Analysis of total precipitable water from radiometric observations

Author: Eric Bataller Garcia
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Yolanda Solà

Abstract: Total precipitable water (TPW) is defined as the total amount of water vapor in the atmospheric column. Measurements of the solar radiation transmitted through the atmosphere in the water vapor absorption band allow the TPW to be estimated. This work analyzes 8 meteorological stations with more than 15 years of data from the AERONET network, located around the world and with different climates according to the Köppen climate classification. At the stations of Banizoumbou and Sede Boker, both with the same climate, there is a clear difference between their TPW values. This is mainly due to their latitudes, since these stations are respectively located on strong convergence and subsidence zones, together with their average temperature differences. The key factor affecting the TPW at the stations of Izaña and Mauna Loa is their high altitude, which determines the columnar amount of water above the ground and also has an inverse proportionality with temperature. The annual temperature variation, along with other particular factors that affect each station, also end up having a more relevant impact on the TPW values than the climate definition. This leads to the conclusion that, while the Köppen-Geiger climate classification system is useful to describe the climate conditions, it is not enough to justify and discuss other meteorological variables, such as the total precipitable water.

I. INTRODUCTION

Water vapor is a key element in the Earth's climate system, since it has a very important role in the hydrological cycle and, at the same time, affects the global radiation budget, being the most important gaseous contributor to the greenhouse effect and providing the largest positive feedback in model projections of climate change [1].

The total precipitable water (TPW) is a commonly used term to express the columnar amount of atmospheric water vapor. It is defined as the amount of water vapor integrated vertically in the air column, from the ground to the top of the atmosphere. TPW is normally expressed in kg/m^2, or the equivalent height in a squared meter, in mm or cm, if all the water vapor in the column were condensed.

Because of its importance, accurate water vapor measurements are necessary to the scientific community in order to conduct meteorological and climatological studies. These measurements can be done through various methods, including Sun photometry, radiosondes, microwave radiometers, Global Positioning System receivers and others [**?**].

The TPW measurements derived from AERONET stations are commonly used to validate satellite products, due to their strict quality controls. Different studies have been performed using this data over USA, China and Europe [2–4].

Sun photometry uses the fact that water vapor transmittance can be obtained from direct Sun radiation in spectral channels within water vapor absorption bands, so TPW can be extracted from the irradiance attenuation. For example, the Aerosol Robotic Network (AERONET), which was created basically for studying columnar aerosol properties, makes measurements of the solar direct irradiance and performs TPW retrievals based on measurements in the water vapor absorption band, around 940 nm.

The objective of this work is to characterize and study the total precipitable water in different AERONET stations located in different climate regions around the world, based on their radiative measurements.

II. DATA & METHODOLOGY

AERONET network's standard instruments are Cimel CE-318-4 Sun photometers, which acquire direct Sun and sky radiance measurements at different wavelengths using various interference filters, one of them centered at 940 nm to retrieve the TPW.

Due to the characteristics of the instruments, the TPW measurement is restricted to daytime. Some recent instruments include lunar light measurements, but they are scarce in the network and have not been used in this work.

The total precipitable water, according to its definition, is normally described by the following formula:

$$TPW = \int_0^{p_0} q(p) \cdot dp \qquad (1)$$

where g is gravity's acceleration, $q(p)$ is the specific humidity and p_0 the surface pressure. Specifically, photometers' response to light in this spectral region is given by the following expression, a modified version of the Beer-Bouguer law:

$$V = V_0 d^{-2} exp(-m_r \delta_{atm}) T_w \qquad (2)$$

where $V_0(940\ nm)$ is the instrument calibration constant (needed to obtain the signal that the instrument would measure if it were placed outside the atmosphere), d is the Earth-Sun distance (in astronomical units) at the time of observation, m_r is the relative optical air mass, $\delta_{atm}(940\ nm)$ is the total atmospheric optical depth (excluding the water vapor absorption), and $T_w(940\ nm)$ is the water vapor transmittance around the 940 nm absorption bands. Although the relative air mass, m_r, is approximately the secant of the solar zenith angle, the AERONET's algorithm result takes this into consideration along with several other estimations, including the Earth's curvature and the atmospheric refraction [5].

Finally, AERONET uses the following simplified expression for T_w:

$$T_w(940nm) = exp(-a(m_w TPW)^b) \quad (3)$$

where m_w is the relative optical water vapor air mass [6] and a and b are two coefficients that depend on the central wavelength position, width and shape of the photometer filter function, as well as the atmospheric pressure-temperature lapse rate and the vertical distribution of water vapor.

The data used in this work corresponds to level 2.0 data, the highest data quality level provided by AERONET, which means that all values are automatically cloud cleared and quality assured, with pre-field and post-field calibration applied.

From the 1294 stations available in AERONET's network, the 8 stations selected in this study are the longest time series with level 2.0 data quality, with a length of between 15 and 25 years. In fact, in climatology, 30 years of data are required to characterize the climate of any region but, given the scarcity of this type of data, even with 15-year time series some conclusions can be drawn.

The considered stations are classified using the Köppen-Geiger climate classification system, which assigns the climate of a region based on seasonal precipitation and temperature. The climate classifications, as well as the average temperatures and relative humidity values, are obtained from the *climate-data.org* website [7], which is based on the European Centre for Medium Range Weather Forecast (ECMWF) model, with 0.1-0.25° resolution.

The work's results are represented using monthly box plot analysis, due to their usefulness to visualize how data is distributed. Box plots consist of 4 key components: a central line representing the median value, the box which represents the interquartile range (IQR), between 25% and 75%, then the whiskers, that include the data from these percentiles up to a maximum range of 1.5 times the IQR, and finally any values beyond the whiskers are considered outliers (represented as dots).

III. RESULTS

All the stations considered in this work, along with relevant information about them, are summarized in Table I.

Station (Country)	Latitude	Altitude (m)	Climate
Banizoumbou (Niger)	13.55° N	274	BWh
GSFC, Greenbelt (USA)	38.99° N	87	Cfa
Ispra (Italy)	45.80° N	235	Cfa
Izaña (Spain)	28.31° N	2401	Csb
Mauna Loa (USA)	19.54° N	3402	Cfb
Sede Boker (Israel)	30.85° N	480	BWh
Sevilleta (USA)	34.35° N	1477	BSk
Venise (Italy)	45.31° N	10	Cfa

TABLE I: Latitude, altitude and climate, according to Köppen classification obtained from *climate-data.org*'s website, for each meteorological station.

Even though the majority of the considered stations are located far apart from each other, some of them have the same climate, according to Köppen classification system. One could expect to find similar values of the TPW between these stations but there are several other factors to take into account, so this is not always true.

Banizoumbou and Sede Boker are classified as hot deserts (BWh) but quite different results for the TPW are found (Fig. 1).

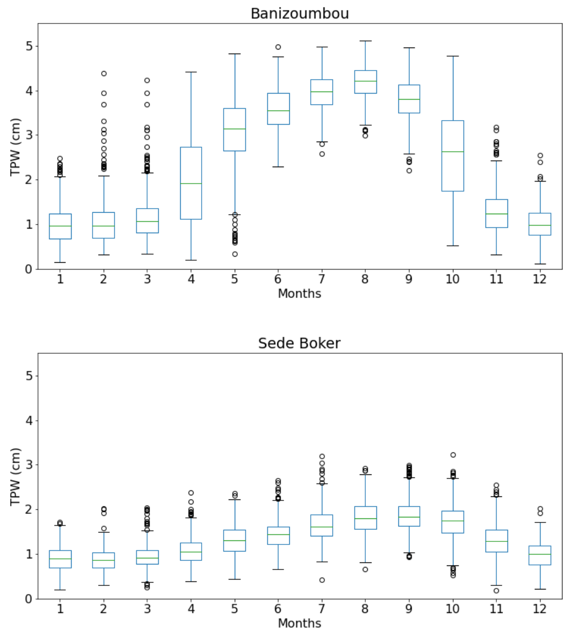

FIG. 1: Monthly box plot analysis of TPW of Banizoumbou's and Sede Boker's stations.

In Banizoumbou's box plot analysis there is a clear seasonal cycle, with low median measurements for the winter, early spring and late autumn months. These values increase, with a large variability, during late spring and early autumn and show very high peaks mostly during the summer months. This cycle can also be observed in Sede Boker's analysis but with lower median values over-all, with less variability throughout the year and narrower interquartile range.

There is a clear difference between these stations' latitudes, that in this case greatly affects the values obtained for the TPW. Sede Boker is located at 30°N, which is right on a strong subsidence zone, caused by the wind currents between the Hadley and Ferrer cells. This phenomenon hinders the cloud formation in this latitude, which in consequence reduces precipitation and therefore the evapotranspiration.

This global-scale atmospheric circulation also offers an explanation for Banizoumbou's TPW, that has very high values for a desert climate. In this case, even though this station is located in the middle of the northern hemisphere's Hadley cell, the air rising in the intertropical convergence zone between the north's and south's Hadley cells, at 0°, oscillates throughout the year up to a maximum latitude of approximately 15°, probably provoking the great increase in total precipitable water that is observed at this station during the summer. This also causes a considerable increase in precipitation and relative humidity (from around 20% up to more than 50%, on average) during this short period of time.

With both stations located in hot deserts, the average temperatures are expected to be high throughout the year, but that is especially true for the Banizoumbou station. This is an important factor to take into consideration, because while the temperature does not impact the TPW by itself, the fact that the saturation vapor pressure is proportional to temperature does have a notable effect on the columnar amount of precipitable water.

In the Nigerien station, the temperatures ranges from 25 up to 34°C on average, reaching the highest values from late spring to early autumn months and coinciding with the largest variability observed. The temperature is only slightly colder in the summer season due to the great increase in precipitation, as previously stated. On the other hand, the temperatures in Sede Boker only oscillate between 10 and 26°C on average, with its maximum values registered in summer, a period of time with almost no precipitation and low relative humidity due to the subsidence zone also commented before.

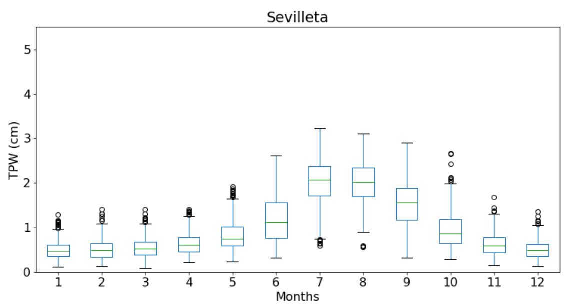

FIG. 2: Monthly box plot analysis of TPW of Sevilleta's station.

Regarding Sevilleta station's values, which is also considered a dry climate but colder and semi-arid (BSk), Fig. 2 shows median values similar to Sede Boker's but with a narrower IQR except for the summer months. During this season, TPW experiences a sudden increase in median values along with larger variability, similarly to Banizoumbou's results, that in consequence shows a much clearer cycle throughout the year.

The values obtained can be partially justified by the region's latitude, which is not far from Sede Boker's strong subsidence zone. They share a similar annual temperature variation as well, which oscillates between 3 and 27°C in this case, also with its maximum values registered during the summer, but not leading to such a decrease in precipitation and relative humidity to completely explain the results obtained.

The key issue is the station's altitude (Table I), which is very important to understand the range of values obtained. According to Eq. 1, TPW is defined from the surface to the top of the atmosphere (TOA), where the pressure is approximately equal to 0. Since this station is located at 1477 m, the values obtained by integrating the column of water vapor have to be significantly lower than those acquired closer to sea level.

Furthermore, stations located at higher altitudes generally tend to have colder climates, due to the decrease of temperature with altitude. This also indirectly decreases the precipitable water in the air column, due to the saturation vapor pressure dependence with temperature, as previously mentioned.

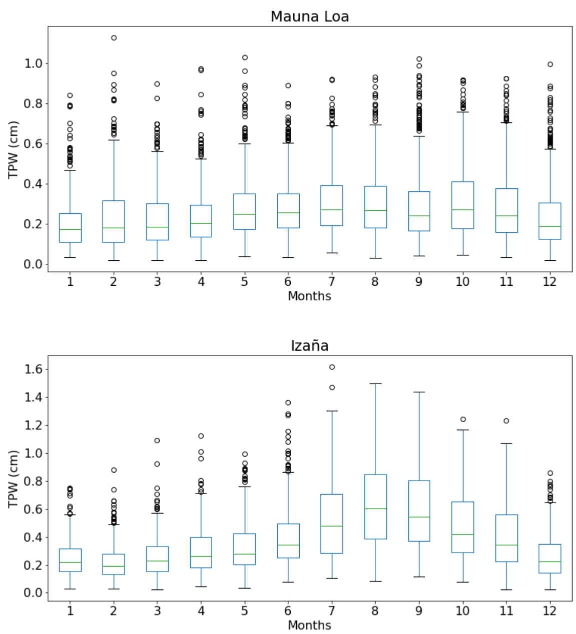

FIG. 3: Monthly box plot analysis of TPW of Mauna Loa's and Izaña's stations. In this case the scale has been modified to match the low values obtained, otherwise the results might be more difficult to observe.

Altitude has such an important role that it characterizes the climate conditions of two of these study's stations, Mauna Loa and Izaña, classified as Csb and Cfb respectively. Both are classified in the b subgroup, since their monthly temperatures are lower than 22°C and at least four months have temperatures higher than 10°C. Their climates are heavily conditioned by their altitude, as observed with the TPW results.

Following the reasoning stated before, the data shown in Fig. 3 has the lowest values of all the stations studied in this work by a clear margin. Because of this fact, the box plot analysis shows the median values biased towards lower values (positive skewness), with an appreciable difference of data distribution between values above and below the median. The quartile and whisker above it have a much broader range throughout the year, due to the large amount of low values obtained. This also justifies the high number of outliers observed, specially for Mauna Loa, because the maximums are slightly biased down by the values considered as "normal". The considerable amount of near-zero data obtained at these stations means that much "unusually high" data ends up classified as outliers.

Analyzing the Izaña station plot in Fig. 3, the seasonal cycle observed previously in all other stations is still visible, with similarly low median values for most of the year and a noticeable increase in median values and TPW variability during the summer months. That is not the case for Mauna Loa, located around 1 km above the already high altitude of Izaña, the region's precipitable water values are so low that a seasonal cycle cannot be observed in the Hawaiian station.

Finally, there is a last group of 3 stations with the same Köppen climate classification, Cfa. These humid and temperate regions are located at low altitudes, so in this case it is not a factor to take into consideration onto the results discussion.

Fig. 4 shows very clear seasonal cycles but with an appreciable difference between the GSFC (located at the NASA's Goddard Space Flight Center, in Greenbelt) and the Italian stations. In fact, Ispra and Venise stations share such similar data that they are treated in the same way hereafter.

Although the TPW values may seem visually unalike when compared, this is mainly due to the high variability of the data obtained at the GSFC station. Regarding the median values exclusively, the results obtained for these regions are actually quite similar, both showing low values throughout half of the year, with an increase from late spring to early autumn and reaching almost the same maximum values during the summer.

Since these stations are located at mid latitudes without any clear convergence or subsidence zones, a key difference between them is their annual temperature variations. While in the Italian stations it ranges from 2 to 27°C, the temperature values at the GSFC station range from -2 up to 30°C, increasing the variability of the TPW. Due to the high number of low values, especially from late autumn to early spring, the GSFC station median values are biased downward as in the mountain climate. This leaves much data represented as outliers because the maximums, constrained by a narrower interquartile range during these months, are also downward biase, as explained before.

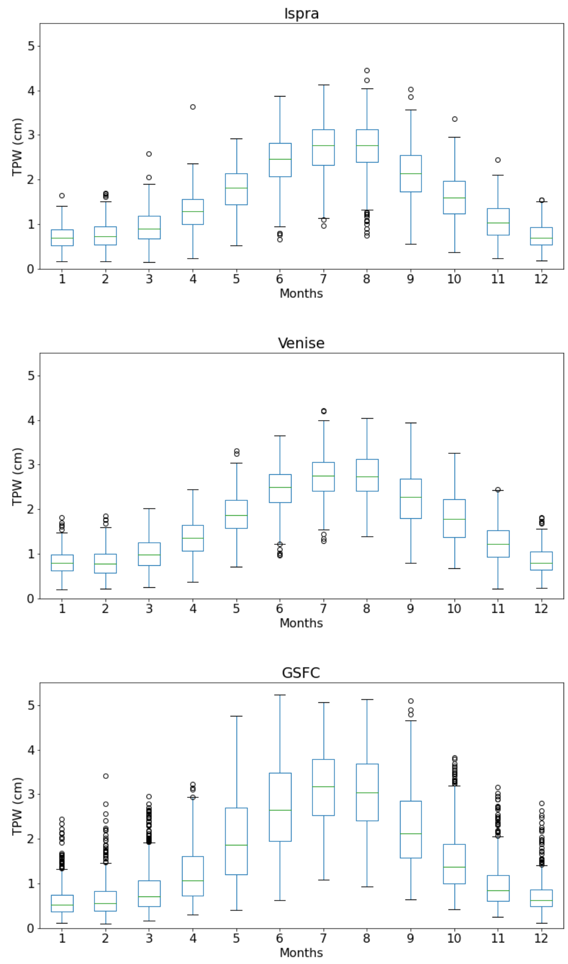

FIG. 4: Monthly box plot analysis of TPW of Ispra's, Venise's and GSFC's stations.

The maximum precipitable water values obtained at GSFC's station are, along with those of Banizoumbou, the highest values observed in this study. Unlike the Nigerien station's case, the median values in the Greenbelt station remain relatively low, while the variability increases.

A plausible explanation for this phenomenon is its geographical location: Greenbelt is located at approximately 40 *km* from Chesapeake Bay, which constitutes a large body of oceanic water. The Ispra station for example, is located even closer to a large body of water, the Lake Maggiore. Following the same reasoning as before, this further affects the TPW variability, since Chesapeake Bay averages temperatures from 5 to 26°C during the year and the lake Maggiore only ranges from 6 to 23°C, according to [8]. The last main difference between these 2 cases is due to the global prevailing winds that transport humid air masses in this latitude range. At the GSFC station, these westerly-dominant winds transport dry air masses from the continent, allowing precipitable water to accumulate without precipitation. Since the Italian stations do not have a comparable amount of continental land to bring dry air, the influence is not so noticeable and the difference between them is accentuated.

IV. CONCLUSIONS

With the aim of analyzing the total precipitable water values obtained by different stations around the globe, in this work we have studied several characteristics of each location that, directly or indirectly, affect the data acquired.

We have concluded that the TPW depends on several different variables, such as the altitude, which not only directly affects the columnar amount of water vapor above ground, but also has an inverse proportionality with temperature. This fact, along with the saturation vapor pressure dependence on temperature, significantly decreases the precipitable water in the air column.

Temperature itself influences the annual variability of total precipitable water and it is essential to partially understand the biased median values in some particular stations.

Although it is not a significant characteristic to discuss in some cases, due to its relatively low impact at some of the stations studied in this work, latitude does play a key role at stations that are directly affected by the global circulation. The existence of strong subsidence or convergence zones at certain latitudes, as between two Hadley cells or the Hadley and the Ferrel cell, clearly characterizes the total precipitable water values in a region.

Finally the last general characteristic considered in this work is the climate classification. Following the Köppen-Geiger climate classification system, which takes into account the seasonal precipitation and temperature of the studied zones, one could expect to observe similar TPW between stations with the same climate. As we have shown in this work, that is generally not true.

In this study, the differences between the precipitable water values have been justified by the particular conditions of each station, with more or less impact of each of the aforementioned variables, but with clear differences (in median values, variability and maximum or minimum values) between stations with the same climate classification.

This leads us to the conclusion that, while the Köppen-Geiger climate classification system is a very useful tool for describing the climate conditions and characterizing some region's properties, it is clearly insufficient to justify and discuss other meteorological variables, such as the total precipitable water.

Acknowledgments

Firstly I would like to acknowledge AERONET for the provision of data and the site managers for their maintenance. Then I would like express my gratitude to my advisor, Y. Solà. She has been very helpful and understanding throughout all the process of this work. Finally I have to thank G. Camacho, for her help at the beginning of the study and for her continuous support during this work, along with that of my friends and family.

[1] Held, I. M., and Soden B. J., *Water vapor feedback and global warming*, Annual Review of Energy and the Environment, 25, 441–475, 2000.

[2] Pérez-Ramírez, D., Whiteman, D.N., Smirnov, A., Lyamani, H., Holben, B. N., Pinker, R., Andrade, M. and Alados-Arboledas, L., *Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites*, Journal of Geophysical Research: Atmospheres, 119(15), 9596-9613, 2014.

[3] Gui, K., Che, H., Chen, Q., Zeng, Z., Liu, H., Wang, Y., Zheng, Y., Sun, T., Liao, T., Wang, H. and Zhang, X., *Evaluation of radiosonde, MODIS-NIR-Clear, and AERONET precipitable water vapor using IGS ground-based GPS measurements over China*, Atmospheric Research, 197, 461-473, 2017.

[4] Campanelli, M., Mascitelli, A., Sanò, P., Diémoz, H., Estellés, V., Federico, S., Iannarelli, A. M., Fratarcangeli, F., Mazzoni, A., Realini, E., Crespi, M., Bock, O., Martínez-Lozano, J. A. and Dietrich, S., *Precipitable water vapour content from ESR/SKYNET sun–sky radiometers: validation against GNSS/GPS and AERONET over three different sites in Europe*, Atmospheric Measurement Techniques, 11, 81–94, 2018.

[5] Kasten, F., *A new table and approximation formula for relative air mass*, Archiv für Meteorologie, Geophysik und Bioklimatologie Serie B, 14, 206-223, 1965.

[6] Smirnov, A., Holben, B.N., Lyapustin A., Slutsker, I. and Eck, T.F., *AERONET processing algorithms refinement*, AERONET Workshop, El Arenosillo, Spain, 2004.

[7] Climate-Data. Available at *https://es.climate-data.org/info/sources/* (Accessed: 16 June 2022).

[8] Sea Temperature. Available at *https://seatemperature.info/* (Accessed: 16 June 2022).