Comparative Analysis of ERA5 Precipitable Water Vapor and Precipitation Data Against Ground Observations and GNSS Measurement in Thailand

Sothyda Khor[1], Chaiyut Charoenphon[2,3] and Supattra Visessri[1,4,*]

[1]Department of Water Resources Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
[2]Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
[3]Mapping and Positioning from Space (MAPS) Technology Research Center, Chulalongkorn University, Bangkok 10330, Thailand
[4]Center of Excellence in Disaster and Risk Management Information Systems, Chulalongkorn University, Bangkok 10330, Thailand
**Corresponding author; E-mail address: Supattra.vi@chula.ac.th*

Abstract

ERA5 reanalysis data and Global Navigation Satellite System (GNSS) provide valuable atmospheric and hydrological data for weather and climate analysis. Both ERA5 and GNSS-derived precipitable water vapor (PWV), along with precipitation data, have potential applications in predicting extreme weather events. This study evaluates the accuracy of ERA5 precipitation data by comparing it with ground-based observations and ERA5 PWV data by comparing it against GNSS PWV measurements at six stations across different geographical regions in Thailand over the period of 2019 to 2023. Each precipitation dataset is used to compute the precipitation efficiency (PE), defined as the percentage of total precipitable water vapor that falls as measurable precipitation. The results indicate that ERA5 effectively captures seasonal precipitation and PWV patterns, but it underestimates high-intensity rainfall, particularly in coastal and complex terrains. In contrast, inland regions show better alignment with ground-based data. For PWV, ERA5 generally underestimates values in mountainous and coastal regions but performs better in flatter, more stable areas. Consequently, PE values from ground-based data show strong regional variation. These results make ERA5 particularly useful for climate monitoring and large-scale weather forecasting in low-lying or stable regions of Thailand. Its consistent spatiotemporal coverage, combined with PWV data, remains valuable for flood forecasting, especially in data-scarce areas for enhanced early warning systems.

Keywords: ERA5, GNSS, PWV, PE, Extreme weather events

1. Introduction

Precipitable Water Vapor (PWV) represents total atmospheric water vapor in a vertical column, reflecting its potential to contribute to precipitation under suitable atmospheric conditions. Processes such as precipitation, floods, and deep moist convection are directly influenced by the amount of PWV, which varies significantly across time and space [1]. Hence, accurate measurement of precipitation and PWV is crucial for studying weather patterns, predicting extreme events, and managing water resources effectively. Typically, precipitation and PWV can be measured directly through ground-based observations or indirectly through satellite, remote sensing, and reanalysis datasets, with each method presenting different advantages and limitations [2, 3].

On the one hand, the most common and reliable way to directly measure precipitation at point scale is ground-based rain gauge stations [2, 4], but this approach is constrained by the uneven distribution and sparse spatial coverage of the stations, especially in developing countries [5]. Satellite remote sensing and reanalysis datasets overcome the shortcomings of the sparse distribution of ground-based sites, however, those methods derive the data from indirect precipitation measurements and simulations. Satellite-derived precipitation products are often affected by errors and uncertainties. The physical constraints of the sensor itself cause some errors and uncertainties in satellite precipitation observations, which in turn affect resolution, sampling, accuracy, precision, and noise [6].

On the other hand, for PWV measurement, radiosondes is the primary ground-based measurement to detect PWV due to its high accuracy; however, it is limited due to its high cost, low spatiotemporal resolution, and ineffective sensor performance in dry and cold conditions [7]. Currently, global navigation satellite systems (GNSS) have proven to be an effective method for determining PWV due to its low cost, resistance to all weather conditions, long-term stability, high temporal resolution, and high precision [8]. Although both radiosondes and GNSS can provide sufficiently accurate estimates of PWV content, their ground-based observations remain restricted to terrestrial regions with limited spatial coverage [7]. In contrast, remote sensing satellites provide wide-area PWV detection, but infrared remote sensing cannot retrieve data in cloudy conditions [9]. While microwave remote sensing can penetrate clouds, PWV retrieval remains challenging due to the complexity of changing land surface conditions [10].

Furthermore, reanalysis product, such as ERA5 released by the European Center for Medium-Range Weather Forecasts (ECMWF) provides hourly estimates for a large number of atmospheric datasets. ERA5 data provides comprehensive atmospheric information across multiple vertical levels, integrating observations with advanced numerical weather models, making it valuable for improving weather prediction accuracy [11]. ERA5 can provide continuous atmospheric observations over multiple decades and serve as an alternative source of data for precipitation analysis and PWV retrievals. ERA5's global availability improves disaster preparedness by improving the understanding of historical extreme events and supporting the development of early warning systems. However, the accuracy of the ERA5 dataset must be evaluated against ground-based data to determine its reliability in extreme event prediction.

Therefore, this study aims to evaluate the accuracy of ERA5 data in Thailand by pursuing the following objectives: (1) to compare ERA5-derived precipitation with ground-based observations, (2) to compare ERA5-derived PWV with GNSS-derived PWV from ground-based stations, and (3) to determine precipitation efficiency (PE) using both ERA5 and ground-based observation datasets. Findings from this study can support improved flood forecasting in Thailand because ERA5 provides continuous, high-resolution atmospheric data, and ERA5-derived PE can serve as an important indicator for evaluating rainfall potential and flood risk at different locations in Thailand.

2. Data and methods

2.1 Study area

The Kingdom of Thailand is located between 5°37′N and 20°37′N latitude-wise and between 97°22′E and 105°37′ E longitudinally. There are 22 major river basins across the country's total area of 514,050 km^2 [12]. This study focuses on six geographical regions of Thailand, which were divided based on their natural features, including landforms and drainage, as well as human cultural patterns [13]. The regions are Northern, Central, Western, Northeastern, Eastern, and Southern. As displayed in **Fig. 1**, the topography of Thailand comprises the highest mountain ranges in the North and West with central plains generally flat and low-lying.

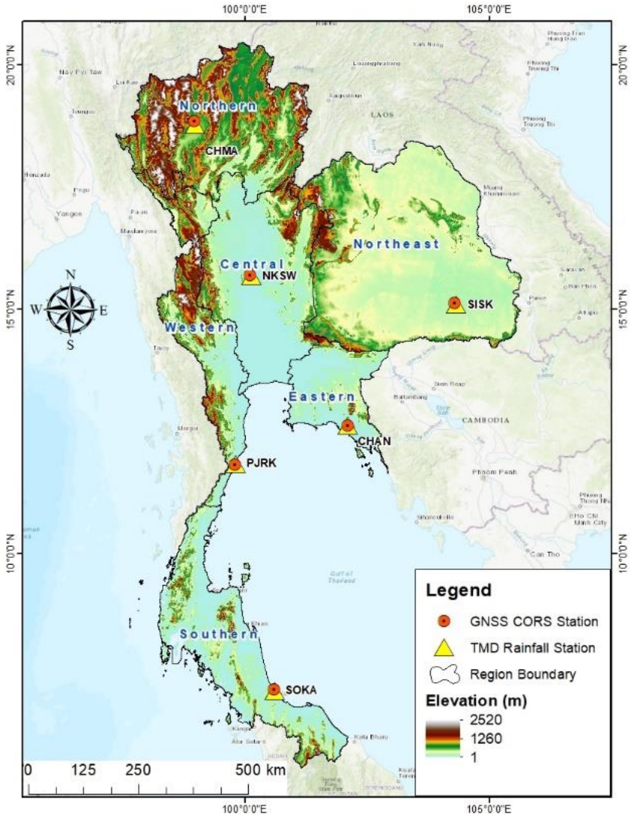

Fig. 1 Study area

Generally, Thailand has three main seasons [14]. Rainy Season [15] (May to October) is caused by the southwest monsoon carrying moist air from the Indian Ocean, impacting the West Coast first, then spreading to central and eastern Thailand. Tropical cyclones bring additional rain from August to October. Winter Season (mid-October to mid-February) is influenced by the northeast monsoon, bringing cool and dry air from the Siberian anticyclone over major parts of the country, while humid eastern winds affect the East Coast of peninsular Thailand. leading to more rain than the western side. December is the driest month. Lastly, Summer Season (mid-February to mid-May) is a transition period between monsoons, with April as the hottest month.

In this study, one ERA5 precipitation and PWV dataset was selected for each of six geographical regions. The closest TMD and GNSS station was used as the primary reference to evaluate the performance of ERA5 precipitation and PWV, accordingly.

The selected referenced GNSS stations have been used in previous GNSS studies in Thailand for various analyses [10, 16-18]. Based on their geographical locations, CHMA represents the northern region, NKSW represents the central region, SISK represents the northeastern region, CHAN represents the eastern region, PJRK represents the western region, and SOKA represents the southern region. CHAN, SOKA, and PJRK stations showed the strongest correlation in Zenith Tropospheric Delay (ZTD) values between the modified BNC software, which was applied to compute GNSS observation data and evaluated its accuracy, and CSRS-PPP as the referenced data [17]. Since PWV is derived from ZTD, these stations were considered suitable for evaluating the correlation between GNSS-based PWV and ERA5-based PWV in this study. Additionally, CHMA, NKSW, and SISK were selected because they exhibited a good correlation in ZTD values between the BNC model and CSRS-PPP while also being geographically representative of the region. Their PWV estimates were also within the accuracy threshold (RMSE < 3 mm) for weather nowcasting [19], making them relevant for this study.

2.2 Dataset

Table 1 provides an overview of all datasets utilized in this study over the period of 2019-2023. Daily precipitation data from six stations were obtained from Thai Meteorological Department (TMD). PWV data were collected from six GNSS stations, with measurements recorded at 5-minute intervals and later converted into daily values for analysis. Additionally, ERA5 reanalysis data were used to obtain both precipitation and PWV values from six grids (and extract as points at center of grids), available at an hourly temporal resolution. Both precipitation and PWV data from ERA5 were aggregated into daily values for comparison. Each ERA5 data point represents the center of a gridded dataset with a resolution of 25° × 25°, with the closest grid point selected to match each GNSS station. Overall, the precipitation dataset from TMD and the PWV dataset from GNSS, serving as ground-based observations, were used as reference data to evaluate the accuracy of ERA5 reanalysis products over a five-year period (2019–2023) at a daily timescale. Furthermore, **Table 2** provides details of the GNSS stations used as reference for PWV measurements, including station names, regions, coordinates, altitude, and annual precipitation.

Table 1 Datasets used for evaluating the accuracy of ERA5 product

Source	Precipitation	PWV	Temporal Resolution	Duration
TMD	6 stations	-	daily	2019-2023
GNSS	-	6 stations	Every 5 minutes	2019-2023
ERA5	6 center of grid points	6 center of grid points	hourly	2019-2023

Table 2 Referenced GNSS station information and average precipitation in 2019-2023

GNSS station	Latitude	Longitude	Altitude (m)	Annual precipitation (mm)
CHMA	18.84	98.97	334.43	1049
NKSW	15.69	100.11	54.82	1293
SISK	15.12	104.29	127.98	1433
CHAN	12.61	102.1	31.18	2293
PJRK	11.81	99.8	16.68	1487
SOKA	7.21	100.6	34.45	2142

2.2.1 Retrieval of PWV from Ground-based GNSS

The calculation of PWV in this study focuses specifically on the tropospheric layer due to its significant impact on signal propagation in the context of GNSS-based PWV estimation. The troposphere contains the majority of the Earth's water vapor, and its impact on signal delay is more directly related to the variations in temperature, pressure, and humidity, which are key factors in PWV estimation [20]. The troposphere introduces an additional delay in GNSS signals as they travel from satellites to ground-based receivers [21]. The GNSS receivers are called Continuously Operating Reference Stations (CORS) (as shown in **Fig. 2**), which were selected as the reference locations in this study.

These delays are typically modeled along each satellite's line of sight and then mapped to a single parameter, known as ZTD. Firstly, to estimate ZTD, Precise Point Positioning (PPP) is applied to raw GNSS data, using high-precision satellite orbit and clock products to determine receiver coordinates and associated atmospheric parameters. Finally, the calculation of PWV relies on ZTD values obtained from PPP with other required meteorological parameters, including surface pressure, temperature, and humidity.

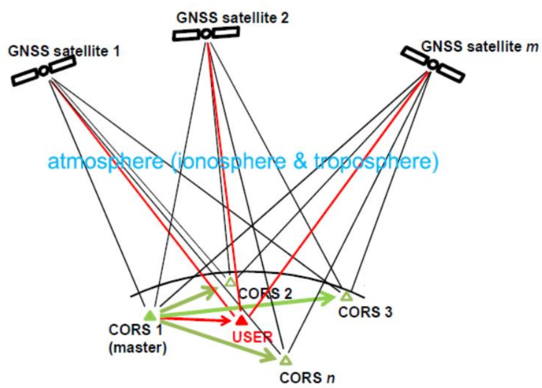

Fig. 2 GNSS signal transmission and CORS station network [22]

ZTD consists of two primary components: the Zenith Hydrostatic Delay (ZHD) and the Zenith Wet Delay (ZWD). Since only ZWD corresponds to atmospheric moisture, it must be separated from ZTD before converting to PWV. To obtain ZWD, ZHD is subtracted from ZTD. This is expressed in Eq. (1).

$$ZWD = ZTD - ZHD \quad (1)$$

The ZHD can be calculated in Eq. (2) as follows [23]:

$$ZHD = 0.0022768 \times \frac{P}{f(\varphi, H)} \quad (2)$$

$$f(\varphi, H) = 1 - 0.00266\cos 2\varphi - 0.00028H \quad (3)$$

where P is the air pressure at the station in hPa, φ is the latitude of the station, and H is the altitude of the station in km.

It is necessary to multiply a dimensionless conversion factor Π in order to convert ZWD to PWV [24], which may be found in Eq. (4).

$$\Pi = \frac{10^5}{R_\upsilon \times (k_3 / T_m + k_2^{'})} \quad (4)$$

where R_υ is the specific gas constant of water vapor which equal to $461\ J.kg^{-1}.K^{-1}$. $k_3, k_2^{'}$ are the atmospheric refractive index constants, which $k_3 = 3.754 \times 105\ \mathrm{K}^2.\mathrm{hPa}^{-1}$ and $k_2^{'} = 71.98\ \mathrm{K.hPa}^{-1}$. T_m is the weighted average temperature of the atmosphere

After calculating Π and ZWD, PWV can be finally derived by the following equation:

$$PWV = \Pi \times ZWD \quad (5)$$

2.2.2 Retrieval of precipitation and PWV data from ERA5

ERA5 is the fifth-generation reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 reanalysis data were used to obtain global atmospheric variables with data available from 1940 to the present. In this study, ERA5 precipitation and PWV datasets were retrieved from the Copernicus Climate Change Service (C3S) Climate Data Store via the Application Programming Interface (API) (Climate Data Store).

The variable "total precipitation" (tp) was used for precipitation, and "total column water vapour" (tcwv) was used for PWV. Both variables are provided at an hourly temporal resolution and a spatial resolution of 0.25 degrees, with each grid cell value representing the model output at a single level (over the Earth's surface). Data spanning from January 2019 to December 2023 were extracted and then aggregated from hourly to daily values for analysis.

It is important to note that this study has not yet adjusted the GNSS station-based PWV data to match the ERA5 grid point height. The GNSS data, which is based on the altitude of the station, has been compared directly with the ERA5 data, which is referenced at the surface level. This may lead to discrepancies due to the elevation differences between the GNSS station's altitude and the surface-level reference of ERA5.

2.3 Statistical evaluation methods

This study estimates the spatial performance of ERA5 precipitation and PWV data using statistical indices of bias and root-mean square error (RMSE). Bias represents the average deviation between ERA5-derived values and ground-based measurements, indicating whether ERA5 tends to overestimate or underestimate the observed data. The Bias can be calculated by Eq. (6).

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(X_{ERA5,i} - X_{ground,i}) \quad (6)$$

where $X_{ERA5,i}$ is the daily ERA5-derived precipitation/PWV, $X_{ground,i}$ is the daily ground-based precipitation/PWV, and n is the number of samples The range of deviation is from −∞ to +∞ (the closer the deviation is to 0, the more accurate the data is). A positive bias reflects overestimation, while a negative bias indicates underestimation.

RMSE measures the overall magnitude of the error between the datasets, providing an indication of how closely ERA5 data matches the observed values by accounting for both variance and mean error. These indices are used to assess the accuracy and reliability of ERA5 data across different regions and stations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{ERA5,i} - X_{ground,i})^2} \quad (7)$$

where the range of RMSE is 0 to ∞, the smaller the value the smaller the overall deviation.

2.4 PE calculation

Precipitation efficiency (PE) is the percentage of total atmospheric water vapor that falls as measurable precipitation [25]. It is defined as the ratio of the mean daily precipitation to the mean daily precipitable water vapor (PWV) in each year, multiplied by 100 as expressed in Eq. (8).

$$PE(\%) = \frac{Mean\ daily\ precipitation}{Mean\ daily\ PWV} \times 100 \quad (8)$$

PE is a key indicator in many weather and climate studies, representing how effectively atmospheric water vapor is converted into precipitation. It plays an important role in convective system analysis and cumulus parameterization schemes, and is essential in forecasting heavy rainfall and flash flood potential [26]. Additionally, PE is significant in understanding cloud–climate feedback processes, and its variation with environmental conditions is an important aspect of climate change research. Therefore, this study focuses on determining precipitation efficiency based on ERA5 and ground-based precipitation data, contributing to the broader understanding of precipitation characteristics in different regions.

3. Results and Discussion

3.1 Ground-based TMD observation vs ERA5 precipitation

Fig. 3 shows the time series comparison between ground-based observed precipitation and ERA5 data across six representative stations for the period of 2019-2023. Spatially, precipitation of both datasets decreases from the eastern and southern coastal regions to the northwestern inland (**Table 2**). ERA5 has similar seasonal rainfall patterns at all stations but tends to underestimate high-intensity rainfall peaks observed in ground-based data. Notably, ERA5 data fails to capture daily precipitation peaks exceeding 50 mm, which are frequently recorded in the observed datasets during rainy season, particularly in stations with high elevation and complex topography (e.g. coastal region) such as CHMA (Northern region), SISK (Northeast), and CHAN (Eastern coastal area). CHMA and SISK (high-elevation), and CHAN (coastal) have frequent peaks in observed precipitation but are only partially reflected in ERA5. In contrast, stations situated in lower-elevation coast and southern regions, such as SOKA and NKSW, show strong seasonal peaks in observed rainfall that ERA5 captures with reduced intensity. The discrepancies are particularly noticeable during heavy rainfall periods.

The summary of the statistical evaluation of ERA5 precipitation data against ground-based observations in terms of mean bias and root mean square error (RMSE) is illustrated in **Table 3**. The results show that the performance of ERA5 is variable with slight underestimation at CHMA, SISK, and CHAN and overestimation for the rest of the stations. The large discrepancy is observed at CHAN with a bias of -1.998 mm and RMSE of 17.431 mm, followed by SOKA station with RMSE of 16.935 mm, though the bias is only 0.098, which is probably due to the high intensity precipitation in those coastal regions [27]. However, PJRK, which is also located near the coast but has lower annual rainfall compared to CHAN and SOKA, shows a much lower RMSE of 8.879 mm. This can be attributed to the lower rainfall intensity at PJRK, which leads to less variability in the precipitation data and consequently a lower RMSE. On the other hand, inland stations such as NKSW show slight overestimation, while CHMA and SISK exhibit underestimation with lower RMSE values below 10 mm. It means that ERA5 struggles to accurately capture high-intensity rainfall events and local precipitation variability in coastal regions and complex geographic settings. These findings suggest that ERA5 performs better in areas with lower rainfall intensity, such as PJRK and inland stations, while coastal and mountainous regions with strong convection or intense rainfall events pose more challenges for the accuracy of ERA5 reanalysis data.

Table 3 Correlation coefficient between ground-based observation and ERA5 precipitation

Station	Bias (mm)	RMSE (mm)
CHMA	-0.548	8.966
NKSW	0.559	9.421
SISK	-0.135	10.040
CHAN	-1.998	17.431
PJRK	1.325	8.879
SOKA	0.098	16.935

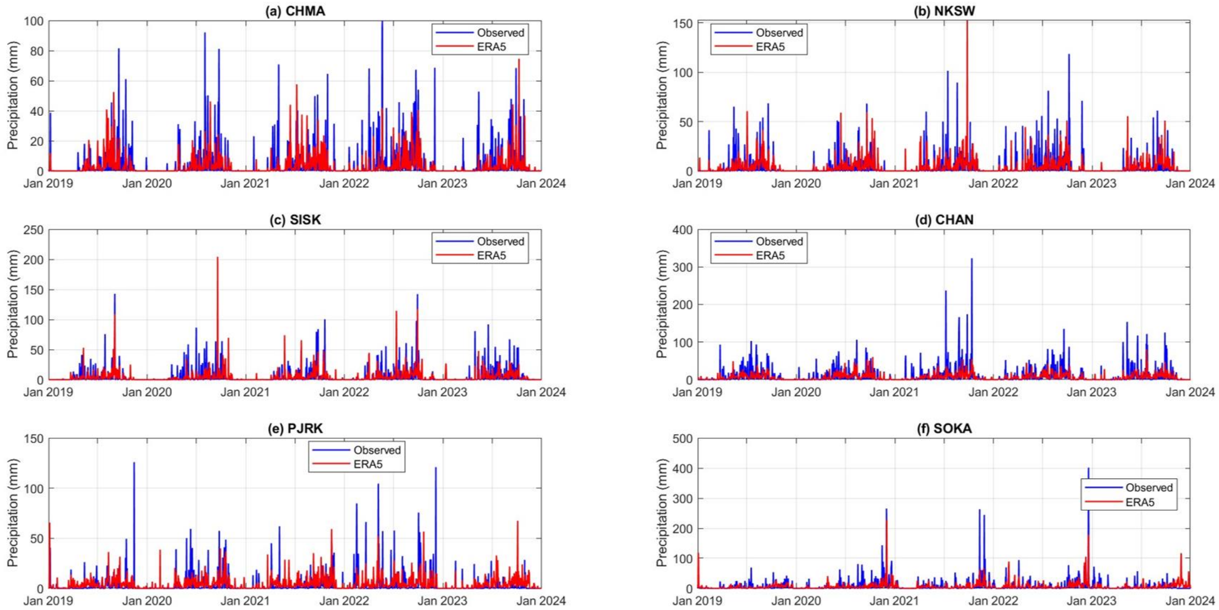

Fig. 3 Daily precipitation from ground-based TMD observation and ERA5 data (2019-2023)

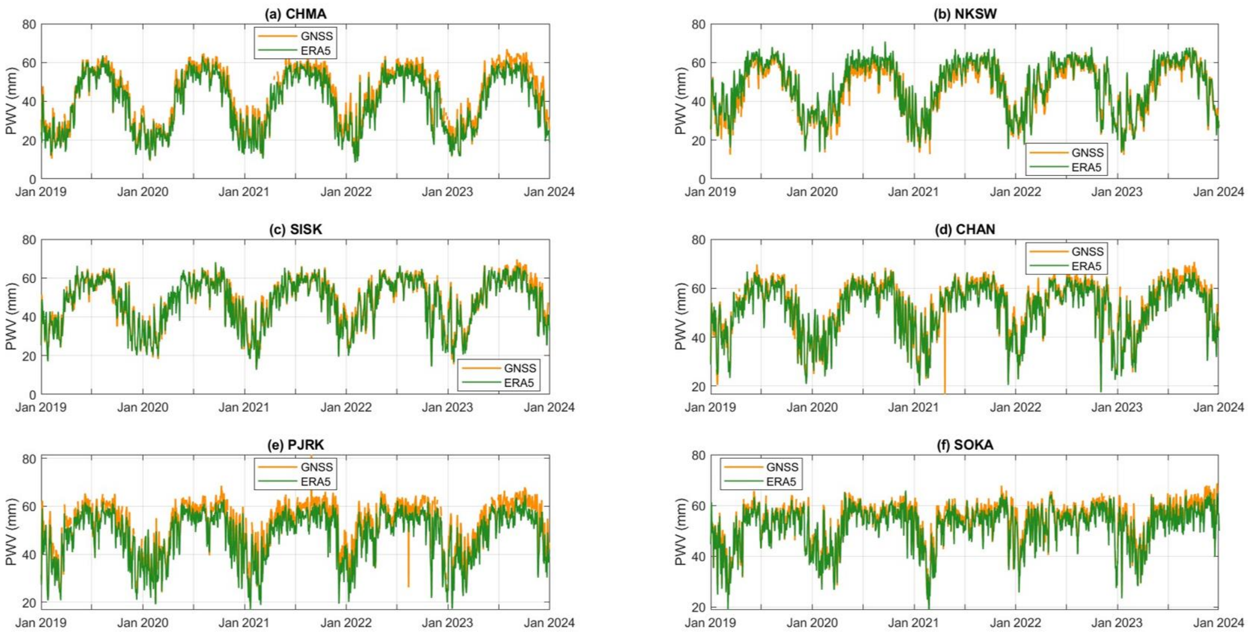

Fig. 4 Daily PWV from ground-based GNSS measurement and ERA5 data (2019-2023)

3.2 Ground-based GNSS measurement vs ERA5 PWV

Fig. 4 illustrates the time behavior (on daily basis from January 2019 to December 2023) of ground-based GNSS PWV and ERA5 PWV at six relevant stations. In general, PWV exhibits distinct seasonal variation, with minima during the winter season and maxima in the rainy season, consistently across all stations.

ERA5 PWV shows a similar seasonal pattern and temporal variation when compared with ground-based GNSS measurements. However, the differences between the two datasets become more pronounced during peak PWV periods, where ERA5 tends to diverge from observed values. Interestingly, unlike its performance in precipitation data, ERA5 PWV exhibits region-specific biases. In particular, at stations characterized by complex topography and coastal influences, such as CHMA, PJRK, and SOKA, ERA5 tends to underestimate PWV values. Conversely, in the central region (e.g., NKSW), ERA5 slightly overestimates PWV compared to GNSS observations.

In this study, GNSS-derived PWV data, obtained from ground-based Continuously Operating Reference Stations (CORS) (as shown in **Fig. 2**), are considered ground-based observations for comparison purposes.

Table 4 summarizes the statistical evaluation of ERA5 PWV accuracy against ground-based GNSS observations using bias and root mean square error (RMSE). Stations located in mountainous regions, such as CHMA in the northern region and PJRK in the western region, exhibit the largest negative biases of -3.398 mm and -3.902 mm, respectively, along with relatively high RMSE values of 4.109 mm and 4.405 mm. These results indicate that ERA5 tends to underestimate PWV in areas characterized by complex terrain and elevated atmospheric variability. In contrast, the central station NKSW shows a positive bias of 1.733 mm and an RMSE of 3.106 mm, reflecting improved agreement between ERA5 estimates and ground-based observations in lower-elevation regions with more stable atmospheric conditions.

Similarly, the northeastern station SISK records a small negative bias of -1.123 mm and the lowest RMSE value of 2.252 mm, suggesting that ERA5 effectively represents moisture distribution in regions with uniform topography. Coastal stations, including CHAN in the east and SOKA in the south, present moderate negative biases of -1.455 mm and -2.122 mm, with corresponding RMSE values of 2.590 mm and 3.185 mm, indicating consistent but slightly underestimated PWV in coastal areas influenced by maritime conditions.

The RMSE values for the PWV comparison presented in Table 4 are generally low, ranging from 2.252 to 4.405 mm indicating good agreement. This is consistent with the findings of Charoenphon (2020) [17] in which the RMSE ranges from 2.1 to 3.6. The higher discrepancies observed in this study may be attributed to differences in the study period, duration, and the number of stations used. Notably, the stations with the highest RMSE value in PJRK and CHMA stations are located in the lowest (coastal) and highest elevations (mountain), respectively (Table 2). This suggests that ERA5 faces challenges in accurately capturing precipitation patterns in regions with complex terrain or localized weather phenomena or strong convection, particularly in coastal and mountainous areas.

3.3 Ground-based and ERA5 PE

Fig. 5 and **Fig. 6** present the annual average daily PE derived from ground-based and ERA5 datasets, respectively, for six stations across Thailand from 2019 to 2023. Despite similar levels of atmospheric PWV (as shown in **Fig. 4**), PE varies across regions, reflecting differences in local atmospheric stability and convective dynamics. The spatial and temporal patterns of PE provide insight into the relative atmospheric stability and moisture conversion dynamics in these regions. High PE values observed at CHAN and SOKA, particularly in 2021 and 2022, indicate atmospheric instability and active convective processes that efficiently convert water vapor into precipitation. In contrast, consistently low PE at CHMA and NKSW reflects stable atmospheric conditions with limited convective activity and lower rainfall efficiency.

Table 4 Correlation coefficient between ground-based observation and ERA5 PWV

Station	Bias (mm)	RMSE (mm)
CHMA	-3.398	4.109
NKSW	1.733	3.106
SISK	-1.123	2.252
CHAN	-1.455	2.590
PJRK	-3.902	4.405
SOKA	-2.122	3.185

Ground-based PE values demonstrate more pronounced variability and higher peaks, particularly in coastal and southern stations such as CHAN and SOKA. The ERA5-derived PE values, while capturing overall spatial trends, are generally lower and smoother than ground-based results, reinforcing ERA5's tendency to underrepresent local convective dynamics. According to Tuller [26], high precipitation totals coupled with high PE, as seen in ground-based data for CHAN and SOKA, indicate dynamic, unstable conditions where abundant water vapor is effectively released. These coastal regions near the Gulf of Thailand are influenced by maritime moisture inflow and strong convection [28], resulting in high precipitation efficiency. Meanwhile, regions with moderate precipitation totals but low PE values, such as SISK and NKSW, reflect stable atmospheric conditions where precipitation formation is less efficient despite the presence of moisture.

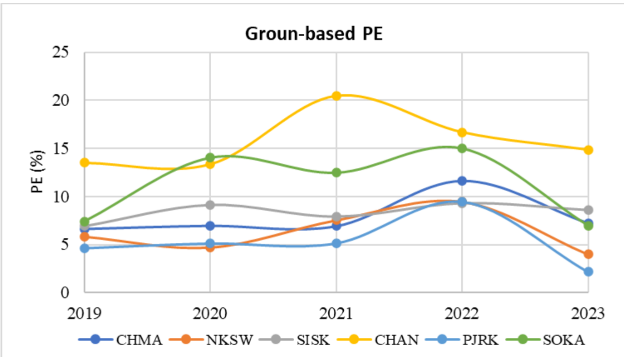

Fig. 5 Annual average daily PE in for ground-based measurements

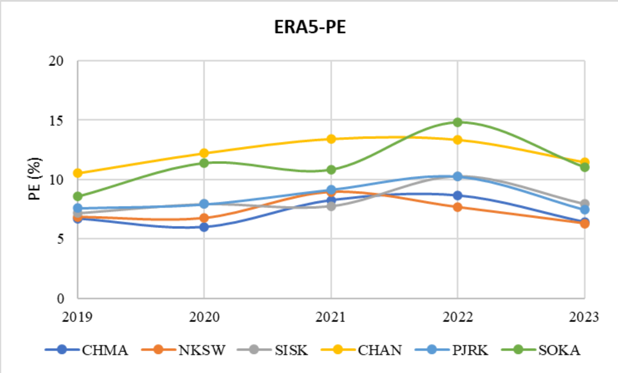

Fig. 6 Annual average daily PE for ERA5 datasets

4. Conclusions

This study assessed ERA5 daily precipitation and PWV data across different regions of Thailand from 2019 to 2023, using ground-based measurements as a reference. The results show that ERA5 effectively captures seasonal precipitation patterns but underestimates high-intensity rainfall, particularly in coastal and complex terrains like CHAN and SOKA. In contrast, inland regions like NKSW show better alignment with ground-based observations.

Additionally, the evaluation of ERA5 PWV against ground-based GNSS observations indicates that ERA5 also accurately captures the temporal patterns of PWV variation across different regions of Thailand. In mountainous regions such as the north (CHMA) and west (PJRK), ERA5 underestimates PWV, with larger negative biases and higher RMSE values. In contrast, flat and stable regions, such as the central (NKSW) and northeastern (SISK) areas, show smaller biases and lower errors, reflecting better agreement with ground-based data. Coastal regions (CHAN and SOKA) exhibit moderate underestimation with consistent seasonal patterns. However, the GNSS station-based PWV data have not been adjusted to the ERA5 grid point height, which may lead to discrepancies in this comparison. Future research should address this adjustment for more accurate comparisons and further improve flood risk assessments.

PE values from ground-based data show strong regional variation, with higher PE in coastal areas like CHAN and SOKA due to active convection, and lower PE in stable inland regions such as CHMA and NKSW. ERA5 captures these spatial trends but underestimates PE and smooths local variability.

With advancements in numerical weather models and data assimilation techniques, ERA5 reanalysis datasets are valuable for weather prediction and hydrological applications. ERA5 can fill gaps in spatially and temporally sparse ground-based measurements, making it reliable for climate monitoring and large-scale weather forecasting in regions such as Thailand's less flood-prone northeastern areas. However, its performance is less accurate in regions with complex topography and high rainfall variability, such as coastal zones and mountainous regions. Despite these limitations, ERA5's broad spatiotemporal coverage and PWV data remain valuable for flood forecasting, particularly when combined with GNSS observations in data-scarce areas for enhanced early warning systems.

It should be noted that reanalysis products, such as ERA5, inherently contain uncertainties, primarily due to model and observation errors in the data assimilation system [29]. These uncertainties can lead to biased precipitation estimates for different locations and events, particularly during extreme weather conditions. Additionally, the height differences between reanalysis grid points and observation stations are a primary cause of discrepancies in precipitable water vapor (PWV) comparisons [7]. Such differences can affect the accuracy of PWV data, which is crucial for flood and drought prediction. To minimize these errors, adjustments for height differences should be considered in future studies. Furthermore, while ERA5 offers valuable long-term data, its spatial resolution may not always capture localized variations in precipitation and moisture, limiting its application in highly detailed water resource management.

Acknowledgement

The authors would like to express gratitude to Chulalongkorn University for the support provided through the ASEAN and Non-ASEAN scholarship and for the data provided through the Thai Meteorological Department (TMD), Department of Land (DOL) and the Department of Public Works and Town & Country Planning (DPT) for providing the GNSS observation data and the International GNSS Service (IGS) for providing final orbit and clock corrections.

References

[1] A. S. Khaniani, H. Motieyan, and A. Mohammadi, "Rainfall forecast based on GPS PWV together with meteorological parameters using neural network models," *Journal of Atmospheric and Solar-Terrestrial Physics,* vol. 214, p. 105533, 2021.

[2] W. Xie *et al.*, "The evaluation of IMERG and ERA5-Land daily precipitation over China with considering the influence of gauge data bias," *Sci Rep,* vol. 12, no. 1, p. 8085, May 16 2022, doi: 10.1038/s41598-022-12307-0.

[3] Y. Wu *et al.*, "Revealing the water vapor transport during the Henan "7.20" heavy rainstorm based on ERA5 and Real-Time GNSS," *The Egyptian Journal of Remote Sensing and Space Sciences,* vol. 27, no. 2, pp. 165-177, 2024, doi: 10.1016/j.ejrs.2024.02.004.

[4] J. de Leeuw, J. Methven, and M. Blackburn, "Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations," *Quarterly Journal of the Royal Meteorological Society,* vol. 141, no. 688, pp. 798-806, 2014, doi: 10.1002/qj.2395.

[5] Y. Xin, Y. Yang, X. Chen, X. Yue, Y. Liu, and C. Yin, "Evaluation of IMERG and ERA5 precipitation products over the Mongolian Plateau," *Sci Rep,* vol. 12, no. 1, p. 21776, Dec 16 2022, doi: 10.1038/s41598-022-26047-8.

[6] V. Maggioni, C. Massari, and C. Kidd, "Errors and uncertainties associated with quasiglobal satellite precipitation products," in *Precipitation science*: Elsevier, 2022, pp. 377-390.

[7] L. Huang *et al.*, "Evaluation of Hourly PWV Products Derived From ERA5 and MERRA‐2 Over the Tibetan Plateau Using Ground‐Based GNSS Observations by Two Enhanced Models," *Earth and Space Science,* vol. 8, no. 5, 2021, doi: 10.1029/2020ea001516.

[8] J. Vaquero-Martinez *et al.*, "Validation of integrated water vapor from OMI satellite instrument against reference GPS data at the Iberian Peninsula," *Sci Total Environ,* vol. 580, pp. 857-864, Feb 15 2017, doi: 10.1016/j.scitotenv.2016.12.032.

[9] S. Wang *et al.*, "Evaluation of Precipitable Water Vapor from Five Reanalysis Products with Ground-Based GNSS Observations," *Remote Sensing,* vol. 12, no. 11, 2020, doi: 10.3390/rs12111817.

[10] C. Trakolkul, C. Charoenphon, and C. Satirapod, "Investigating the Relationship between Precipitable Water Vapor and Rainfall Data during Flood Events: A GNSS-Based Study in Thailand," *International Journal of Geoinformatics,* vol. 20, no. 11, pp. 1-10, 2024.

[11] Q. Jiang *et al.*, "Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland," *Journal of Hydrology,* vol. 595, 2021, doi: 10.1016/j.jhydrol.2020.125660.

[12] N. Sriwongsitanon, T. Suwawong, S. Thianpopirug, J. Williams, L. Jia, and W. Bastiaanssen, "Validation of seven global remotely sensed ET products across Thailand using water balance measurements and land use classifications," *Journal of Hydrology: Regional Studies,* vol. 30, 2020, doi: 10.1016/j.ejrh.2020.100709.

[13] S. A. Martin and R. J. Ritchie, "Sourcing Thai geography literature for ASEAN and international education," *Singapore Journal of Tropical Geography,* vol. 41, no. 1, pp. 61-85, 2019, doi: 10.1111/sjtg.12296.

[14] J. Khedari, A. Sangprajak, and J. Hirunlabh, "Thailand climatic zones," *Renewable Energy,* vol. 25, no. 2, pp. 267-280, 2002.

[15] V. Manisan, "Geography and climatology in every season of various parts in Thailand," *Technical document of Meteorological Department, Bangkok,* 1995.

[16] C. Trakolkul and C. Satirapod, "Variations of precipitable water vapor using GNSS CORS in Thailand," *Survey Review,* vol. 53, no. 376, pp. 90-96, 2020, doi: 10.1080/00396265.2020.1713611.

[17] C. Charoenphon and C. Satirapod, "Improving the accuracy of real-time precipitable water vapour using country-wide meteorological model with precise point positioning in Thailand," *Journal of Spatial Science,* vol. 67, no. 2, pp. 313-329, 2020, doi: 10.1080/14498596.2020.1758969.

[18] C. Charoenphon and C. Satirapod, "A Preliminary study for real-time estimation of precipitable water vapor using GNSS precise point positioning in Thailand."

[19] S. De Haan, *National/regional operational procedures of GPS water vapour networks and agreed international procedures*. WMO Geneva, Switzerland, 2006.

[20] B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle, "GNNS–Global Navigation Satellite Systems," *GPS, GLONASS,* 2008.

[21] P. Meunram and C. Satirapod, "Spatial variation of precipitable water vapor derived from GNSS CORS in Thailand," *Geodesy and Geodynamics,* vol. 10, no. 2, pp. 140-145, 2019, doi: 10.1016/j.geog.2019.01.003.

[22] D. Odijk and P. J. Teunissen, "Improving the speed of CORS Network RTK ambiguity resolution," in *IEEE/ION Position, Location and Navigation Symposium*, 2010: IEEE, pp. 79-84.

[23] J. Saastamoinen, "Atmospheric correction for the troposphere and stratosphere in radio ranging satellites," *The use of artificial satellites for geodesy,* vol. 15, pp. 247-251, 1972.

[24] B. Chen, W. Dai, Z. Liu, L. Wu, C. Kuang, and M. Ao, "Constructing a precipitable water vapor map from regional GNSS network observations without collocated meteorological data for weather forecasting," *Atmospheric Measurement Techniques,* vol. 11, no. 9, pp. 5153-5166, 2018, doi: 10.5194/amt-11-5153-2018.

[25] S. E. Tuller, "The world distribution of annual precipitation efficiency," *Journal of Geography,* vol. 70, no. 4, pp. 219-223, 1971.

[26] C.-H. Sui, X. Li, and M.-J. Yang, "On the definition of precipitation efficiency," *Journal of the Atmospheric Sciences,* vol. 64, no. 12, pp. 4506-4513, 2007.

[27] K. Torsri, Z. Lin, V. N. Dike, T. Thodsan, and P. Wongsaming, "Evaluation of Spatial-Temporal Characteristics of Rainfall Variations over Thailand Inferred from Different Gridded Datasets," *Water,* vol. 14, no. 9, 2022, doi: 10.3390/w14091359.

[28] W. Nounmusig, "Analysis of rainfall in the eastern Thailand," *GEOMATE Journal,* vol. 14, no. 46, pp. 150-155, 2018.

[29] S. Nie, Y. Luo, T. Wu, X. Shi, and Z. Wang, "A merging scheme for constructing daily precipitation analyses based on objective bias-correction and error estimation techniques," *Journal of Geophysical Research: Atmospheres,* vol. 120, no. 17, pp. 8671-8692, 2015.