIRTA are also used to compute the mean tropospheric temperature (more details in section 3).

2.3 Med-CORDEX simulations

The list of regional climate models (RCMs) and details about the settings are given in Table 1. All the simulations use the 6-hourly European Center for Medium-Range Weather Forecast (ECMWF) reanalyses ERA-Interim (Dee et al., 2011) as RCM boundary conditions. They cover at least the period 1989–2008 as initially recommended in MED-CORDEX project (Ruti et al. 2015). For LMDZ, which is a global model with regional zoom capability, temperature, wind speed and specific humidity are nudged towards the ERA-Interim fields outside the MED-CORDEX domain. It must be noted that the mesh of LMDZ is not regular within the zoom region and the resolution varies between 50 km and 30 km. All other RCMs are forced at the boundaries using 3-dimensional re-analyses of wind, humidity, temperature or potential temperature and geopotential height. For CCLM, cloud ice and liquid water are additionally prescribed at the domain boundaries. The IPSL WRF simulation uses nudging at all scales within the domain for temperature, wind and humidity above the planetary boundary layer (Salameh et al. 2010; Omrani et al. 2013, 2015). The other models did not use nudging in the Med-CORDEX domain.

Simulations used here were produced from five models (ALADIN V5.2, Colin et al. 2010; CCLM, Rockel et al. 2008; WRF V3.1.1, Skamarock et al. 2008; LMDZ V4, Hourdin et al. 2006; PROMES, Dominguez et al., 2010 ; 2013) with a horizontal resolution around 50 km (0.44° for most models) on the MED-CORDEX domain. In the following the simulations are referred to by the name of the modeling group and resolution (see Table 1, first column). All the models provide daily values of IWV, precipitation, 2-m temperature, and temperature at 850, 500 and 200 hPa.

2.4 Others datasets

The GPS observations are supplemented by the HadISD v.2.0.1.2016 subdaily dataset of surface parameters (e.g temperature, dew point temperature, wind, pressure - Dunn et al., 2012). It is global and based on the Integrated Surface Database (ISD) dataset from NOAA's National Climatic Data Center. Stations were selected on the basis of their length of record and reporting frequency before they are passed through a suite of quality control tests. It is a joint effort from the MetOffice Hadley Center and the National Center for Atmospheric research (NCAR).

3 Methods

To compare models and observations, we consider differences as the 'model minus observation' results throughout the manuscript.

3.1 Comparison between GPS dataset and model outputs

Each modeling group has provided a file containing the gridded IWV over the 1995-2008 period at daily resolution on its native grid. The IWV is either computed online or offline the model. The offline computation can introduce some errors due to vertical integration over the discretized vertical grid. For each model, we extracted the value of IWV at the closest grid point of GPS stations. We did it using the native grid, and also after having regridded all the model outputs to ERA-Interim (ERAI) grid, to have a fair comparison with ERAI when necessary. The difference of altitude between the GPS station and the closest grid point is difficult to take into account and can introduce strong bias over complex terrain (Hagemann et al., 2003; Zhang and Wang, 2009). As a consequence, the stations where the difference of altitude is higher than 500m were removed from the analysis. Note the number of stations that are removed depends on each model since the models do not use the same topography and the same projection. Then, a linear correction is applied on model outputs to reduce the bias due to orography: for each model and each month, we plotted the difference of the monthly averaged IWV values between the model and GPS as a function of the difference of altitude and we concluded that a linear correction can be applied to take into account the difference of altitude. For each different month, the slope of the linear regression between these two differences is computed and we apply the corresponding correction to IWV values of the model. The values of the slope of the linear regression for each model and each month are indicated in Table 2, both for the native grid of the model and for the regridded outputs. Various evaluation metrics (Table 3) have been computed with and without correction for these two grids. The correction does not impact the variability scores (e.g interannual variability), but does affect the mean bias (not always by an improvement) and slightly reduces the standard deviation of the difference (Table 3). It does not affect the ranking of performance between models.

Note that at SIRTA, the difference of altitude is weak and results are thus not impacted by this problem.

3.2 Comparison with SIRTA observations

i) Tropospheric temperature: to be as consistent as possible between model outputs and radiosoundings, the mean tropospheric temperature corresponds here to the daily average value of 2m-temperature, and temperatures at 850 hPa, 500 hPa and 200 hPa. So we extracted the temperature values at these pressure levels for each radiosounding launched from Trappes (a few kilometers away from SIRTA) and computed the mean tropospheric temperature as the average of these 4 values. Then, the daily mean is the average of the two daily radiosoundings. The same method is applied to ERA-Interim reanalysis to compute the mean tropospheric temperature over the other European sites. The impact of using ERA-Interim instead of radiosoudings data has been evaluated at SIRTA where both are available. Due to the width of temperature bins considered in this study, results are not sensitive to the use of one or another.

ii) Precipitation: a first step consists of comparing the precipitation statistics obtained using the single raingauge located at SIRTA, from the closest grid point (1 km*1 km) of COMEPHORE over the SIRTA and an average of COMEPHORE over an area centered over the SIRTA and covering 2500 km^2 (i.e a 50-km resolution model grid point). Figure 2 shows the mean annual cycle of the frequency of occurrence of different light precipitation regimes. The annual cycle is computed from 30-day means from January to December over the years 2004-2007, which is the common period of all datasets. The SIRTA raingauge and COMEPHORE product do not indicate the same frequency of occurrence of non-precipitating days, and very light precipitations, but they show similar frequency of occurrence for light precipitation which corresponds mainly to large scale precipitation. Note the different values of the 50^{th} percentile of precipitation when considering only precipitating days, which emphasize the impact of heavy precipitation on the estimate of this index. The SIRTA raingauge estimate is between the estimate by COMEPHORE at the two different resolutions. The difference between the frequency of occurrence of non–precipitating days (higher with COMEPHORE at the closest grid point) and that of very light precipitation (higher with SIRTA raingauge) likely come from the coarse resolution of the coding of reflectivity data of the radars used at low levels which is a limiting factor for the precise estimation of precipitation at low rain rates (Laurentin et al., 2012). The difference is stronger in winter than in summer. The average over 50*50 km^2 grid cell shows a decrease in the number of non-precipitating days and an increase in the occurrence of very light precipitations: it is expected since the probability that a system passes through a wider area is higher but the total precipitation averaged over a wide area is often much weaker as the strongest precipitation rates are generally localized.

Table 4 presents an estimate of the frequencies of occurrence of non-precipitating days, very light precipitation and light precipitation for winter (julian day 1 to 100) and summer (julian day 151 to 251) when considering either the common period of the two datasets (2004-2007) or the full period of each dataset (i.e 1997-2007 for COMEPHORE and 2003-2015 for SIRTA raingauge). The number of dry days increases for the two estimations from COMEPHORE when considering a longer period than the common period. For ReOBS, the statistics remain similar for the two different periods. However, when considering the most recent years only, that will be used in the next section (2008-2015), the number of dry days increases. The influence of the number of years, the years considered and the products used to estimate these frequencies of occurrences is generally small but significant. Even though model errors are most of the time beyond this uncertainty, it has to be kept in mind in the following analysis.

3.3 Method to establish the relationship between IWV and precipitation as a function of tropospheric temperature

The objective is to characterize how precipitation depends on IWV for different range of mean tropospheric temperature. To do that, we divided our datasets into four different bins of temperature : the first bin is for temperature less or equal to 254.5 K, the second bin if for temperature between 254.5 K and 258 K, the third one corresponds to temperature between 258 K and 262.5 K and the fourth one for temperature higher than 262.5 K. This choice has been made to ensure a high number of samples in each bin of temperature to proceed to the next step. Doing this way, the mean tropospheric temperature of one bin is similar more or less 1 K for all the datasets. Then, in each bin of temperature, the daily mean precipitation rates are sorted according increasing values of daily mean IWV. IWV bins are then defined such that they contain an equal number of pairs of precipitation rates-IWV (40 samples) to ensure a reasonable number of days to compute the 50th quantile of precipitation which indicates if there are more precipitating days than non-precipitating days. The range of each IWV bin is thus not constant but allows to identify quite easily the transition between mostly non-precipitating days and mostly precipitating days. For each model and for observations, a critical value of IWV (w_c) is then determined for each different bin of temperature by using a very simple algorithm. This one identifies the value(s) of IWV over which 50th quantile of precipitation is greater than 0.1 mm/day. In some cases, two values are obtained, that are represented by an errorbar to indicate the uncertainty of the estimate of this critical value.

4 Humidity biases in the Med-CORDEX simulations

4.1 Comparison with GPS dataset at regional scale

Figure 1 indicates the mean values of IWV retrieved from GPS measurements in winter (Fig.1a) and in summer (Fig.1b). In winter, higher values are observed along the Atltantic and Mediterranean coasts while Central and Eastern Europe exhibit very low values of IWV. In summer, there are two different regimes: i) north of 45°N showing a decrease of IWV values while going to Scandinavia, following the temperature gradient; ii) around the Mediterranean, the structure is more patchy with an alternance of low and high values. Most low values correspond to higher topography but not systematically.

Table 3 shows the mean bias and the standard deviation (SD) of ERA-interim and RCMs IWV on their native grid or regridded to ERAI grid, in comparison to GPS estimates. Daily datasets are used to compute these statistics and exactly the same sampling is used between models and observations. It shows that the mean bias ranges from 0.5 $kg.m^{-2}$ to 1.0 $kg.m^{-2}$ except for CNRM50 and UCLM50 for which the biases are stronger (up to 1.8 $kg.m^{-2}$ for UCLM50 on its native grid). It is to be noted that the mean biases are generally not better on the native grid (50 km resolution) than when regridded on ERAI grid (75 km resolution). The correction of topography improves the comparison for all models but ERAI and LMD50 on its native grid.

The standard deviation indicates a large spread around observations for all models (~3.5 $kg.m^{-2}$ to 4.0 $kg.m^{-2}$), except IPSL which is even better than ERA-interim. The nudging towards ERA-interim used in this simulation likely explains this behaviour. This large standard deviation, which is reduced when using the corrected topograhy, does not mean that the model is wrong in a climatological context, but it is the result of internal

variability in absence of nudging, generating deviations from its driving data in the interior of the domain (e.g Kida et al. 1991), partly compensated by potential artificial flows created to achieve physical consistency with the lateral boundary conditions (Becker et al., 2015 ; Omrani et al., 2015). However, in the configuration of Med-CORDEX, the domain is quite small and the jump of resolution between lateral boundary conditions (75 km) and the RCM (50 km) is also weak, so that we could have expected smaller differences between model using nudging and not (Matte et al., 2017).

To the first order, the IWV variability is dominated by the seasonal cycle (shown on Fig.1), which is underestimated by models (not shown), and which explains part of the model standard deviation: indeed, Figure 3 shows the percentage of simulated daily mean IWV values which overestimate GPS values at each station for the ensemble of the five models regridded to the ERAI grid, in winter and summer. Figures 3a and 3b present results without height-correction, while Figs.3c and 3d are done with height-corrected data. In winter, more than 70% of values are overestimated over most stations, with or without height corrections. In summer, this percentage decreases appreciably in most stations and in almost half of them it reaches values below 50% (Fig.3b). The use of height correction homogenizes the results in summer between stations and the very low or very high percentages do not appear anymore when this correction is applied. UCLM50 simulation is the moistest, especially in summer (not shown), which explains its high standard deviation in Table 3.

The IWV variability also comes from the interannual variability. For each month, we computed its anomaly by substracting the average value of the month over all the years. We then computed the correlation between the anomalies of the GPS IWV estimates and those from the models. The minimum and maximum values of these correlations for each model when all stations are considered are indicated in Table 3 (last column). As indicated by these numbers that range between 0.78 and 0.99, the interannual variability is well captured by model, which is not surprising since this variability is mostly driven by large scale advection of air masses (all models use 6-hourly ERA-Interim parameters as lateral boundary conditions). Note that the maximum correlation is very high for all models, while the minimum values are higher for ERAI, LMD50 and IPSL50 than for the three others. This is mostly due to a specific month (not shown): in january 1996, three models have an anomaly very different from the observations and, as a consequence, the interannual correlation for january goes down.

To conclude, the RCM configuration allows a reasonable representation of the large scale advection of air masses by the models, which is an important driver of humidity within the RCM domains (see also Trenberth et al., 2005). This is further improved when the model is nudged towards reanalysis, as in the IPSL model (Tab.3).

Nevertheless RCM errors are significant, with a difficulty to reproduce low values of IWV, generating a mean positive bias for all models. Most of the contribution of humidity to the integrated water vapor comes from the surface, and the interaction between environment and clouds and boundary layer. Small scale processes are thus important to reproduce moisture sources and sinks (precipitation, mesoscale circulations, evaporation and evapotranspiration, clouds and microphysics). To better understand these errors, we assess the link between surface humidity and IWV at different time scales. Figure 4 displays the monthly mean values of IWV versus monthly mean values of 2-m specific humidity (Q2) averaged over all stations where and when both IWV from GPS and Q2 from HadISD are available. Monthly means are computed if at least 60 concommitant (both IWV from GPS and Q2 from surface station) values are available (i.e about 2 values per day out of 4 possible). A total of 3238 months are obtained, spread over 42 different stations. The average number of stations per month is 19 with a maximum of 30 stations. Figure 4 shows that 2-m specific humidity is a very good proxy for IWV at the

monthly scale. All models but IPSL have a similar relationship between the two variables to the observed one (slope of $2.4*10^3$ $kg.m^{-2}$). However, for a given surface humidity, IWV is generally overestimated by models, especially for the driest conditions. The IPSL model presents a different behavior: for low values of surface humidity, IPSL shows the same bias than the other models, while at higher Q2 values, its bias strongly increases, generating a different regression slope than observations. IPSL compensates its underestimation of surface humidity in summer (Bastin et al., 2018) by a steeper slope. This compensation may be the result of the use of nudging or it can be due to a deep boundary layer so that the total humidity contained within the boundary layer is similar to that of other models (not possible to check that). Since the nudging is only used above the boundary layer, and since most of humidity is contained within the boundary layer, there is little reason that the nudging totally explains this compensation. It means that for this model, Q2 is not a good proxy of total column humidity, even at monthly scales. Note also that the spread between models is a bit higher in summer than in winter, most probably because of more active boundary layers and increased entrainment of humidity from the free troposphere at their top. These processes will also affect the link between surface humidity and IWV at scales shorter than a month.

Despite the strong correlation between the annual cycle of Q2 and those of IWV, Q2 is not necessarily a good proxy for IWV at other time scales or to tackle model biases (for instance IPSL bias for surface humidity is strongly negative while it is weak and slightly positive for IWV). While in the wintertime, humidity variability mostly originates from the air mass advection, summertime variability is mainly affected by land-surface interactions and boundary layer processes. Several studies have shown the existence of a large spread in the representation of the surface fluxes and land-atmosphere coupling strength between models over Europe, due to the fact that Europe is a zone of transition between the regime of « energy limited » with low land-atmosphere coupling strength and those of « soil moisture limited » with high land-atmosphere coupling strength. The difficulty to represent soil conditions and surface fluxes is then increased (Cheruy et al., 2015 ; Boe and Terray, 2014 ; Fischer et al., 2007 ; Knist et al., 2017). The interannual correlations between IWV and Q2 summertime anomalies are indicated for each model in Table 5, as averaged values over all GPS stations and minimum/maximum values across all GPS stations (an attempt was done for GPS IWV and HadISD Q2 but there were too many missing values). For most stations, the correlation is higher than 0.5 with a mean value around 0.8 for the 5 models. The standard deviation of the difference between models is around 0.15, which reveals a good agreement between them. Some stations however indicate higher differences, as indicated by the minimum values that strongly differs between models. IPSL50 and LMD50 models present strong correlation at all stations ($r > 0.52$ and 0.43, respectively) while it is very weak at some stations for UCLM50, CNRM50 and ERAI. It can be explained by a stronger availability of surface humidity in summer, and then a weaker sensivity of IWV to the surface moisture availability. For the drier models (e.g IPSL50), a dry interannual anomaly of soil moisture will have an impact on the surface evaporation and then on the IWV anomaly. On the other hand, over areas where the advection of air masses from the sea/ocean is a more important driver, the interannual variability of the large scale dynamics affects more strongly both IWV and surface humidity than surface processes. Depending on the area and on the model, Q2 and IWV thus convey different but complementary information about the model behaviour.

4.2 Scaling of IWV with temperature

A way to check the behavior of models is to consider the relationship between IWV and temperature. At global scale, the scaling of IWV with temperature is expected to follow the Clausius-Clapeyron (C-C) law, at a rate of about 6 %.°C^{-1} to 7 %.°C^{-1}. At regional scale, this relationship deviates from C-C due to the strong influence of dynamics (air masses advection) and moisture availability (e.g Drobinski et al., 2016). Figure 5a illustrates the scaling of IWV with temperature over the SIRTA station, which is representative of most stations over Europe. For the lower temperature, the scaling follows C-C law. Above a critical temperature (T_{break}), IWV stops to increase at this rate. This critical temperature, defined as the temperature when the slope of the relationship deviates from C-C, presents spatial variations that are displayed on Fig.5c for observations. For most stations, the critical value is between 15°C and 18°C. It is higher for stations located around the Mediterranean sea, the Black Sea or at the eastern edge of the domain, in the Dniepr and Volga basins. The IWV value corresponding to this critical temperature (which is close to the maximum IWV value reached at each station) is more variable (Fig.5d), due to a combinated effect of T_{break} value, orography, and latitudinal differences.

Physically, T_{break} corresponds to the value when the relative humidity significantly decreases due to a lack of humidity sources.

To determine the slope after T_{break}, we can approximate the expression of IWV:

$$IWV(T) = \int_{Z=0}^{Z=TOA} \rho_v(T,z)dz = -\frac{1}{g}\int_{Ps}^{P_{TOA}} Q(T,p)dP$$

with Q the specific humidity at altitude z, ρ_v the density of water vapor in kg.m^{-3}, P_s the surface pressure and P_{TOA} the pressure at the Top Of Atmosphere (TOA).

As observed in average (e.g Ruzmaikin et al., 2014), the troposphere can be separated into two layers, one being the boundary layer (BL) and the other one the free troposphere (FT): assuming constant humidity within the two layers, IWV can be expressed as:

$$IWV(T) \approx \frac{1}{g}((P_s\text{-}P_{BL})*Q_{BL}(T_{BL}) + (P_{BL}\text{-}P_{TOA})*Q_{FT}(T_{FT}))$$

P_s is about 1000 hPa, P_{TOA} between 0 and 10 hPa, and P_{BL} around 800 hPa-850 hPa, so that we can write:

$(P_s\text{-}P_{BL}) = \alpha\ (P_{BL}\text{-}P_{TOA}) \cong \alpha\ P_{BL}$ with $\alpha \cong 1/6\text{-}1/4$

For an air mass at temperature T and specific humidity Q, we can write $Q(T) = RH* Q_s(T)$, with subscript s for saturation and RH for relative humidity.

Which gives:

$$IWV(T) \approx \frac{1}{g}(\alpha\ P_{BL}*RH_{BL}*Q_s(T_{BL}) + P_{BL}*RH_{FT}*Q_s(T_{FT})) \qquad (1)$$

With our simplified profile in two layers, we have:

$Q_s(T_{FT}) \approx \alpha Q_s(T_{BL})$ if we consider that T_{BL} is around 280 K-290 K (Q_s then ranges between ~6 g.kg^{-1} and ~12 g.kg^{-1} according to August-Magnus-Roche formula) and T_{FT} around 260 K (for T<260 K, the value of T does not affect a lot the value of Q^s since it has nearly reached its minimum value, being inferior to 2 g.kg^{-1}).

RH_{FT} can approximately be considered as a constant, close to 30% in average over Europe (Ruzmaikin et al., 2014)

Equation (1) can thus be approximated by:

$$IWV(T) \approx \frac{\alpha}{g} P_{BL} \, (RH_{BL} + RH_{FT})^{*} Q^{s}(T_{BL})$$

And finally, by differentiation we obtain equation (2)

$$\frac{1}{IWV}\frac{\partial IWV}{\partial T} \approx \frac{1}{Q^{s}_{BL}}\frac{\partial Q^{s}_{BL}}{\partial T} + \frac{1}{RH_{BL}+RH_{FT}}\frac{\partial RH_{BL}}{\partial T} \qquad (2)$$

The first term of the right hand side (RHS) of Eq. (2) is the variation of water holding capacity of the atmosphere at temperature T and is thus Clausius-Clapeyron rate (~7%.$C^{\circ-1}$ at 25°C).

At temperature lower than T_{break}, RH is nearly constant, so that the second term of RHS of Eq. (2) is close to zero and the slope of IWV as a function of temperature follows C-C. According to Eq. (2), the deviation of the slope from C-C after T_{break} should correspond more or less to the rate of RH decrease, which depends on the considered station. Figure 5b shows RH and IWV as a function of T for SIRTA observations. RH starts to decrease significantly at T~13°C (hereafter called T_{b1}), while IWV curve deflects at T~16°C (hereafter called T_{b2}). Table 6 indicates the values of the left hand side term of Eq(2) and the second term of the right hand side of Eq. (2) before T_{b1} (at 10°C) and after T_{b2} (at 20°C). In spite of the important approximations done to obtain Eq. (2), the RH variations in the boundary layer thus explain to the first order the scaling of IWV with T. The determination of humidity sources that explain the RH variability are thus crucial.

Figure 5a indicates that despite the bias at low temperature, models generally capture the deviation from C-C, and the IWV maximum value that is reached at high temperature, except the UCLM model for which IWV slightly continues to increasing with T. Note that the model slopes at low T are a bit lower than that of the C-C law and that derived from the observations. They are also often slightly different above the critical value, indicating some difficulties to simulate the relative humidity decrease. Figure 6a shows the model ensemble mean of the scaling at the SIRTA station. It confirms that the deviation from C-C exists but the transition is smoother in the models than in the observations which makes the estimate of T_{break} more uncertain at this station. However, some stations show more abrupt transitions (not shown). For the ensemble, due to the smooth transition, T_{break} is thus defined as the temperature value when IWV stops to increase (i.e when it reaches its hiatus value). Figure 6b and c indicates the model ensemble mean value of T_{break} and the corresponding IWV_{break}. The models capture the spatial pattern of T_{break}, especially the higher values close to the Mediterranean, the Black Sea and in the eastern edge of the domain, but tend to overestimate it compared to the observations (Fig5c) but as already said, the uncertainty on the T_{break} estimate is quite high. The maximum value of IWV is also generally well simulated by the models, except over the northern part of the domain where UCLM50 simulation does not always capture this break (as seen on Fig.5a), indicating an overestimation of relative humidity.

The link between the IWV and RH evolutions for the model ensemble is shown on Fig.6a. Once again, the transition from one regime to another is smoother than for observations (Fig.5b), but the decrease of RH starts around the same range of temperature as for the transition of IWV-T relationship. Table 6 confirms that also in models RH variations in the boundary layer explain to a good degree the scaling of IWV with T.

In conclusion to this section models tend to overestimate low values of IWV. Although they generally well capture the IWV scaling with temperature, small scale processes also explain part of standard deviation (not only induced by the deviations from driving data) when considering the differences with GPS IWV data. In the next section, we consider the impact of these humidity biases in models on light precipitation occurrence.

5 Impact of model biases on light precipitation

5.1 Over SIRTA

Figure 7 displays the 50th percentile of precipitation (computed including days without precipitation) as a function of IWV for four different bins of mean tropospheric temperature for both observations and the models (see section 3 for the details of the methodology). The observations are considered for the period 2008-2015 and the different models for the period 2001-2008 (i.e. the same number of years, despite the time shift). For all these datasets, there exists a critical value of IWV, w_c, over which precipitation starts to increase. The values of these critical values are determined from Figure 7 following methodology indicated on section 3.3. The critical value depends on temperature (it increases with temperature) and on the models, as shown on Figure 8a. UCLM50 presents lower critical values of IWV than observations at all temperatures, and CNRM50 also at high temperatures. This means that in these models the precipitation begins in a drier atmosphere than that which begins to produce rain in the observations. On the contrary, for CMCC50, at high temperature, the atmosphere needs to be as wet or wetter than what is observed to trigger precipitation. IPSL50 and LMD50 present similar values than observations for this critical value. CNRM50 and UCLM50 are the two models which present too often light precipitation and not enough days without precipitation (Fig.2).

The second reason of the high frequency of occurrence of very light precipitation in these models is that the probability of exceeding this critical value of IWV is strongly overestimated in these two models in comparison with observations (Fig.8b). Statistically speaking, it means that the models that are too humid have a positive bias in light precipitation. For CMCC, the dipole between winter and summer observed in the estimate of non-precipitating days is observed here also : though the critical value of IWV is correct at low and moderate T, the probability to exceed it is too strong, i.e typically during winter, which explains why it rains too much during winter, while it is not the case during summer, with very low probability to exceed the high critical value of IWV. IPSL model is dry both in terms of humidity and precipitation, and LMD50 follows observations but with a slightly higher probability to exceed the critical value. It is also important to note that the underestimation of non-precipitating days for all models in winter is consistent with the systematic overestimation of low values of IWV (Figs 3, 4, 5).

In addition to period 2001-2008 discussed above, we tested two other 8-year periods (1989-1996 and 1995-2002) and the entire period (1989-2008) to assess the influence of the considered period on the results. Figure S1 and Table 7 show that results are rather robust among models and periods though there exists some uncertainty in both w_c and the probability to exceed it. The maximum relative variability in w_c for one model is around 25%, which is small compared to the maximum difference when considering all values from all models that is 58% for bin 1 (253 K), 73% for bin 2 (257 K), 63% for bin 3 (261 K), and 64% for bin 4 (264 K). The warming of the tropospheric temperature due to climage change is also visible on Fig. S1 and in Table 7 since for all models and

for the three warmer temperature bins, the warmer temperature is obtained during the most recent period (2001-2008 ; i.e period 3). On the contrary, lower temperatures (first bin) tends to decrease, indicating a tendency of the distribution to become wider. We can note that due to the variability of the critical value of IWV and the variability of T inside each bin, the maximum value of w_c inside each bin is not always obtained during the period of maximum temperature. For the first bin, 4 models out of 6 indicate a higher value of w_c during the first period, which is the warmer for this bin. For the other three bins, the maximum temperature is obtained for all six models during the most recent period, i.e period 3 (100% of values) while this period gets only 39% (7 out of 18 cases) of maximum w_c values, and the three other periods represent about 20% each.

5.2 Generalization

To have an idea of the models' behavior over other parts of Europe, several other stations are considered in this section. Their locations are shown on Fig. 1b by black diamonds and details are given in Table 8. Except for Marseille located in the south of France where the COMEPHORE product associated with GPS has been used, the model outputs are not compared with observations for the other stations. Figure 9 displays the occurrence of non-precipitating days for the different models and from COMEPHORE product at the station location when available (over France), w_c values as a function of temperature and the percentage of IWV values that exceed w_c as a function of temperature. The observed frequency of occurrence of non-precitating days is slightly higher in Marseille, located in the dry southern France, than in the northern part of France (SIRTA). Most of the models (except CNRM50) overestimate this north-south gradient. For instance, for UCLM50, the frequency is between 80 and 90% at Marseille, while it is around 15% at SIRTA. Several other characteristics of the models' behavior do not depend much on the station: CNRM50 always simulates less occurrence of non-precipitating days than other models and simulates a rather flat annual cycle ; the annual cycle simulated by CMCC50 is always very intense, with much higher frequency of occurrence of non-precipitating days in summer than in winter, IPSL is always the model with the highest frequency of occurrence of non-precipitating days ; UCLM50 and LMD50 are between CNRM50 and IPSL50, with a tendency to simulate many days with very light precipitation at the northern stations (Fig.2 and Fig.9g, m) but not at the southern ones (Fig.9a, d). In the eastern part of the domain (Fig.9j), the annual cycle of precipitation simulated by models is a bit different than elsewhere with two drier periods in spring and fall.

To relate these caracteristics to temperature and humidity, we reproduced the analysis done at SIRTA. The value of the critical value of IWV over which precipitation starts to increase is generally similar among models, despite increasing dispersion with temperature. This value depends on the stations, indicating the influence of local specificities in the estimation of this relationship. For instance, UCLM50, which is the model with the most important difference between southern and northern stations for the annual cycle of dry days also indicates strong differences in the w_c for a similar temperature, with a critical value around 25 $kg.m^{-2}$ in Marseille for a tropospheric temperature of 260 K (Fig.9b), while it is ~12 $kg.m^{-2}$ in the Netherlands (Fig.9n). The nature of precipitation, more or less convective, likely explains these differences. The probability to exceed the critical value is the most discriminant parameter between models and between seasons. Results are robust and confirm the importance of the relationship between temperature, IWV and light precipitation: for a given bin of temperature if the model is too humid (higher probability to exceed the threshold), it rains too often.

6 Conclusion

This work uses GPS integrated water vapor measurements associated with temperature and precipitation measurements to i) estimate the biases of six regional climate models over Europe in terms of humidity ; ii) understand their origins ; iii) and finally assess the impact of these biases on the occurrence of precipitation.

The first part of the study aimed at evaluating the mean bias and standard deviations of IWV in models compared to GPS measurements at interannual, seasonal, and daily time scales. An interesting result is that all models overestimate the lower values of IWV (nighttime, wintertime) at all stations. The spread among models is increased during summertime. Our analysis suggests that the model physics mostly explain the mean biases, while dynamics affects the variability. The use of nudging towards reanalyses thus improves the representation of the large scale advections of air masses and reduces the standard deviation of differences between GPS retrieved IWV and simulated ones. The land surface/atmosphere interactions are crucial in the estimation of IWV over most part of Europe, especially in summer, and explain part of the mean biases. However, the relationship between IWV and temperature, that deviates from the Clausius-Clapeyron law after a critical value of temperature, is generally well captured by models. This critical temperature presents a spatial variability since it corresponds to the value when relative humidity starts to decrease. It is thus strongly dependent on local processes which drive the local humidity sources (from evaporation and advection). This explains why the maximum values of IWV are not necessarily observed over warmer areas, which often corresponds to dry areas, where soil moisture limited regime is dominant.

The improvement in humidity representation may also help in the representation of precipitation distribution. Indeed, in the second part of this study, it is shown that the biases in IWV and most importantly IWV's distributions as a function of temperature strongly impact the occurrence of light precipitation over France, and most generally over areas where convection is the main process of precipitation triggering. For each range of mean tropospheric temperature, there exists a critical value of IWV over which a pickup in precipitation occurs. This is observed and simulated by models, but the critical values and the probability to exceed them vary between models and observations. Models which present too often light precipitation generally show lower critical values and higher probability to exceed them. Thus, a better knowledge and representation of the triggering thresholds of precipitation and of their variablity should potentially help to improve the representation of the whole precipitation distribution in models. The ensemble of simulations with implicit and explicit convection that will be performed in the framework of the Flagship Pilot Study Convective-Permitting Climate Simulation of CORDEX project will allow us to assess the sensitivity of precipitation triggering and distribution to the model resolution. Issues that will be explored in more details following this work will be the role of humidity in i) the triggering of precipitation in simulations at different resolution; ii) the low precipitation rates (precipitation efficiency); iii) the impact of too easy triggering in the entire precipitation distribution.

Acknowledgements: This work is a contribution to the HyMeX program (HYdrological cycle in The Mediterranean EXperiment) through INSU-MISTRALS support and the MEDCORDEX program (Coordinated Regional climate Downscaling EXperiment—Mediterranean region). This research has received funding from the French National Research Agency (ANR) project REMEMBER (grant ANR-12- SENV-001) and is a

contribution to the VEGA project through LEFE/INSU support, to the EECLAT project through LEFE-INSU and CNES supports and to the GNSS4SWEC COST action ES1206 through EU support. It was supported by the IPSL group for regional climate and environmental studies, with granted access to the HPC resources of GENCI/IDRIS (under allocation i2011010227), by SIRTA Working Group 'Climate studies' and by the national infrastucture ACTRIS-FR, identified on the French road map for Research Infrastructures, published by the Ministry of Research. SIRTA-ReOBS effort also benefited from the support of the L-IPSL funded by ANR under the "Programme d'Investissements d'Avenir (Grant ANR-10-LABX-0018) and by the EUCLIPSE project funded by the European Commission under the Seventh Framework Program (Grant no 244067). To process the data, this study benefited from the IPSL mesocenter ESPRI facility which is supported by CNRS, UPMC, Labex L-IPSL, CNES and Ecole Polytechnique. We would like to acknowledge the SIRTA team for collecting data; Cindy Lebeaupin-Brossier and Marc Stefanon for providing simulation outputs; the CNES (Centre National d'Etudes Spatiales) for partially funded M. Chiriaco research, Emmanuele Lombardi and ENEA for the Med-CORDEX database, Samuel Somot for his role in Med-CORDEX coordination, and Thomas Noel for its help in post-processing data. The work to carry out the simulations of the UCLM model was funded by the Spanish Ministry of Education and Science and the European Regional Development Fund, through grant CGL2007-66440-C04-02.

References

Allan, R.P. and Soden, B.: Atmospheric warming and the amplification of precipitation extremes, Science, 321 (5895), 1481-1484, 2008.

Allen, M.R. and Ingram, W.J.: Constraints on future changes in climate and the hydrological cycle, Nature, 419:224–32, 2002.

Baldauf, M. and Schulz, J.P.: Prognostic precipitation in the Lokal– Modell (LM) of DWD, COSMO Newsletter, 4:177–180, 2004.

Bastin, S., Champollion, C., Bock, O., Drobinski, P. and Masson, F.: Diurnal cycle of water vapor as documented by a dense GPS network in a coastal area during ESCOMPTE-IOP2, J Appl Meteorol Clim, 46, 167-182, 2007.

Bastin, S., Champollion, C., Bock, O., Drobinski, P. and Masson F.: On the use of GPS tomography to investigate the water vapor variability during a Mistral/sea-breeze event in southeastern France, Geophys Res Lett, 32, No. 5, L05808, d01:10.1029/2004GL021907, 2005.

Bastin, S., Chiriaco, M. and Drobinski, P.: Control of radiation and evaporation on temperature variability in a WRF regional climate simulation: comparison with colocated long-term ground based observations near Paris, Clim Dynam, 51: 985-1003. doi:10.1007/s00382-016-2974-1. insu-01282710, 2018.

Becker, N., Ulbrich, U. and Klein, R.: Systematic large-scale secondary circulations in a regional climate model, Geophys.Res.Lett., 42(10), 4142–4149, doi:10.1002/2015GL063955, 2015.

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: doi: 10.1029/92JD01517, issn: 0148-0227, 1992.

Bock O, Keil C, Richard E, Flamant C, Bouin MN. Validation of precipitable water from ECMWF model analyses with GPS and radiosonde data during the MAP SOP. Q. J. R. Meteorol. Soc. 131:3013–3036, 2005.

Bock, O., et al.: West African Monsoon observed with ground-based GPS receivers during African Monsoon Multidisciplinary Analysis (AMMA), J Geophys Res, 113, D21105, doi:10.1029/2008JD010327, 2008.

Bock, O., Bosser, P., Bourcy, T., David, L., Goutail, F., Hoareau, C., Keckhut, P., Legain, D., Pazmino, A., Pelon, J., Pipis, K., Poujol, G., Sarkissian, A., Thom, C., Tournois, G., and Tzanos, D.: Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP, Atmos. Meas. Tech., 6, 2777-2802, https://doi.org/10.5194/amt-6-2777-2013, 2013.

Boé, J. and Terray, L.: Land–sea contrast, soil-atmosphere and cloud-temperature interactions: interplays and roles in future summer European climate change, Clim Dynam, 42:683–699. doi:10.1007/s00382-013-1868-8, 2014.

Bougeault, P.: A simple parameterization of the largescale effects of cumulus convection, Mon Weather Rev, 113:2108–2121, 1985.

Champollion, C., Flamant, C., Bock, O., Masson F., Turner D. and Weckwerth T.: Mesoscale GPS tomography applied to the 12 June 2002 convective initiation event of IHOP_2002, Q J Roy Meteor Soc, 135, 645 – 662, 10.1002/qj.386, 2009.

Chen, C.T. and Knutson, T.: On the verification and comparison of extreme rainfall indices from climate models, J Climate, 21:1605–1621, 2008.

Cheruy, F., Dufresne, J.L., Hourdin F. and Ducharne, A.: Role of clouds and land-atmosphere coupling in midlatitude continental summer warm biases and climate change amplification in CMIP5 simulations, Geophys Res Lett, 41:6493–6500, 2015.

Chiriaco, M., Bastin, S., Yiou P., Haeffelin M., Dupont, J.-C. and Stéfanon M.: European heatwave in July 2006: observations and modeling showing how local processes amplify conducive large-scale conditions, Geophys Res Lett, 41, 5644–5652, doi:10.1002/2014GL060205, 2014.

Chiriaco, M., Dupont, J.-C., Bastin, S., Badosa, J., Lopez, J., Haeffelin, M., Chepfer, H. and Guzman, R.: ReOBS, a new approach to synthetize long-term multi-variable dataset and application to the SIRTA supersite, Earth Syst Sci Data, 10, 919-940, doi:10.5194/essd-10-919-2018, 2018.

Colin, J., Déqué, M., Radu, R. and Somot, S.: Sensitivity study of heavy precipitation in limited area model climate simulations: influence of the size of the domain and the use of the spectral nudging technique, Tellus, 62A:591–604, 2010.

Cuxart, J., Bougeault, P. and Redelsperger, J.L.: A turbulence scheme allowing for mesoscale and large-eddy simulations, Q J Roy Meteor Soc, 126:1–30, 2000.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Koehler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q J Roy Meteor Soc, 137(656), 553–597, doi:10.1002/qj.828, 2011.

Domínguez, M., Gaertner, M.A., De Rosnay, P. and Losada, T.: A regional climate model simulation over West Africa: parameterization tests and analysis of land-surface fields. Clim Dynam, 35:249–265, 2010.

Domínguez, M., Romera, R., Sánchez, E. et al.: Present-climate precipitation and temperature extremes over Spain from a set of high resolution RCMs. Climate Res, 58:149–164, 2013.

Doms, G., Förstner, J., Heise, E., Herzog, H.-J., Raschendorfer, M., Schrodin, R., Reinhardt, T. and Vogel G. : A description of the nonhydrostatic regional model LM. Part II: physical parameterization. http://www.cosmomodel.org/content/model/documentation/core/cosmoPhysParamtr, 2007.

Drobinski, P., et al.: HyMeX, a 10-year multidisciplinary program on the Mediterranean water cycle, B Am Meteorol Soc, 95:1063–1082, 2014.

Drobinski, P., Alonso, B., Bastin, S., Da Silva, N. and Muller, C.: Scaling of precipitation extremes with temperature in the French Mediterranean region: what explains the hook shape? J Geophys Res-Atmos, 121 (7), pp.3100-3119. <10.1002/2015JD023497> - insu-01298539, 2016.

Drobinski, P., Da Silva, N., Panthou, G., Bastin, S., Muller, C., Borga, M., Conte, D., Fosser, G., Güttler, I., Kotroni, V., Li, L., Morin, E., Onol, B., Quintana-Segui, P. and Romera, R.: Temperature-precipitation extremes

relationship in the Mediterranean: past climate assessment and projection in anthropogenic scenarios, Clim Dynam, 51:1237-1257, <10.1007/s00382-016-3083-x> - insu-01294073, 2018.

Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. J Atmos Sci, 46:3077–3107, 1989.

Dunn, R. J. H., et al.: HadISD: A Quality Controlled global synoptic report database for selected variables at long-term stations from 1973-2011, Clim Past, 2012.

Emanuel, K.A.: A cumulus representation based on the episodic mixing model: the importance of mixing and microphysics in predicting humidity, AMS Meteorol Monogr, 24:185–192, 1993.

Fischer, E.M., Seneviratne, S.I., Vidale, P.L., Lüthi, D. and Schär C.: Soil moisture–atmosphere interactions during the 2003 European summer heat wave. J Climate, 20, pp. 5081-5099, 2007.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M.: Evaluation of climate models. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Doschung, J., Nauels, A., Xia, Y., Bex, V., and P.M. Midgley, Eds. Cambridge University Press, 741-882, doi:10.1017/CBO9781107415324.020, 2013.

Fouquart, Y. and Bonnel, B. : Computations of solar heating of the Earth's atmosphere: a new parametrization. Contrib Atmos Phys, 53:35–62, 1980.

Fumière, Q., Déqué, M., Nuissier, O., Somot, S., Alias, A., Seity, Y., Laurantin, O. : Extreme rainfall in Mediterranean France during autumn: added value of the AROME convection permitting regional climate model. Sub. to Clim Dynam

Haeffelin, M., et al.: SIRTA, a ground-based atmospheric observatory for cloud and aerosol research. Ann Geophys, 23:253–275, 2005.

Hagemann, S., Bengtsson, L., and Gendt, G.: On the determination of atmospheric water vapor from GPS measurements, J Geophys Res, 108, *4678, doi:*10.1029/2002JD003235, D21, 2003.

Held, I. M., and Soden, B. J.: Robust Responses of the Hydrological Cycle to Global Warming, J Climate, 19:5686–5699, doi: http://dx.doi.org/10.1175/JCLI3990.1, 2006.

Holloway, C. E., and Neelin, J. D.: Moisture vertical structure, column water vapor, and tropical deep convection, J Atmos Sci, 66, 1665–1683, 2009.

Hong, S.Y., Dudhia, J., and Chen, S.H. : A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, Mon Weather Rev, 132:103–120, 2004.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.L., Fairhead, L., Filiberti, M.A., Friedlingstein and P., Grandpeix, J.Y.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, Clim Dynam, 27:787–813, 2006.

Kain, J.S. : The Kain–Fritsch convective parameterization: an update, J Appl Meteorol, 43:170–181, 2004.

Kharin, V.V., Zwiers, F.W., Zhang, X. and Hegerl, G.C. : Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, J Climate, 20:1419–1444, 2007.

Kida, H., Koide, T., Sasaki, H. and Chiba, M.: A New Approach for Coupling a Limited Area Model to a Gcm for Regional Climate Simulations, J. Meteorol. Soc. Jpn., 69(6), 723–728, 1991.

Knist, S., et al.: Land-atmosphere coupling in EURO-CORDEX evaluation experiments, J Geophys Res-Atmos, 122 ,79–103, doi:10.1002/2006JD025476, 2017.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, C. : A dynamic global vegetation model for studies of the coupled atmosphere biosphere system. Glob Biogeochem Cycles 19:GB1015. doi:10 .1029/2003GB002199k, 2005.

Laurantin, O., Tabary, P., Dupuy, P., L'Henaff, G., Merlier, C., Soubeyroux, J.M.: A 10-year (1997–2006) reanalysis of quantitative precipitation estimation over France. ERAD 2012: the 7th European conference on radar meteorology and hydrology.

Li, Z.X.: Ensemble atmospheric GCM simulation of climate interannual variability from 1979 to 1994. J Climate, 12:986–1001, 1999.

Louis, J.-F. : A parametric model of vertical eddy fluxes in the atmosphere, Bound-Lay Meteorol, 17:187–202, 1979.

Matte , D., Laprise, R., Thériault, R. M., Lucas-Picher, P. : Spatial spin-up of fine scales in a regional climate model simulation driven by low□resolution boundary conditions, Clim Dynam, 49:563-574. DOI 10.1007/s00382-016-3358-2

Mlawer, E.J., Taubnam, S.J., Brown, P.D., Iacono, M.J., Clough, S.A.: A validated correlated k-model for the longwave, J Geophys Res, 102:16663–16682, 1997.

Morcrette, J.J., Barker, H., Cole, J., Iacono, M. and Pincus, R.: Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. Mon Weather Rev, 136:4773–4798, 2008

Morcrette, J.J.: Impact of changes to the radiation transfer parameterizations plus cloud optical properties in the ECMWF model, Mon Weather Rev, 118:847–873, 1990.

Morcrette, J.J., Smith, L. and Fouquart, Y.: Pressure and temperature dependence of the absorption in longwave radiation parametrizations. Contrib Atmos Phys 59:455–469, 1986.

Muller, C.: Impact of convective organization on the response of tropical precipitation extremes to warming, J Climate, 26:5028–5043, 2013.

Neelin, J. D., Peters, O., and Hales, K.: The Transition to Strong Convection. J Atmos Sci, 66(8), 2367–2384. https://doi.org/10.1175/2009JAS2962.1, 2009.

Ning, T., Wang, J., Elgered, G., Dick, G., Wickert, J., Bradke, M., Sommer, M., Querel, R., and Smale, D.: The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations, Atmos. Meas. Tech., 9, 79-92, https://doi.org/10.5194/amt-9-79-2016, 2016.

Noh, Y., Cheon, W.G., Hong, S.Y. and Raasch, S.: Improvement of the k-profile model for the planetary boundary layer based on large eddy simulation data, Bound-Lay Meteorol, 107:401–427, 2003.

Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterization scheme. Glob Planet Change, 13:145–159, 1996.

Noilhan, J. and Planton, S.: A simple parameterization of land surface processes for meteorological models. Mon Weather Rev, 117:536–549, 1989.

O'Gorman, P.A. and Schneider, E.: Scaling of precipitation extremes over a wide range of climates simulated with an idealised GCM, J Climate, 22:5676–5685, 2009.

Omrani, H., Drobinski, P. and Dubos, T.: Using nudging to improve global-regional dynamic consistency in limited-area climate modeling: what should we nudge?, Clim Dynam, 44:1627–1644, 2015.

Omrani, H., Drobinski, P. and Dubos, T.: Optimal nudging strategies in regional climate modelling: investigation in a Big-Brother Experiment over the European and Mediterranean regions, Clim Dynam, 41:2451–2470, 2013.

Pall, P., Allen, M.R. and Stone, D.A.: Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming, Clim Dynam, 28:351–363, 2007.

Panthou, G., Vrac, M., Drobinski, P., Bastin, S. and Li, L.: Impact of model resolution and Mediterranean sea coupling on hydrometeorological extremes in RCMs in the frame of HyMeX and MED-CORDEX, Clim Dynam, <10.1007/s00382-016-3374-2>. identifiant HAL : <insu-01378375>, 2016.

Parracho, A. C., Bock, O., and Bastin, S.: Global IWV trends and variability in atmospheric reanalyses and GPS observations, Atmos Chem Phys, 18, 16213-16237, https://doi.org/10.5194/acp-2018-137, 2018.

Randall, D.A., et al: Climate models and their evaluation. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquiz, M., Averyt, K.B., Tignor, M., Miller, H.L. (eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom / New York, NY, USA, 2007.

Ricard, J.L. and Royer, J.F. : A statistical cloud scheme for use in an AGCM, Ann Geophys, 11:1095–1115, 1993.

Ritter, B. and Geleyn, J.-F. : A comprehensive radiation scheme of numerical weather prediction with potential application to climate simulations, Mon Weather Rev, 120:303–325, 1992.

Rockel, B., Will, A. and Hense A. (eds): The regional climate model COSMO-CLM (CCLM), Meteorol Z 17:347–348, 2008.

Ruti, P., et al.: Med-CORDEX initiative for Mediterranean climate studies. Bul Amer Met Soc, 97 (7), pp.1187-1208. <10.1175/BAMS-D-14-00176.1> - insu-01241426, 2016.

Ruzmaikin, A., Aumann, H. H., Manning, E. M.: Relative Humidity in the troposphere with AIRS. J Atmos Sci, 71, 2516-2533, 2014.

Salameh, T., Drobinski, P. and Dubos, T.: The effect of indiscriminate nudging time on the large and small scales in regional climate modelling: application to the Mediterranean Basin, Q J Roy Meteorol Soc, 136:170–182, 2010.

Sahany, S., Neelin, J.D., Hales, K. and Neale, R.B.: Temperature–moisture dependence of the deep convective transition as a constraint on entrainment in climate models. J Atmos Sci, 69, 1340-1358, https://doi.org/10.1175/JAS-D-11-0164.1, 2012.

Singleton, A. and Toumi, R.: Super-Clausius–Clapeyron scalnig of rainfall in a model squall line, Q J Roy Meteorol Soc, 139:334–339, 2013.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.Y., Wang, W. and Powers, J.G.: A description of the advanced research WRF version 3. Technical Report, NCAR, 2008.

Smirnova, T.G., Brown, J.M. and Benjamin, S.G.: Performance of different soil model configurations in simulating ground surface temperature and surface fluxes, Mon Weather Rev, 125:1870–1884, 1997.

Smith, R.N.B. : A scheme for predicting layer clouds and their water content in a general circulation model. Q J Roy Meteorol Soc, 116:435–460, 1990.

Sugiyama, M., Shiogama, H. and Emori, S.: Precipitation extreme changes exceeding moisture content increases in MIROC and IPCC climate models. Proc Natl Acad Sci 107:571–575, 2010.

Sun, Y., Solomon, S., Dai, A. and Portmann R.W.: How often does it rain? J Climate, 19, 916-934, 2006.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon Weather Rev, 117:1779–1799, 1989.

Trenberth, K.E, Dai, A., Rasmussen, R.M. and Parsons D.B.: The changing character of precipitation. B Am Meteorol Soc, 84:1205–121, 2003.

Trenberth, K. E.: Changes in precipitation with climate change. Climate Res, 47, 123–138, 2011.

Trenberth, K. E., Fasullo, J., and Smith L.: Trends and variability in column-integrated atmospheric water vapor, Clim Dynam, 24, 741-758, 2005.

Zhang, L., Wu, L., and Gan, B.: Modes and Mechanisms of Global Water Vapor Variability over the Twentieth Century. J Climate, 26(15), 5578-5593, 2013.

Wang, J., Zhang, L., Dai, A., Van Hove, T., and Van Baelen, J.: A near-global 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements, J Geophys Res, 112, D11107, DOI 10.1029/2006JD007529, 2007.

Wang, J. and Zhang, L.: Climate applications of a global, 2-hourly atmospheric precipitable water dataset derived from IGS tropospheric products. J Geodesy, 83:209–217 DOI 10.1007/s00190-008-0238-5, 2009.

Table 1 : List of models used in the study

Name in this paper	Model/resolution	Number of levels	Radiation scheme	Convective scheme	Microphysics scheme	Land surface scheme	PBL scheme	Land use
CMCC50	CCLM4-8-19/MED44	45	Ritter and Geleyn (1992)	Tiedtke (1989)	Doms et al., (2007) and Baldauf and Schulz (2004)	Doms et al. (2007)	Louis (1979)	GLC2000
CNRM50	ALADIN52/MED44	31	Morcrette (1990)	Bougeault (1985)	Ricard and Royer (1993) and Smith (1990)	Noilhan and Planton (1989) and Noilhan and Mahfouf (1996)	Ricard and Royer (1993)	GLC2000
IPSL50	WRF311/MED44	28	Mlawer et al. (1997) and Dudhia (1989)	Kain (2004)	Hong et al. (2004)	Smirnova et al., (1997)	Noh et al. (2003)	USGS
LMD50	LMDZ4NEMO8/MED44	19	Fouquart and Bonnel (1980) Morcrette et al. (1986)	Emmanuel (1993)	Li (1999) and Hourdin et al. (2006)	Krinner et al. (2005)	Louis (1979)	None
UCLM50	PROMES/MED44	37	Mlawer et al. (1997), Morcrette et al. (2008)	Kain (2004)	Hong et al. (2004)	Krinner et al. (2005)	Cuxart et al. (2000)	IGBP

Table 2 : values of linear regression slopes obtained for each model/each month when plotting IWVmod-IWVobs as a function of difference of altitudes between model and GPS at each station. The first figure is for native grid of the model and the second figure is for model regridded to ERAI grid. It is expressed in 10^{-3} kg.m^{-3}

	CMCC50	*CNRM50*	*IPSL50*	*LMD50*	*UCLM50*
January	*-4.1/-5.5*	*-5.8/-5.6*	*-4.8/-5.9*	*-4.5/-5.1*	*-4.7/-5.3*
February	*-4.1/-5.2*	*-5.3/-5.3*	*-4.7/-5.5*	*-4.5/-5.1*	*-4.7/-5.2*
March	*-4.6/-5.0*	*-5.4/-5.2*	*-5.1/-5.3*	*-4.9/-4.8*	*-4.9/-5.0*
April	*-5.7/-5.0*	*-5.6/-5.0*	*-5.4/-5.4*	*-5.8/-4.8*	*-5.8/-4.7*
May	*-7.0/-6.2*	*-7.0/-6.2*	*-6.6/-6.6*	*-7.0/-6.2*	*-6.8/-5.8*
June	*-8.3/-5.9*	*-7.6/-6.6*	*-8.2/-6.8*	*-7.5/-6.1*	*-7.9/-5.6*
July	*-9.8/-6.8*	*-8.9/-7.7*	*-8.5/-7.3*	*-8.2/-6.7*	*-8.6/-6.3*
August	*-11.2/-8.2*	*-8.2/-8.5*	*-9.0/-8.0*	*-8.6/-7.5*	*-8.6/-6.7*
September	*-8.4/-7.4*	*-7.6/-7.9*	*-8.1/-7.8*	*-7.2/-6.9*	*-8.1/-6.9*
october	*-8.0/-7.4*	*-7.7/-7.6*	*-7.9/-7.8*	*-7.2/-7.0*	*-7.7/-7.0*
November	*-5.4/-5.8*	*-6.2/-6.1*	*-5.8/-6.3*	*-5.7/-5.7*	*-6.0/-5.9*
December	*-4.4/-5.6*	*-5.8/-5.8*	*-5.0/-5.9*	*-4.9/-5.3*	*-5.2/-5.4*

Table 3: Mean bias and standard deviation (SD) in kg.m^{-2} of the differences of IWV between models and GPS observations using daily time resolution, and using the data obtained with the correction due to altitude difference or not. Computation has been done using models on their native grid, or bilinearly remapped on ERA-Interim grid. The rightmost column indicates the minimum and maximum values of the correlation of the interannual variability of monthly GPS anomalies and model anomalies (one correlation computed by month).

	Bias (daily)		SD of difference (daily)		Interannual correlation (min/max)
	uncorrected	corrected	uncorrected	corrected	
ERAI	0.57	0.61	1.87	1.30	0.92/0.99
LMD50	0.71	0.77	3.28	3.09	0.91/0.99
LMD50 grid ERAI	0.77	0.56	3.25	2.96	0.89/0.99
IPSL50	0.95	0.80	1.80	1.38	0.91/0.99
IPSL50 grid ERAI	0.94	0.76	1.92	1.44	0.97/0.99
CNRM50	1.35	1.11	3.35	3.10	0.77/0.98
CNRM50 grid ERAI	1.39	1.14	3.34	3.07	0.79/0.98
CMCC50	0.60	0.55	3.38	3.16	0.78/0.96
CMCC50 grid ERAI	0.59	0.52	3.30	3.05	0.79/0.98
UCLM50	1.81	1.56	3.62	3.47	0.80/0.98
UCLM50 grid ERAI	1.71	1.52	4.30	4.16	0.84/0.97

Table 4 : Occurrence (%) of non –precipitating days (first number), very light precipitation (second number) and light precipitation (third number) for different datasets (columns) computed over different years (rows) for two different periods of the year: W is for winter (Julian day from 1 to 100) and S is for summer (julian day from 151 to 251). Native grids are used.

		REOBS	COM01	COM50	LMD50	UCLM50	CMCC50	CNRM50	IPSL50
2004-2007	W	45/24/8	56/7/8	34/24/9	20/43/11	13/41/18	5/55/13	2/37/18	33/39/8
	S	58/16/4	67/10/5	47/27/6	29/41/9	18/38/15	41/40/5	2/43/11	59/21/4
1989-1996	W	-	-	-	24/40/11	12/41/17	7/56/10	1/42/14	35/38/7
	S	-	-	-	36/42/6	22/38/16	49/35/4	0/48/8	65/20/4
1995-2002	W	-	-	-	23/38/10	23/34/14	7/53/11	1/42/14	36/35/8
	S	-	-	-	33/41/7	24/36/15	45/37/4	1/45/9	62/20/4
2001-2008	W	-	-	-	23/28/11	14/37/17	6/54/13	2/37/16	35/37/8
	S	-	-	-	31/40/7	16/41/15	42/38/5	1/44/10	62/19/4
1989-2008	W	-	-	-	24/38/11	17/38/16	7/55/11	1/41/15	36/37/8
	S	-	-	-	33/42/6	22/38/15	45/37/4	0/46/9	62/21/4
1997-2007	W	-	62/6/7	44/22/8	24/38/11	22/36/16	7/54/12	1/40/15	36/37/8
	S	-	69/8/5	53/21/6	31/40/7	23/37/14	44/37/4	1/45/9	61/21/4
2008-2015	W	53/17/6	-	-	-	-	-	-	-
	S	63/13/5	-	-	-	-	-	-	-
2003-2015	W	47/18/6							
	S	57/13/4							

Table 5 : Summertime interannual correlation between IWV and Q2 at GPS stations. In bold, the mean value for each model, followed by min and max values. The other values in the table corresponds to the standard deviation of the difference of correlation between two models.

	CMCC50	CNRM50	IPSL50	LMD50	UCLM50	ERAI
CMCC50	**0.81/0.30/0.96**	0.18	0.14	0.13	0.18	0.16
CNRM50	0.18	**0.76/0.05/0.97**	0.16	0.15	0.19	0.20
IPSL50	0.14	0.16	**0.81/0.52/0.98**	0.11	0.15	0.20
LMD50	0.13	0.15	0.11	**0.83/0.43/0.96**	0.13	0.16
UCLM50	0.18	0.19	0.15	0.13	**0.76/0.24/0.96**	0.21
ERAI	0.16	0.20	0.20	0.16	0.21	**0.76/0.13/0.97**

Table 6 : Values of the left hand side (LHS), 2nd term of the right hand side (RHS) and total RHS of Eq. (2) computed for the SIRTA site for two different temperatures (10°C and 20°C), according to Fig.5d for observations and Fig.6a for models.

	10°C			20°C		
	LHS	RHS, term 2	RHS total	LHS	RHS, term2	RHS total
OBS	*IWV~15 kg.m^{-2} Slope = 1 kg.m-2. C$^{°-1}$* ***Value = 6.6%. C$^{°-1}$***	*RH~85%+30% Slope = -0.8%.C$^{°-1}$ Value = -0.7%. C$^{°-1}$*	6.6-0.7=**5.9 %. C$^{°-1}$**	*IWV~23 kg.m^{-2} Slope = 0.5 kg.m-2. C$^{°-1}$* ***Value = 2.2%. C$^{°-1}$***	*RH~60%+30% Slope = -3%.C$^{°-1}$ Value = -3.3%. C$^{°-1}$*	*6.2-3.3 =* ***2.9 %. C$^{°-1}$***
IPSL50	*IWV~15 kg.m^{-2} Slope = 0.8 kg.m-2. C$^{°-1}$* ***Value = 5.2 %. C$^{°-1}$***	*RH~80%+30% Slope = -0.8%.C$^{°-1}$ Value = -0.7%. C$^{°-1}$*	6.6-0.7=**5.9 %. C$^{°-1}$**	*IWV~22 kg.m^{-2} Slope = 0.2 kg.m-2. C$^{°-1}$* ***Value = 1.1%. C$^{°-1}$***	*RH~47%+30% Slope = -3%.C$^{°-1}$ Value = -3.9%. C$^{°-1}$*	*6.2-3.9 =* ***2.3 %. C$^{°-1}$***
Model ensemble	*IWV~16.5 kg.m^{-2} Slope = 1 kg.m-2. C$^{°-1}$* ***Value = 6.1 %. C$^{°-1}$***	*RH~90%+30% Slope = -0.6%.C$^{°-1}$ Value = -0.5%. C$^{°-1}$*	6.6-0.5=**6.1 %. C$^{°-1}$**	*IWV~25 kg.m^{-2} Slope = 0.4 kg.m-2. C$^{°-1}$* ***Value = 1.6%. C$^{°-1}$***	*RH~70%+30% Slope = -4%.C$^{°-1}$ Value = -4%. C$^{°-1}$*	*6.2-4 =* ***2.2 %. C$^{°-1}$***

Table 7 : Period for which the maximum value of the temperature bin (T), the critical value of IWV (w_c) is reached. Period 1 is for 1989-1996 ; P2 for 1995-2002 ; P3 for 2001-2008. The second number in the column of Wc corresponds to the maximum relative variability of IWV computed as ((max(Wc)-min(Wc))/mean(Wc)) for each temperature bin and each model

	Bin1-253 K		Bin2-257K		Bin3-261K		Bin4-264K	
	T	Wc	T	Wc	T	Wc	T	Wc
IPSL50	P1	P1/7%	P3	P3/10%	P3	P2/11%	P3	P1/9%
CMCC50	P1	P3/15%	P3	P2/7%	P3	P1/15%	P3	P1/6%
LMD50	P1	P1/22%	P3	P3/5%	P3	P3/17%	P3	P3/19%
CNRM50	P1	P1/20%	P3	P3/7%	P3	P3/8%	P3	P2/16%
UCLM50	P1	P1/4%	P3	P2/11%	P3	P3/6%	P3	P3/24%

Table 8 : Latitude/longitude/altitude of the closest grid point of each model for the GPS stations used.

	LMD50	CNRM50	UCLM50	CMCC50	IPSL50
SIRTA 48.7/2.2/156	48.7/2.3/99	48.6/2.1/125	48.6/2.5/84	48.5/2.1/123	48.7/2.6/103
MARSEILLE 43.3/5.4/12	43.3/5.4/132	43.5/5.1/35	43.1/5.6/0	43.2/5.5/112	43.1/5.6/61
MADRID 40.4/-4.2/777	40.3/-4.4/731	40.5/-4.2/1130	40.4/-4.2/759	40.5/-4.5/922	40.4/-4.2/750
DRESDE 51.0/13.0/160	51.1/13.7/228	51.1/13.8/177	51.1/13.9/256	50.9/14.0/277	51.1/14.0/273
KOOTWIJK 52.2/5.8/53	52.0/5.8/17	52.2/5.8/32	52.2/5.5/15	52.1/5.9/22	52.1/6.1/13
POLTAVA 49.6/34.5/160	49.6/34.6/125	49.7/34.5/145	49.8/34.6/132	49.8/34.6/132	49.4/34.8/118

Figure captions :

Figure 1: Mean values of IWV in winter (a) and summer (b) retrieved from GPS network. Note that color scales are different between winter and summer. Circles indicate stations with more than 10 years of observations between 1995 and 2008 and squares indicate stations with 5 to 10 years of observations. SIRTA observatory is shown by the black triangle. Black diamonds indicate the location of stations considered in section 5.2. Topography higher than 500 m above sea level is shaded in grey.

Figure 2 : Occurrence for 30-day periods from January to December over the years 2004-2007 of a. Non precipitating days ; b. precipitation rates between 0 and 1 mm/day ; c. precipitation rates between 1 and 2 mm/day. d. Value of the 50th quantile of precipitation (excluding non-precipitating days). Each color corresponds to a different model at 50 km resolution regridded to the LMD grid (see legend for details). Black line is for observations at SIRTA supersite (ReOBS). Dashed black line is for COMEPHORE product at 1*1 km2 resolution at the closest grid point of SIRTA (COM01) and black line with squares is for COMEPHORE averaged over a square of 50*50 km2 around SIRTA (COM50).

Figure 3: Percentage of simulated daily mean IWV values which overestimate the GPS ones a) in winter and b) in summer. Simulated values are taken from 5 models at 50 km resolution (LMD50, IPSL50, CNRM50, CMCC50, UCLM50) regridded to the ERAI grid. c) and d) same as a) and b) but for height-corrected values.

Figure 4: Monthly values of IWV as a function of monthly values of 2-m specific humidity (Q2) averaged over all stations where and when both IWV from GPS and Q2 from HadISD are available (Monthly mean are computed if at least 60 co-existing values exist (i.e about 2 values per day over 4 possible). A total of 3238 months are obtained, spread over 42 different stations. The average number of stations per month is 19 with a maximum of 30 stations). The color of circles corresponding to each model is indicated on the legend. All model outputs have been regridded to the ERAI grid. The first number indicates the slope of the regression obtained when considering the same months than observations at each station, while the second number is the slope of the regression when considering all months at all grid points (only models).

Figure 5: **(a)** IWV – T(2m) relationship at SIRTA station from observations and models. Median values are plotted for observations and models. The grey band represents the interval between the 20th and 80th quantiles for observations. The black vertical dashed line shows T_{break}, the black horizontal dashed line shows IWV_{break}, and the red slant dashed line shows C-C scaling **(b)** Median values of IWV and RH(2m) as a function of T(2m) at SIRTA from observations only. The vertical grey dashed lines indicate the values of Tb1 and Tb2 (see text section 4). **(c)** Map of T_{break} at all the GPS stations. **(d)** Map of IWV_{break} at all the GPS stations.

Figure 6 : a) Scaling of IWV (black) and RH (blue) with temperature for the model ensemble at SIRTA station. Solid lines are for quantile 50 and dashed lines are for quantiles 20 and 80 of the distributions. The red slant dashed line shows C-C scaling. b) and c) As Fig5c,d but for the model ensemble.

Figure 7 : 50th quantile of precipitation as a function of IWV for 4 different bins of tropospheric temperature at latitude 48.7°N and longitude 2.2°N (SIRTA). **(a)** SIRTA-ReOBS observations. **(b)** CMCC50. **(c)** IPSL50. **(d)** LMD50. **(e)** CNRM50. **(f)** UCLM50. Observations are analysed for the period 2008-2015, and models for the period 2001-2008.

The values of T indicated in the legends correspond to the mean values in the bins that have been choosen as indicated in the text.

Figure 8 : **(a)** Critical value of IWV at latitude 48.7°N and longitude 2.8°E over which 50th quantile precipitation significantly increases as a function of the mean tropospheric temperature. Each color corresponds to a different model (see legend for details). The period considered is 2001-2008. Solid black line is for observations from the SIRTA-ReOBS dataset between 2008 and 2015 and dashed black line is for SIRTA-ReOBS with ERA-Interim tropospheric temperature instead of tropospheric temperature from radiosoundings. **(b)** Probability for IWV to exceed the critical value for each dataset.

Figure 9 : first column : Annual cycle of occurrence of non-precipitating days for the different models and COMEPHORE in black in panel a (same legend than Fig.8 for models). Second column : Value of w_c (maximum value) as a function of tropospheric temperature. Third column : Probability to exceed w_c value. First row is for the station located in southern France (Marseille), second row for the one located in Central Spain (Madrid), third row for station in eastern Germany (Dresde), fourth row in Ukraine, and fifth row in Netherlands (see table 8 and Fig.1 for details on the locations of the stations).

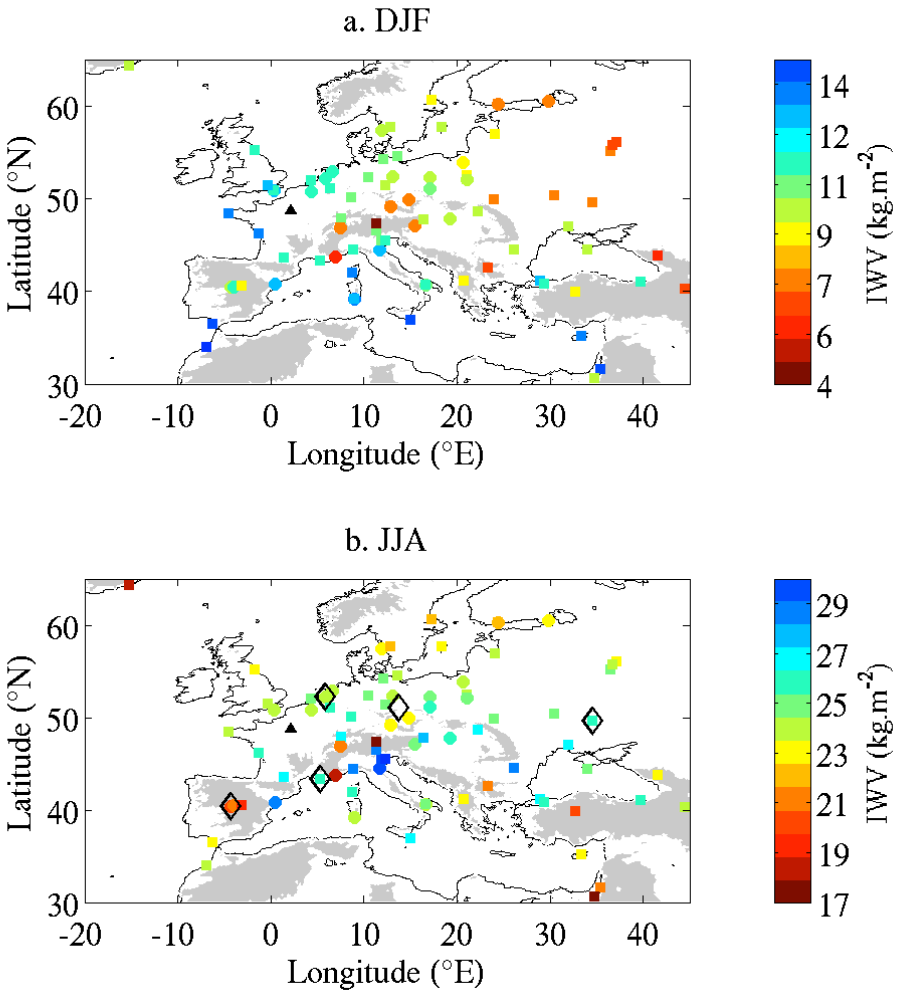

Figure 1: Mean values of IWV in winter (a) and summer (b) retrieved from GPS network. Note that color scales are different between winter and summer. Circles indicate stations with more than 10 years of observations between 1995 and 2008 and squares indicate stations with 5 to 10 years of observations. SIRTA observatory is shown by the black triangle. Black diamonds indicate the location of stations considered in section 5.2. Topography higher than 500 m above sea level is shaded in grey.

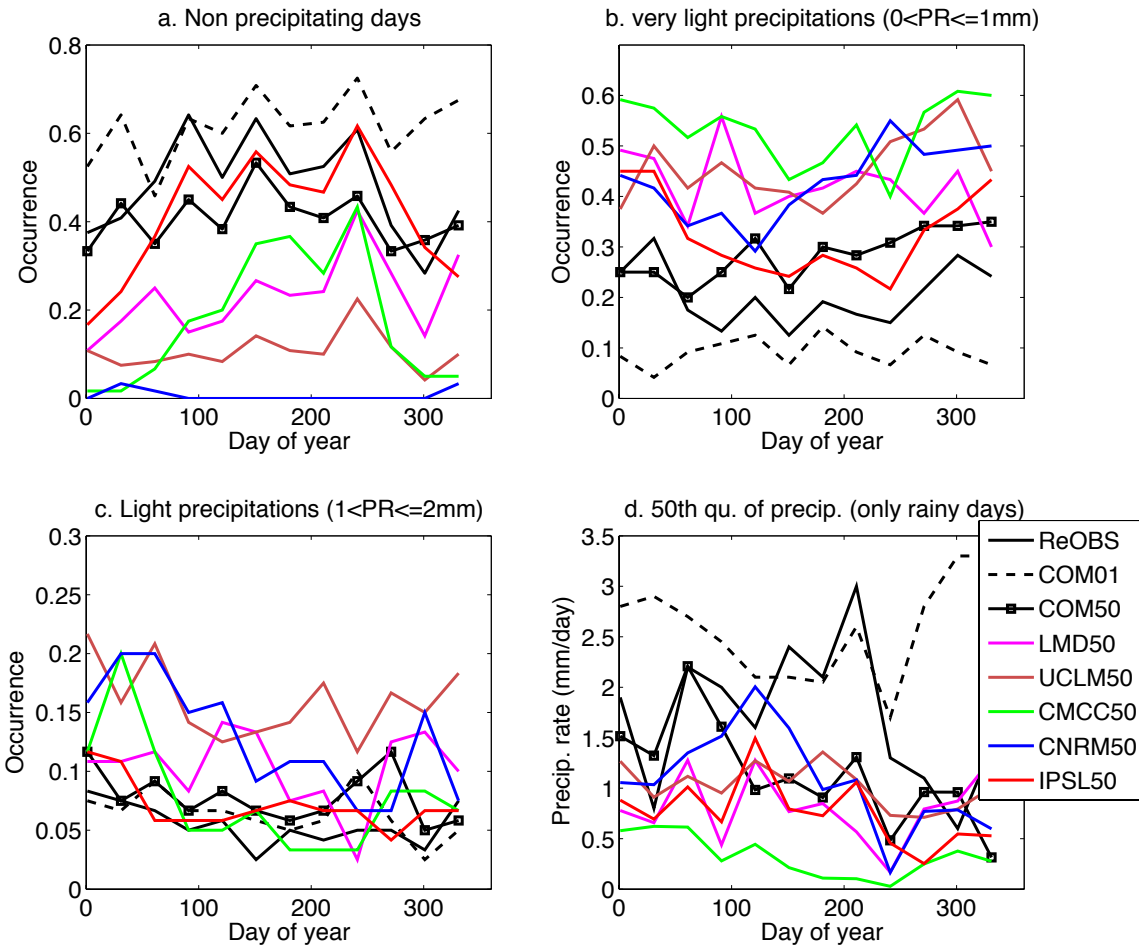

Figure 2 : Occurrence for 30-day periods from January to December over the years 2004-2007 of a. Non precipitating days ; b. precipitation rates between 0 and 1 mm/day ; c. precipitation rates between 1 and 2 mm/day. d. Value of the 50th quantile of precipitation (excluding non-precipitating days). Each color corresponds to a different model at 50 km resolution regridded to the LMD grid (see legend for details). Black line is for observations at SIRTA supersite (ReOBS). Dashed black line is for COMEPHORE product at 1*1 km2 resolution at the closest grid point of SIRTA (COM01) and black line with squares is for COMEPHORE averaged over a square of 50*50 km2 around SIRTA (COM50).

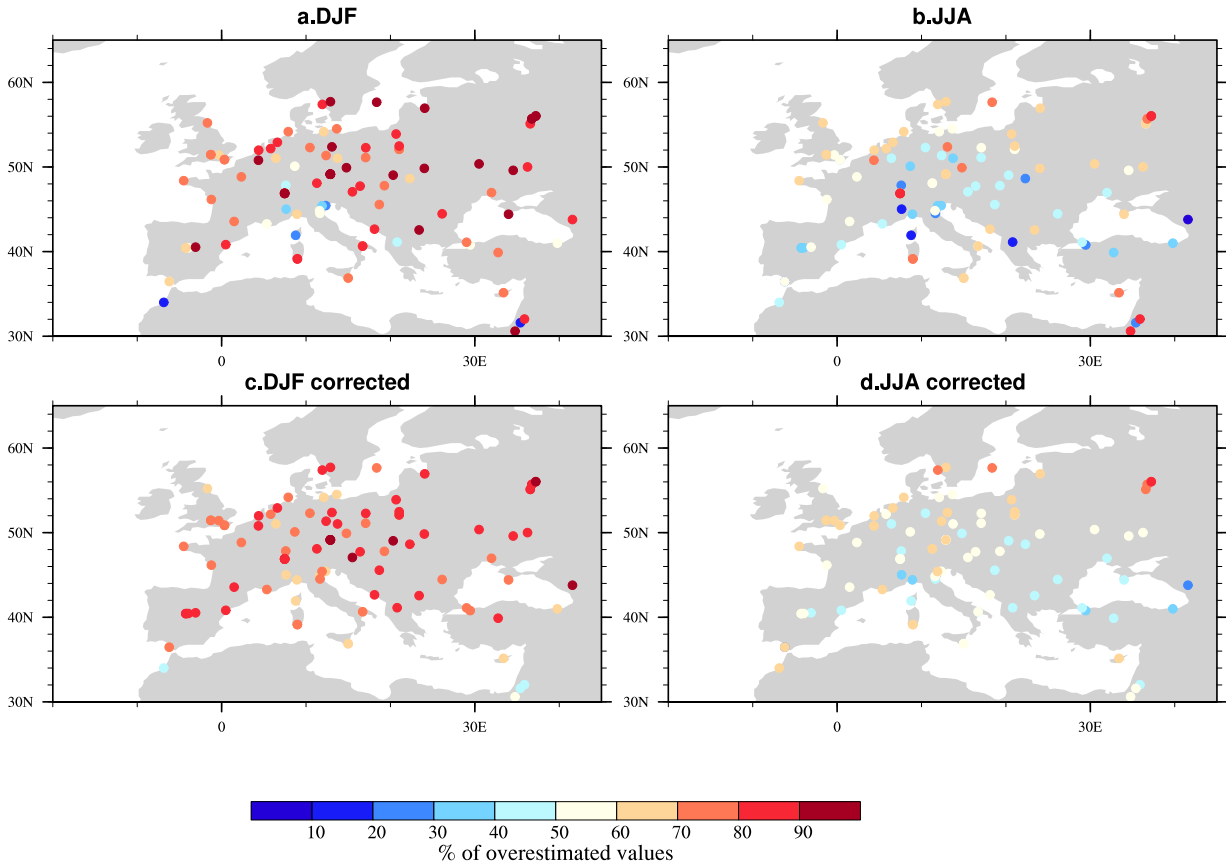

Figure 3: Percentage of simulated daily mean IWV values which overestimate the GPS ones a) in winter and b) in summer. Simulated values are taken from 5 models at 50 km resolution (LMD50, IPSL50, CNRM50, CMCC50, UCLM50) regridded to the ERAI grid. c) and d) same as a) and b) but for height-corrected values.

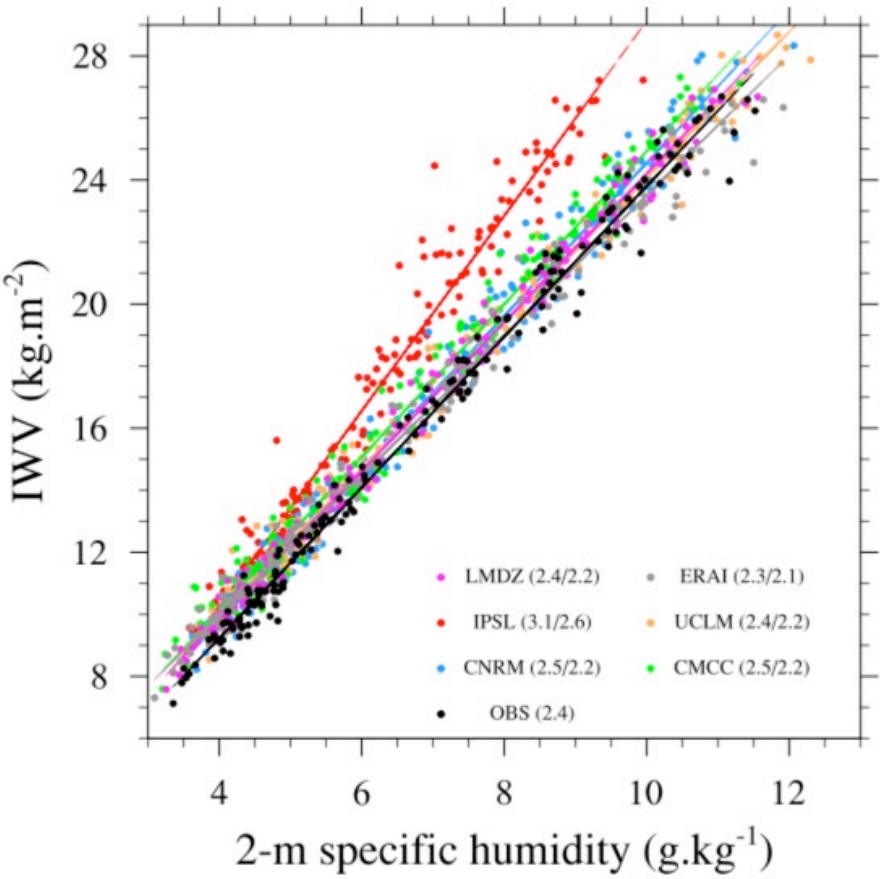

Figure 4: Monthly values of IWV as a function of monthly values of 2-m specific humidity (Q2) averaged over all stations where and when both IWV from GPS and Q2 from HadISD are available (Monthly mean are computed if at least 60 co-existing values exist (i.e about 2 values per day over 4 possible). A total of 3238 months are obtained, spread over 42 different stations. The average number of stations per month is 19 with a maximum of 30 stations). The color of circles corresponding to each model is indicated on the legend. All model outputs have been regridded to the ERAI grid. The first number indicates the slope of the regression obtained when considering the same months than observations at each station, while the second number is the slope of the regression when considering all months at all grid points (only models).

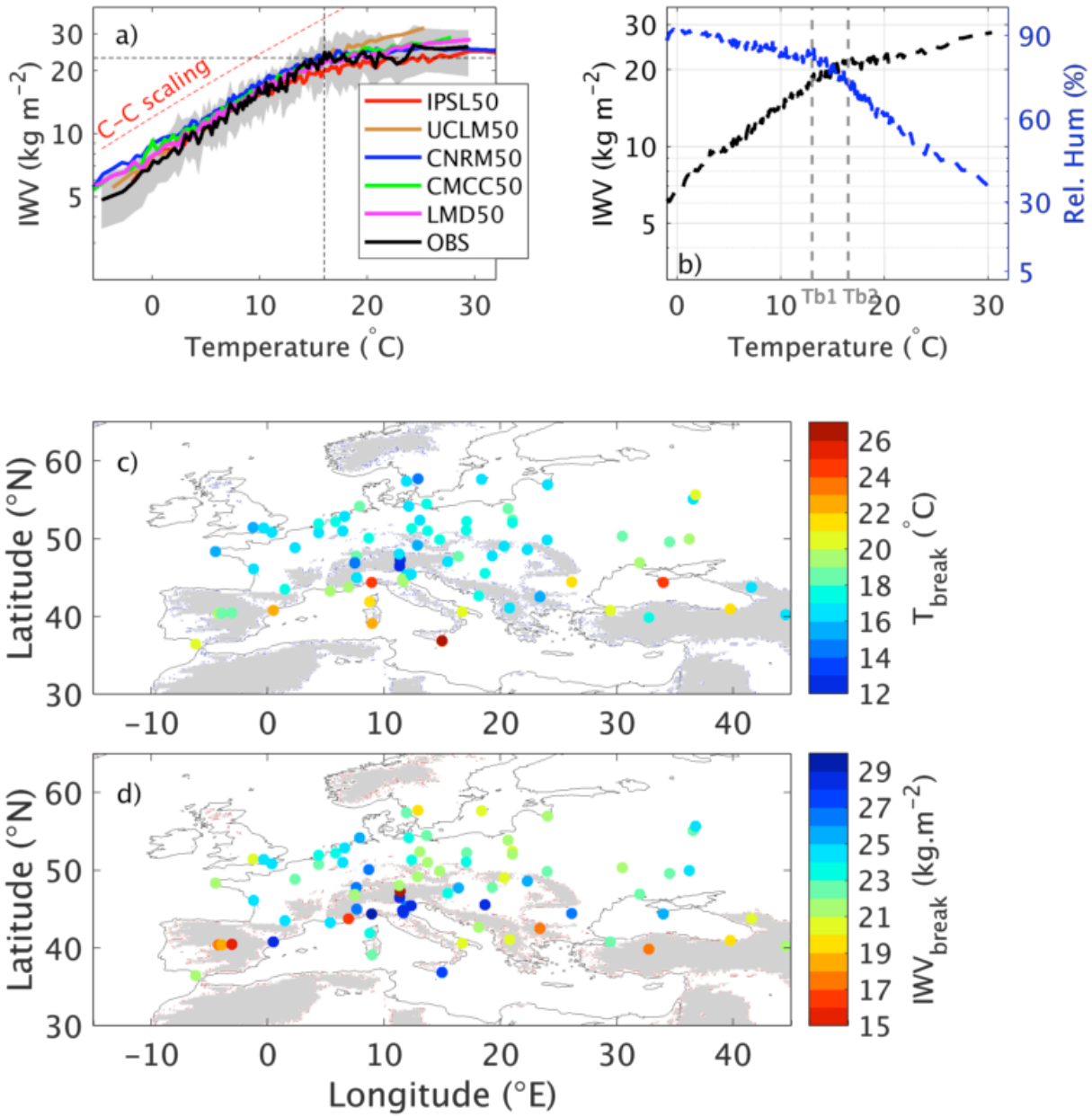

Figure 5: **(a)** IWV – T(2m) relationship at SIRTA station from observations and models. Median values are plotted for observations and models. The grey band represents the interval between the 20th and 80th quantiles for observations. The black vertical dashed line shows T_{break}, the black horizontal dashed line shows IWV_{break}, and the red slant dashed line shows C-C scaling **(b)** Median values of IWV and RH(2m) as a function of T(2m) at SIRTA from observations only. The vertical grey dashed lines indicate the values of Tb1 and Tb2 (see text section 4). **(c)** Map of T_{break} at all the GPS stations. **(d)** Map of IWV_{break} at all the GPS stations.

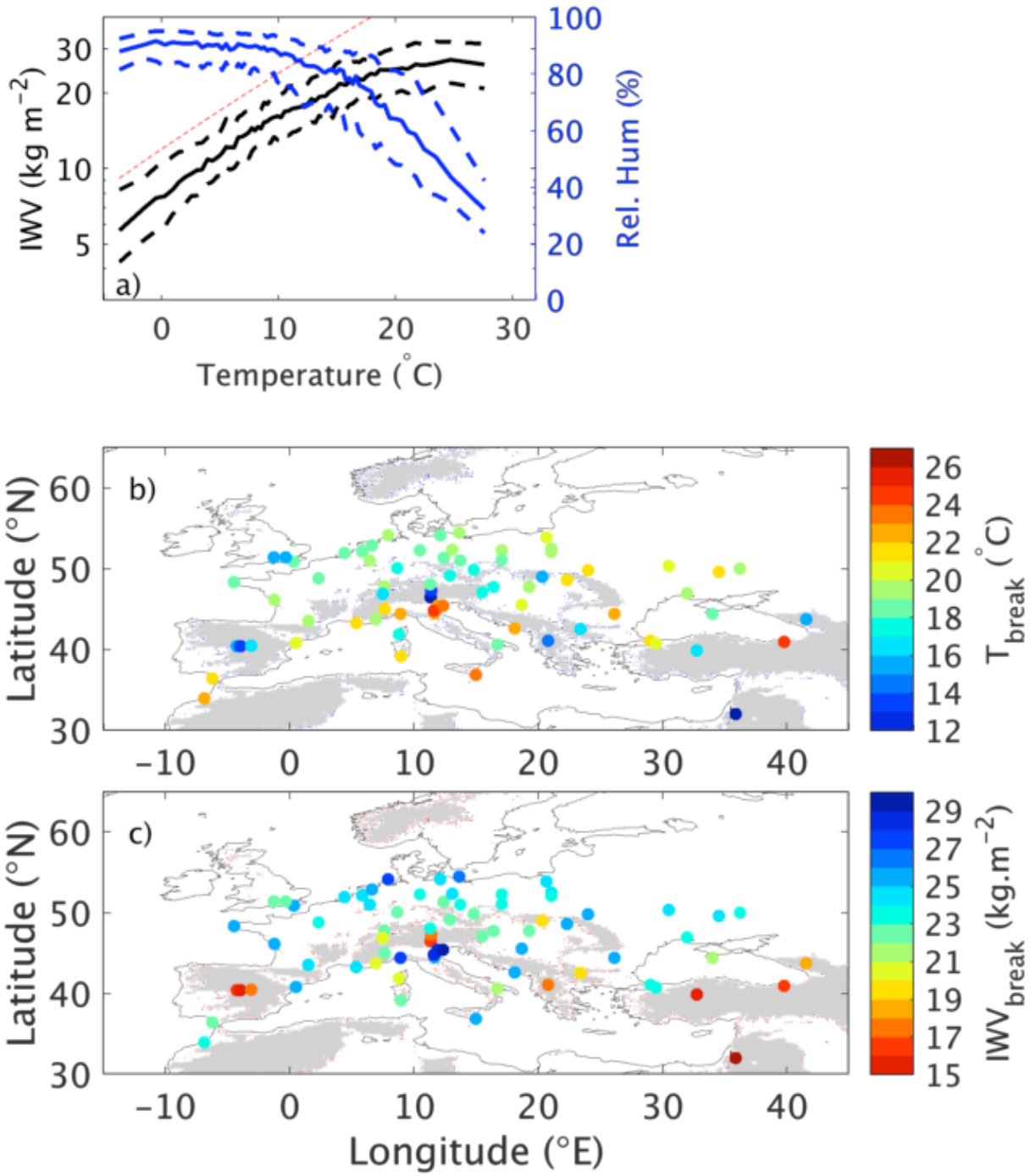

Figure 6 : a) Scaling of IWV (black) and RH (blue) with temperature for the model ensemble at SIRTA station. Solid lines are for quantile 50 and dashed lines are for quantiles 20 and 80 of the distributions. The red slant dashed line shows C-C scaling. b) and c) As Fig5c,d but for the model ensemble.

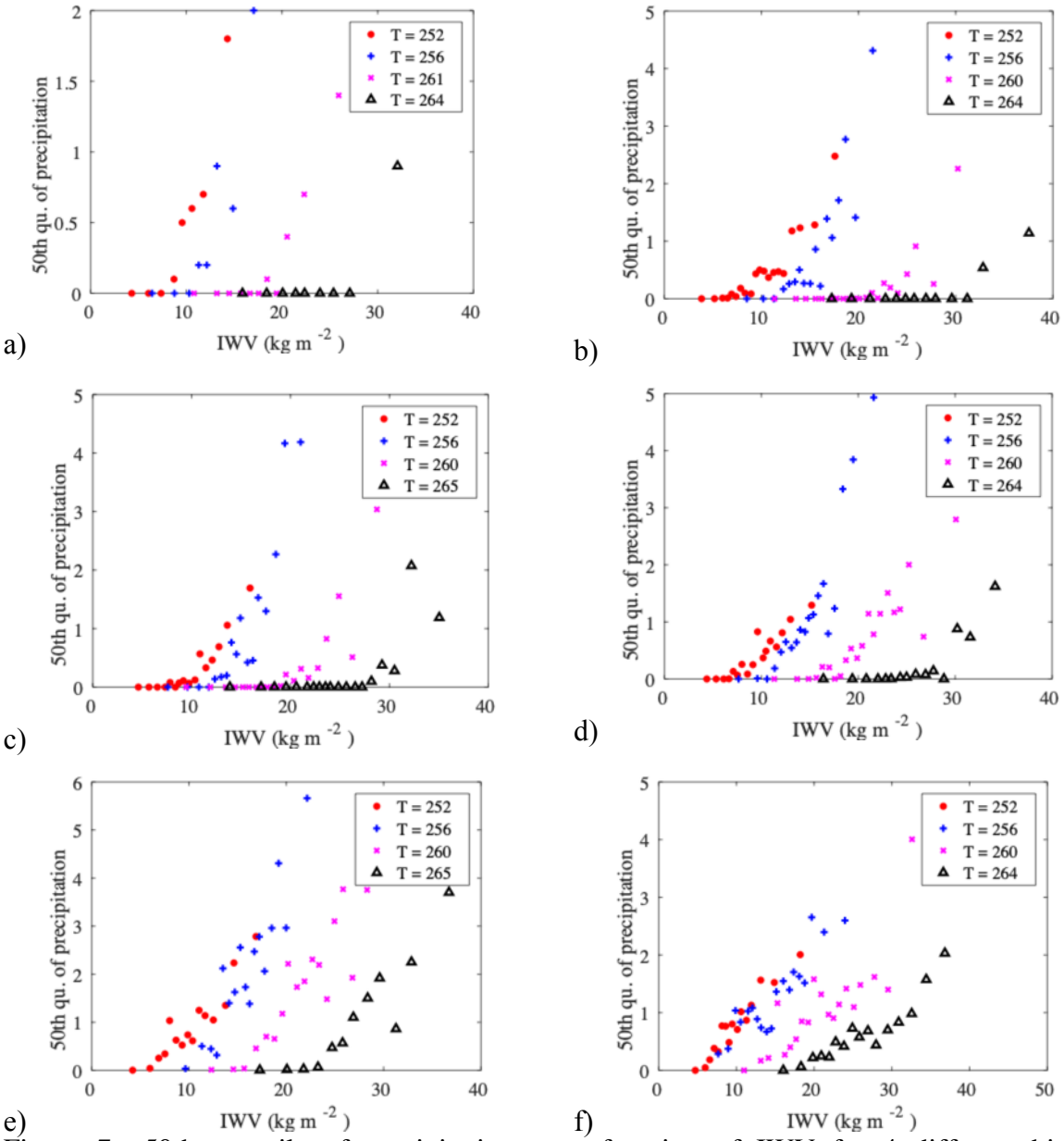

Figure 7 : 50th quantile of precipitation as a function of IWV for 4 different bins of tropospheric temperature at latitude 48.7°N and longitude 2.2°N (SIRTA). **(a)** SIRTA-ReOBS observations. **(b)** CMCC50. **(c)** IPSL50. **(d)** LMD50. **(e)** CNRM50. **(f)** UCLM50. Observations are analysed for the period 2008-2015, and models for the period 2001-2008. The values of T indicated in the legends correspond to the mean values in the bins that have been choosen as indicated in the text.

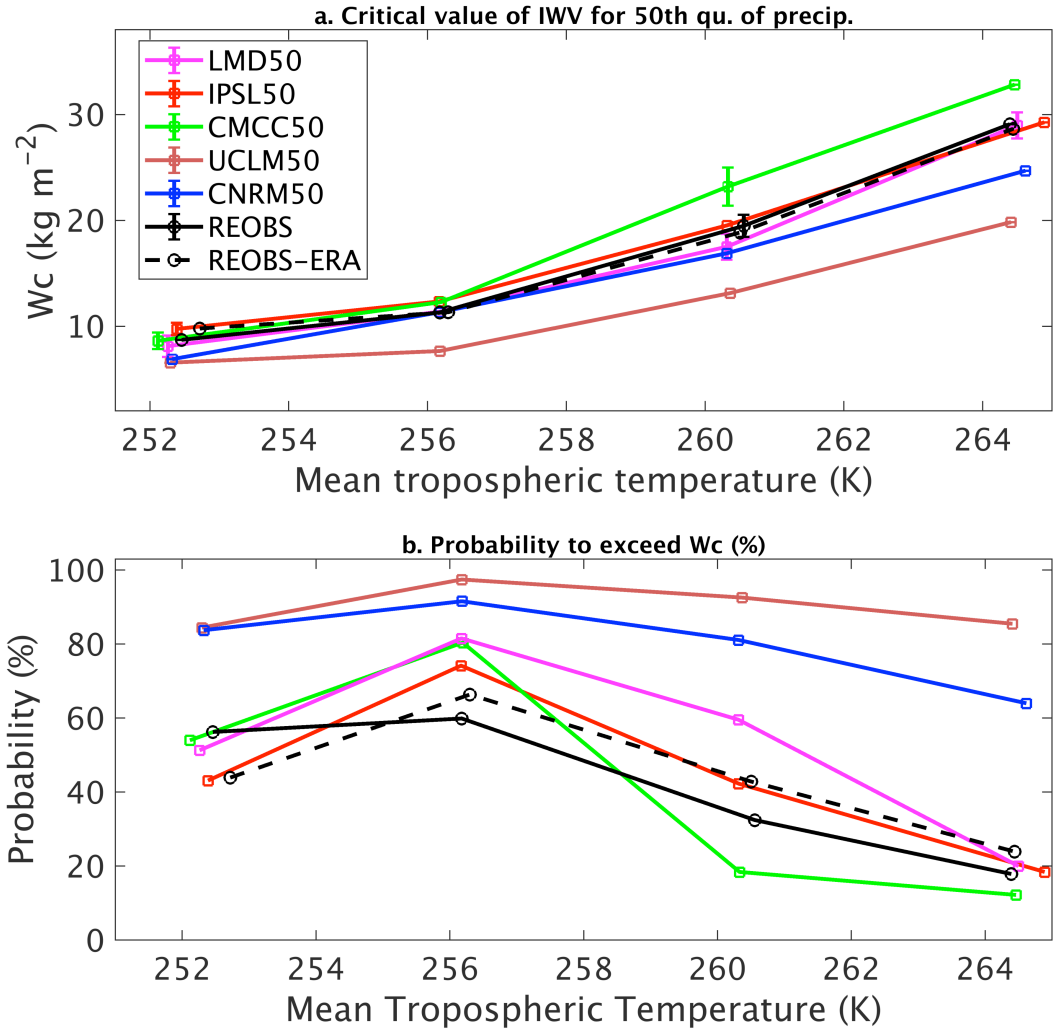

Figure 8 : **(a)** Critical value of IWV at latitude 48.7°N and longitude 2.8°E over which 50th quantile precipitation significantly increases as a function of the mean tropospheric temperature. Each color corresponds to a different model (see legend for details). The period considered is 2001-2008. Solid black line is for observations from the SIRTA-ReOBS dataset between 2008 and 2015 and dashed black line is for SIRTA-ReOBS with ERA-Interim tropospheric temperature instead of tropospheric temperature from radiosoundings. **(b)** Probability for IWV to exceed the critical value for each dataset.

5

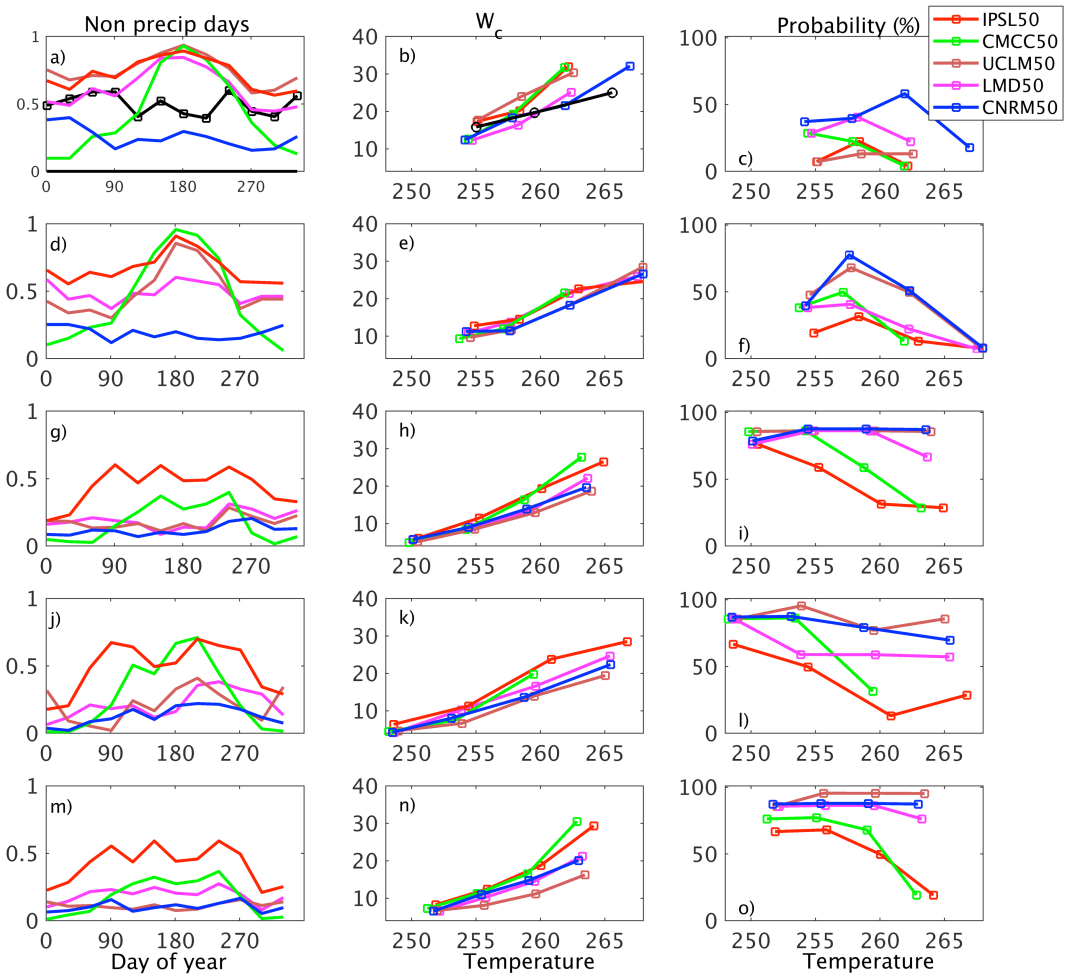

Figure 9 : first column : Annual cycle of occurrence of non-precipitating days for the different models and COMEPHORE in black in panel a (same legend than Fig.8 for models). Second column : Value of w_c (maximum value) as a function of tropospheric temperature. Third column : Probability to exceed w_c value. First row is for the station located in southern France (Marseille), second row for the one located in Central Spain (Madrid), third row for station in eastern Germany (Dresde), fourth row in Ukraine, and fifth row in Netherlands (see table 8 and Fig.1 for details on the locations of the stations).