High-Resolution Integrated Water Vapor Estimation Using the Gaussian Mixed Long Short-Term Memory Network: A Satellite-Based Inter-Comparison and Data-Fusion

Lingke Wang, Duo Wang, Joseph Awange *and* Hansjörg Kutterer

***Abstract*— Water vapor, the most influential greenhouse gas, is central to Earth's climate system, affecting the hydrological cycle, energy balance, and atmospheric dynamics. Integrated Water Vapor (IWV) is a key variable for understanding these processes. However, conventional IWV retrieval methods—such as ground-based sensors, satellite observations, and numerical weather models (NWM)—are often limited by spatial resolution, temporal continuity, and retrieval accuracy. To address these challenges, this study introduces a novel deep learning method GMLSTM-HIM, a High-resolution IWV estimation Model (HIM) based on a Gaussian Mixture Long Short-Term Memory (GMLSTM) framework. By integrating Global Navigation Satellite System (GNSS) and NWM inputs, including weighted mean temperature, GMLSTM-HIM utilizes a bidirectional LSTM structure and probabilistic output sequences to improve IWV estimation accuracy while quantifying uncertainty arising from spatial heterogeneity. Compared to ERA5 and VMF3, the model achieves average RMSE reductions of 68.44% and 36.15%, respectively. The model's performance is further evaluated through inter-comparisons with MODIS and Fengyun satellite-derived IWV products, highlighting both the accuracy of GMLSTM-HIM and the complementary strengths of satellite observations. The results suggest that, of the satellite datasets examined in this case study, the MODIS 5 km product exhibits the highest consistency with the GMLSTM-HIM model estimates, outperforming the higher-resolution MODIS 1 km and FY-3D 1 km products in terms of product reliability (measured by root mean square error and correlation). A data fusion strategy is also proposed, combining model and satellite estimates to preserve fine-scale details and enhance robustness. Overall, GMLSTM-HIM provides a robust framework for high-resolution IWV retrieval, with significant potential to advance atmospheric studies, climate surveillance, and operational weather forecasting within the remote sensing community.**

Index Terms— Integrated Water Vapor (IWV) estimation, Global Navigation Satellite System (GNSS), Gaussian Mixture Long Short-Term Memory (GMLSTM), Inter-comparison, Data Fusion, MODIS, Fengyun Series

I. INTRODUCTION

Water vapor is the predominant greenhouse gas in Earth's atmosphere, playing a crucial role in shaping weather and driving climate change [1], [2]. Its variations influence key processes such as the hydrological cycle, global energy distribution, and biogeochemical dynamics. Therefore, studying water vapor is essential for understanding climate systems at both regional and global levels [3]. Integrated Water Vapor (IWV), also referred to as Precipitable Water Vapor (PWV), quantifies the total atmospheric water vapor content within a vertical column extending from the Earth's surface to the tropopause. As a fundamental parameter in climate research, IWV is essential for evaluating energy budgets, investigating hydrological cycles, and enhancing numerical weather prediction models [4], [5], [6].

IWV retrieval has traditionally relied on both ground-based and satellite-based techniques. Ground-based instruments—such as radiosondes, Global Navigation Satellite System (GNSS), sun photometers, lunar photometers, star photometers, and microwave radiometers—are widely used [7], [8], [9], [10], [11]. While radiosondes provide accurate vertical profiles, their spatial and temporal limitations (typically twice-daily launches) and high operational costs hinder their suitability for continuous or large-scale monitoring [12], [13]. Consequently, they are inadequate for applications requiring high-resolution or real-time IWV data. Satellite-based observations, including those from the Moderate-Resolution Imaging Spectroradiometer (MODIS) and Fengyun (FY) missions, retrieve IWV using radiance measurements in near-infrared (NIR) and infrared (IR) bands via onboard Medium Resolution Spectral Imager (MERSI) sensors [14], [15], [16], [17], [18]. These methods have improved through advancements in spatial resolution, radiometric sensitivity, and processing capabilities [4]. However, limitations remain: data are only available at satellite overpass times, and retrieval accuracy is reduced under cloudy conditions due to NIR/IR sensitivity [6]. Additionally, global coverage is incomplete, particularly in persistently cloudy or high-latitude regions, restricting their application in high-precision or operational meteorology.

In contrast to direct measurement techniques, numerical and statistical methods—such as those based on the European Centre for Medium-Range Weather Forecasts (ECMWF) offer an alternative method for large-scale IWV estimation. The fifth-generation ECMWF reanalysis (ERA5), for instance, enables global gridded IWV retrieval by interpolating hourly pressure-level meteorological variables and applying empirical retrieval models [19], [20], [21], [22]. While such approaches offer considerable spatial coverage and continuity, they face notable limitations in accurately resolving short-term or localized variations in IWV, particularly under conditions of rapid atmospheric change. This is primarily due to the coarse spatial resolution and the inherent smoothing effects of interpolation,

This work was supported by funding of the first author through the China Scholarship Council (CSC). Duo Wang and Joseph Awange also acknowledge the support of The Hong Kong Polytechnic University Fund, Grant 283 No. P0054005. (Corresponding author: Duo Wang.)

Lingke Wang and Hansjörg Kutterer are with the Geodetic Institute (GIK), Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany. (E-Mail: lingke.wang@kit.edu; hansjoerg.kutterer@kit.edu).

Duo Wang and Joseph Awange are with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong. (E-Mail: duo.wang@polyu.edu.hk; joseph.awange@polyu.edu.hk).

which reduce the model's sensitivity to transient or extreme IWV events [23], [24]. In light of these limitations, it is necessary to develop a high-resolution IWV estimation model to substantially improve the accuracy, reliability, and applicability of IWV retrievals in large-scale atmospheric and climate research.

Accurately modeling IWV is challenging due to the complexity of water vapor dynamics and atmospheric turbulence. Deep learning, a data-driven method utilizing deep neural networks, has shown strong potential with its nonlinear fitting and feature extraction capabilities [25], [26], [27]. Several models have been developed, such as the Tropospheric Delay Network, which combines Geostationary Operational Environmental Satellite-R (GOES-R) data with Global Forecast System (GFS) simulations to predict Zenith Wet Delay (ZWD) [28]. Although it outperforms traditional approaches, its accuracy often exceeds 10 mm, and GOES-R's limited spatial coverage and sensitivity to weather reduce its generalizability. Other supervised methods, like Least Squares Support Vector Machine (LSSVM), use GNSS-derived IWV and auxiliary variables (e.g., latitude, altitude, temperature) for regional estimation [24], but their effectiveness depends on dense GNSS station coverage and lacks temporal modeling, limiting applicability in sparse or dynamic environments.

In response to the inherent limitations of both direct integrated water vapor measurements and reanalysis-based retrieval methods, as well as the challenges associated with existing deep learning approaches, this study presents a novel framework for high-resolution regional IWV estimation. The proposed model is specifically designed to capture complex spatiotemporal dependencies in atmospheric water vapor, thereby enhancing retrieval accuracy and ensuring robustness against observational uncertainties, such as those induced by cloud contamination. This study builds upon the Gaussian Mixed Long Short-Term Memory (GMLSTM) architecture introduced by Reference [29], and extends it into a comprehensive framework tailored for high-resolution IWV estimation. The extended model, referred to hereafter as GMLSTM-HIM (Gaussian Mixed Long Short-Term Memory framework for High-resolution IWV estimation Model), integrates spatial and temporal correlations to establish an adaptive mapping between numerical weather model (NWM)-interpolated ZWD and GNSS-derived ZWD. By incorporating weighted mean temperature (T_m) into the IWV retrieval process [30], the model is capable of generating accurate IWV estimates at arbitrary locations within the study domain. Unlike the original work by Reference [29], which focused primarily on Zenith Total Delay (ZTD) prediction under different weather scenarios, our study advances the model's application to the direct estimation of IWV fields. From a methodological perspective, GMLSTM-HIM uses a bidirectional LSTM (Bi-LSTM) structure to better capture the temporal dynamics of ZWD sequences [31]. In addition, a distinguishing feature of this model is its use of probability density distribution sequences—instead of deterministic numerical outputs—to characterize IWV predictions. This probabilistic formulation not only improves the model's robustness to input noise but also allows for a more nuanced representation of rapid and transient variations in atmospheric water vapor. Consequently, the model effectively captures rapid and transient variations in atmospheric water vapor, highlighting its robustness and applicability in dynamic meteorological environments. A key innovation of the present work lies in the quantitative validation of the estimated IWV fields through systematic comparisons with satellite-based products, including MODIS and the FY series—an aspect not addressed in previous studies. These inter-comparisons not only verify the performance of the GMLSTM-HIM but also serve to evaluate the reliability and consistency of satellite IWV retrievals over the German region. Furthermore, this study introduces a data fusion strategy that combines satellite-derived IWV with model-based estimates from GMLSTM-HIM. This fusion framework, proposed for the first time, addresses the spatial discontinuities and temporal sparsity commonly found in satellite observations, while simultaneously preserving the fine-scale features provided by the deep learning model. The resulting fused IWV fields exhibit enhanced spatial and temporal continuity, thereby supporting broader applications in both operational weather forecasting and atmospheric research.

Section II describes the background and proposed methodology. Section III presents the dataset employed in this study. The experimental results and validation are discussed in Section IV, followed by the concluding remarks in Section V.

II. Background and Methodology

A. Retrieval of IWV from Multiple Data Sources

This section presents a comprehensive overview of prevalent techniques for IWV retrieval, encompassing both ground-based observation systems and NWM–based approaches. Ground-based techniques include GNSS, with openly accessible data provided by the Nevada Geodetic Laboratory (NGL) (https://geodesy.unr.edu/), and radiosonde observations, available via the University of Wyoming's upper-air sounding archive (https://weather.uwyo.edu/). In parallel, NWM-based estimations—such as those derived from the ERA5 reanalysis product and the Vienna Mapping Functions 3 (VMF3)—offer modeled IWV data, with access provided by ECMWF at https://www.ecmwf.int and the VMF3 server at https://vmf.geo.tuwien.ac.at/, respectively. The relevant information is summarized in TABLE I.

TABLE I
A COMPARATIVE OVERVIEW OF DIFFERENT APPROACHES TO IWV CALCULATION

IWV Retrieval Methods	Time Resolution	Spatial Resolution
GNSS IWV	30s	Station-wise
Radiosonde IWV	Twice per day	Station-wise
ERA5 IWV	1 hour	0.25°x0.25°
VMF3 IWV	6 hours	1°x1°

GNSS satellites transmit L-band microwave signals that refract in the non-dispersive troposphere, causing an increased signal path length and resulting in slant tropospheric delay (STD) [32]. This delay is converted to ZTD via mapping functions. Since water vapor is the only atmospheric constituent with a permanent dipole moment, total refractivity (N) is

decomposed into hydrostatic (N_{Hydro}) and wet (N_{Wet}) components [33], corresponding to Zenith Hydrostatic Delay (ZHD) and ZWD, respectively. The ZWD is obtained by subtracting ZHD from ZTD, with its accuracy depending on both components. Using precise orbit and clock data from International GNSS Service (IGS), GNSS Precise Point Positioning (GNSS-PPP) approach provides high-temporal-resolution, accurate ZTD_{GNSS} estimates, widely accepted in geodetic studies. These can be computed with software like Bernese or accessed from global GNSS tropospheric products (e.g., NGL). Since GNSS alone cannot separate ZTD into wet and hydrostatic parts, the stable ZHD is typically modeled (e.g., via VMF3), yielding ZHD_{VMF3}, which allows derivation of ZWD_{GNSS}:

$$ZWD_{GNSS} = ZTD_{GNSS} - ZHD_{VMF3} \quad (1)$$

introducing the relationship between ZWD_{GNSS} and GNSS IWV (IWV_{GNSS}):

$$IWV_{GNSS} = Q \cdot ZWD_{GNSS} \quad (2)$$

$$Q = \frac{10^6}{R_v \cdot \left[k_2' + \frac{k_3}{T_m}\right]} \quad (3)$$

where Q is the conversion factor and the parameter R_v=461.522 $J \cdot kg^{-1}K^{-1}$ is the specific gas constant for water vapor [34]. The parameters k_2'=22.1 K/hPa and k_3=373900 K^2/hPa are atmospheric refractivity constants [17]. In (3), T_m is the weighted mean temperature and represents the pressure-weighted average atmospheric temperature that links ZWD to IWV. Since IWV is inversely proportional to T_m, using ERA5-derived T_m improves the accuracy of IWV retrieval by better capturing its spatial and temporal variability. Then, the methodology of obtaining T_m is outlined as follows [35]:

$$T_m = \frac{\sum_{H_s}^{H_{top}} \frac{e_j}{T_j} \Delta H_j}{\sum_{H_s}^{H_{top}} \frac{e_j}{T_j^2} \Delta H_j} \quad (4)$$

here, e and T represent the water vapor pressure (in hPa) and temperature (in K), respectively, extending vertically from the altitude of the GNSS station (H_s) to the upper boundary of the troposphere (H_{top}). The T_m (in K) can subsequently be derived using the ERA5 pressure-level dataset, which provides high-resolution data on an hourly basis with a spatial grid of 0.25° × 0.25° and covers 37 vertical pressure levels.

Radiosonde-derived IWV (IWV_{RS}), as a reliable geodetic reference, is widely utilized to validate IWV estimates from other methods. Radiosonde observations provide vertical profiles of atmospheric parameters, typically including temperature (in °C), dew point temperature (in °C), pressure (in Pa), geopotential height (in m), wind direction (in degrees from north), and wind speed (in m/s). To compute IWV_{RS}, the integral method is applied over the vertical extent of the observational domain. In this research, quality control is performed on the radiosonde profiles by removing incomplete or missing values. If the cleaned profiles contain fewer than five valid layers, the corresponding data for that epoch are excluded from further analysis. The basic calculation principle of IWV_{RS} is outlined by Reference [36]:

$$IWV_{RS} = \int_{H_{low}}^{H_{top}} \rho_w dh = \sum_{i=1}^{n} \rho_{d,i} r_i \Delta h_i \quad (5)$$

here, ρ_w represents the water vapor density, H_{low} denotes the altitude of the radiosonde station, and H_{top} corresponds to the altitude of the highest point reached by the radiosonde. The variable n indicates the number of atmospheric layers in the profile. Additionally, $\rho_{d,i}$, r_i and Δh_i denote the dry air density, the water vapor mixing ratio, and the altitude difference for each layer, respectively. The detailed calculation methodology can be found in Reference [36].

NWM like ERA5 and VMF3 provide large-scale IWV estimates with lower temporal resolution—1 hour for ERA5 and 6 hours for VMF3. ERA5, known for assimilating global observations using physical laws, offers high-quality hourly reconstructions of atmospheric variables on a 0.25° global grid, covering from the surface to the tropopause in 37 pressure-based layers. By assuming uniformity within each atmospheric layer as represented in ERA5, the ZWD (ZWD_{ERA5}) and IWV (IWV_{ERA5}) can be integrated as follows:

$$ZWD_{ERA5} = 10^{-6} \sum_{i=i_0}^{36} \frac{1}{2}\left[k_2 \left(\frac{e_i}{T_i} + \frac{e_{i+1}}{T_{i+1}}\right) + k_3 \left(\frac{e_i}{T_i^2} + \frac{e_{i+1}}{T_{i+1}^2}\right)\right] d_{i,i+1} \quad (6)$$

$$IWV_{ERA5} = \int \frac{q}{g} dP \approx \sum_{i=i_0}^{36} \frac{q_i}{g_i} \Delta P_i \quad (7)$$

where i_0 is the initial layer which corresponds to the height of GNSS station. T_i are the pressure and temperature at i-th layer pressure level obtained from ERA5 directly. e_i is the water vapor pressure at i-th layer pressure level, q_i and g_i represent the mean specific humidity ($kg \cdot kg^{-1}$) and the mean local gravitational acceleration ($m \cdot s^{-2}$) at the i-th atmospheric layer, respectively. ΔP_i denotes the pressure difference (Pa) between the (i+1)-th and the i-th layers. The parameters k_2=77.4 K/hPa is an atmospheric refractivity constant. To ensure consistency between ERA5 and GNSS, height alignment is initially performed. In this study, ZWD_{ERA5} and IWV_{ERA5} are computed relative to the height of the GNSS station. If the GNSS station's elevation is lower than the surface layer height of ERA5, the surface height is used as the initial reference, and the i_0 in (6) and (7) is determined accordingly. Conversely, if the GNSS station's height exceeds the ERA5 surface height, the GNSS station's height is selected as the initial reference. Due to space constraints, the detailed procedures for calculating IWV_{ERA5} are outlined in Reference [37] with the corresponding interpolation methods based on those described by Reference [38].

Another widely used method for obtaining ZWD and IWV is through the VMF3 surface products. VMF3 is a prominent empirical tropospheric model provided by TU Vienna. It offers tropospheric delay and corresponding mapping function data for the Earth's surface grid, with a spatial resolution of 1° and a temporal resolution of 6 hours [39]. As described by Reference [40], estimating the ZWD at a target location using VMF3 involves an initial linear temporal interpolation across epochs for the four surrounding grid points. This step in this research adjusts the temporal resolution to 1 hour, ensuring consistency with the ERA5 products. Since VMF3 provides surface products, the elevation of the target location may differ from that of the surrounding grid points. Therefore, a vertical adjustment must be applied after time alignment. Finally, the ZWD_{VMF3} at the target location can be derived using the procedures and equations outlined in Reference [41], applying bilinear interpolation for spatial adjustments. Subsequently, the

IWV estimated by VMF3 (IWV_{VMF3}) can be determined using (2) and (3).

B. Satellite Instruments and their IWV Products

MODIS, onboard NASA's Terra and Aqua satellites [42], provides near-global coverage every 1–2 days via a 2330 km swath. Terra operates in a morning orbit (north to south), while Aqua follows an afternoon orbit (south to north). MODIS acquires data across 36 spectral bands, with specific bands (890–920 nm, 931–941 nm, 915–965 nm) dedicated to water vapor retrieval. Though native spatial resolution is 1 km, Level 2 IWV products aggregate to 5 km × 5 km using onboard MERSI, which employs multispectral NIR and IR channels and radiative transfer algorithms for both day and night retrievals [5]. Daytime IWV is derived using NIR bands (channels 2, 5, 17–19), based on two-/three-channel ratios and radiative transfer look-up tables, accounting for solar and viewing geometry. In cloudy conditions, 0.8–2.5 μm channels are applied to retrieve water vapor above/within clouds [43]. At night, thermal IR bands are used, employing statistical retrievals refined by physical models, correlating MODIS radiances with radiosonde profiles to estimate IWV [44]. Both 1 km and 5 km IWV products: MOD05_L2 (Terra) and MYD05_L2 (Aqua), are used in this study, with details summarized in TABLE II.

TABLE II
INSTRUMENT INFORMATION FOR MODIS AND FY-3D

Feature	MODIS	FY -3D
Focus	Integrated water vapor	Global atmospheric monitoring; Numerical weather prediction; Environmental observation
Bands Used	NIR (daytime) and IR (nighttime)	Visible and NIR channels for MERSI;
Output Type	Single value (total water vapor)	Single value (total water vapor)
Resolution	5 km × 5 km spatial 1 km × 1 km spatial	1 km × 1 km spatial
Cloud Handling	Clear-sky conditions	Clear-sky conditions
Use Cases	Water vapor studies; Atmospheric correction	Monitoring atmospheric water vapor; Land surface parameter retrieval
Data Source	https://earthdata.nasa.gov/	https://satellite.nsmc.org.cn

The Chinese meteorological satellite system, initiated in 1988 with the launch of the first Fengyun satellite, comprises both sun-synchronous and geostationary satellites (National Satellite Meteorological Center). The FY-3 series represents the second generation of China's polar-orbiting meteorological satellites, offering global observations of atmospheric temperature, humidity profiles, and related variables, supporting applications in climate research and numerical weather prediction [45], [46]. Multiple FY-3 satellites have since been launched, each equipped with advanced sensors to enhance atmospheric monitoring. The FY-4 series, consisting of geostationary satellites, is not applicable to our case and thus excluded from this study. Among the FY-3 satellites, FY-3B ceased operations after 2021, and FY-3C mainly covers low-latitude regions, limiting its relevance for German applications. FY-3D, by contrast, provides global coverage with a 12-hour revisit cycle and includes the MERSI sensor for IWV retrieval, particularly over East Asia and adjacent regions [47], [48]. FY-3F, launched in 2023, offers improved meteorological monitoring, but its IWV products became publicly available only in 2025, beyond the scope of this study. Therefore, this study employs FY-3D data (TABLE II) to assess the performance of Fengyun satellite products and to compare them with GMLSTM-HIM results.

C. GMLSTM-HIM Based IWV Retrieval

In general, the accuracy of IWV estimation is largely dependent on the precision of ZWD. While the ZHD provided by VMF3 is highly reliable, significant discrepancies exist between the ZTD derived from VMF3 or ERA5 and the ZTD observed from GNSS. This indicates that inaccuracies in ZWD are the primary contributors to these differences, with ZWD_{VMF3} or ZWD_{ERA5} failing to achieve the necessary precision [29]. To address this limitation and achieve accurate IWV estimation at any location, a deep learning approach GMLSTM-HIM using the GMLSTM framework is employed. This method is applied within a 200 km radius around the study area, where the presence of GNSS stations is essential. The ZWD_{GNSS} at each station can be directly obtained, while ZWD_{VMF3} and ZWD_{ERA5} are estimated at the same GNSS station locations. Subsequently, the method aims to learn the mapping relationship $\{ZWD_{VMF3}, ZWD_{ERA5}\} \rightarrow \{ZWD_{GNSS}\}$, operating under the assumption that this mapping pattern is consistent across a large-scale area. In this model, ZWD_{VMF3} is incorporated to enhance the accuracy and spatiotemporal adaptability of the input, particularly in regions where ZWD_{ERA5} may be less reliable. It also enables effective cross-validation and integration of multiple data sources. By training the GMLSTM-HIM with supervised data, specifically, ZWD_{GNSS} observations paired with corresponding ZWD_{VMF3} and ZWD_{ERA5} values at the same locations. The learned model can generalize this mapping pattern to the entire study region. Once trained, the GMLSTM-HIM enables accurate IWV estimation in high resolution across the region by incorporating temperature parameters into the estimation process. The overall architecture of the GMLSTM-HIM is illustrated in Fig.1 and further details of the GMLSTM framework can be found in Reference [29].

This study begins with the preparation of supervised data, which consists of ZWD_{GNSS} derived from GNSS stations within the study area. Quality control is applied prior to training, retaining only hourly GNSS data with daily integrity >75%, with spline interpolation for minor gaps. After completing the cleaning process and verifying data integrity, the target variable $Y = \{ZWD_{GNSS}\}$ calculated using (1). Following data cleaning, the dataset Y is aggregated into a 24×14 matrix, with the 14 columns representing the total number of valid observation days across the selected GNSS stations. According to the temporal epochs and spatial locations of the selected GNSS stations, the corresponding input features $X = \{ZWD_{VMF3},$

$ZWD_{ERA5}\}$ are generated following the procedures detailed in Section II, in order to ensure consistency and alignment between the input data and the supervised target for effective model training. As illustrated in Fig. 1, the proposed GMLSTM-HIM is developed based on the GMLSTM framework, which integrates Gaussian Mixture Models (GMM) with Long Short-Term Memory (LSTM) networks, integrating GMM with three stacked Bi-LSTM layers to capture bidirectional temporal dependencies and mitigate atmospheric turbulence [31], [49].

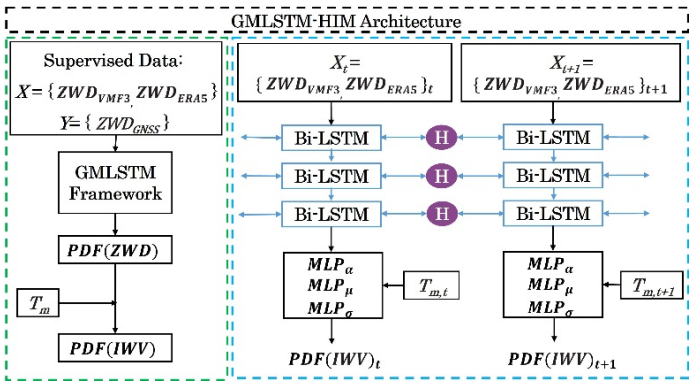

Fig. 1. Architecture of GMLSTM-HIM. H is a hidden state that contains historical information, and PDF denotes the probability density function of the Gaussian Mixture Distribution inferred by GMLSTM at time t. In addition, α, μ and σ are three fully connected perceptrons.

In our implementation, model training parameters are listed in TABLE III. To enhance generalization and reduce overfitting, dropout is applied to the hidden states of all Bi-LSTM layers except the final one in the feature extractor. Each input sequence X is thus encoded into a 4096-dimensional feature vector, which is then projected by three fully connected perceptrons to generate the parameters of a five-component GMM: mixing coefficients (α), means (μ), and standard deviations (σ). The Adam optimizer is used with an initial learning rate of 0.001 over 10,000 iterations. To ensure convergence, the learning rate decays exponentially by a factor of 0.99 every 500 iterations. The SoftMax function is applied to normalize α, ensuring they sum to one and represent a valid probability distribution.

$$\sum_{i=1}^{k} \alpha_i = 1 \tag{8}$$

$$\alpha_i = SoftMax(h_{\alpha_i}) = \frac{e^{h_{\alpha_i}}}{\sum_k e^{h_{\alpha_k}}} \tag{9}$$

where h_{α_i} represent the output from the mixing coefficients perceptron. Furthermore, the output from the standard deviation perceptron h_{σ_i} must remain strictly positive to prevent pathological configurations in which multiple standard deviation components approach zero. Such scenarios amplify errors propagated from the standard deviation term, resulting in exponential growth in gradients. This gradient explosion can ultimately destabilize the training process and lead to its collapse. To overcome this issue, we propose utilizing a modified activation function, the Piecewise Exponential Linear Unit (PELU). Its mathematical formulation is provided in (10). This approach ensures that h_{σ_i} remains strictly positive and avoids scenarios that can destabilize training, enhancing both convergence and model performance.

$$\sigma_i^2 = PELU(h_{\sigma_i}) = \begin{cases} h_{\sigma_i}, h_{\sigma_i} > 1 \\ e^{h_{\sigma_i}-1} - 1 + 10^{-8}, h_{\sigma_i} \le 1 \end{cases} \tag{10}$$

TABLE III
SUMMARY OF TRAINING PARAMETERS USED FOR THE GMLSTM-HIM

Model Parameter	Value
Hidden layer dimension of each LSTM cell	2048
Dropout rate	0.2
Number of fully connected perceptrons	3
Number of Gaussian mixture components	5

In this model, IWV is represented as a conditional probability distribution function rather than a single numerical value. To optimize the model, the Maximum Likelihood Estimation (MLE) approach is utilized as the loss function, enabling the determination of the optimal parameters. Letting θ represent the parameters of this model, the probability density function of the GMM predicted by GMLSTM-HIM at time t is expressed as:

$$PDF_t(X_t|\,\theta) = \sum_{i=1}^{K} \alpha_i(X_t|\,\theta) \frac{1}{\sqrt{2\pi}\sigma_i(X_t|\,\theta)} e^{\left(-\frac{(X_t-\mu_i(X_t|\,\theta))^2}{2\sigma_i^2(X_t|\,\theta)}\right)} \tag{11}$$

furthermore, to enhance the model's precision and minimize uncertainty, we aim for each Gaussian component to effectively characterize a more localized region within the IWV space. To achieve this, a penalty term is incorporated into the loss function to constrain each Gaussian component to be as narrow as possible, the detailed formulation of the penalty term is provided in Reference [29] After completing the training process, the model, referred to as $IWV_{GMLSTM\text{-}HIM}$, is capable of estimating IWV at any location within the study area. To perform an estimation for a specific point, the input feature vector $X = \{ZWD_{VMF3}, ZWD_{ERA5}\}$ corresponding to the desired location, is provided to the model. The model then generates the inferred $IWV_{GMLSTM\text{-}HIM}$ for the given position. At the same time, the uncertainty associated with the inference can be assessed by the standard deviation of the mixture model, as illustrated in:

$$s = Q\sqrt{\sum_{i=1}^{K} \alpha_i[\sigma_i^2 + (\mu_i - ZWD_{GMLSTM-HIM})^2]} \tag{12}$$

Finally, the overall flowchart of GMLSTM-HIM proposed in this paper and inter-comparison as well as data fusion based on satellite products is shown in Fig. 2.

D. Validation Methods

The integration of datasets with differing temporal and spatial resolutions poses a significant challenge in atmospheric studies. For example, VMF3 data are provided at 6-hour intervals with a spatial resolution of 1°, while ERA5 offers data at an hourly

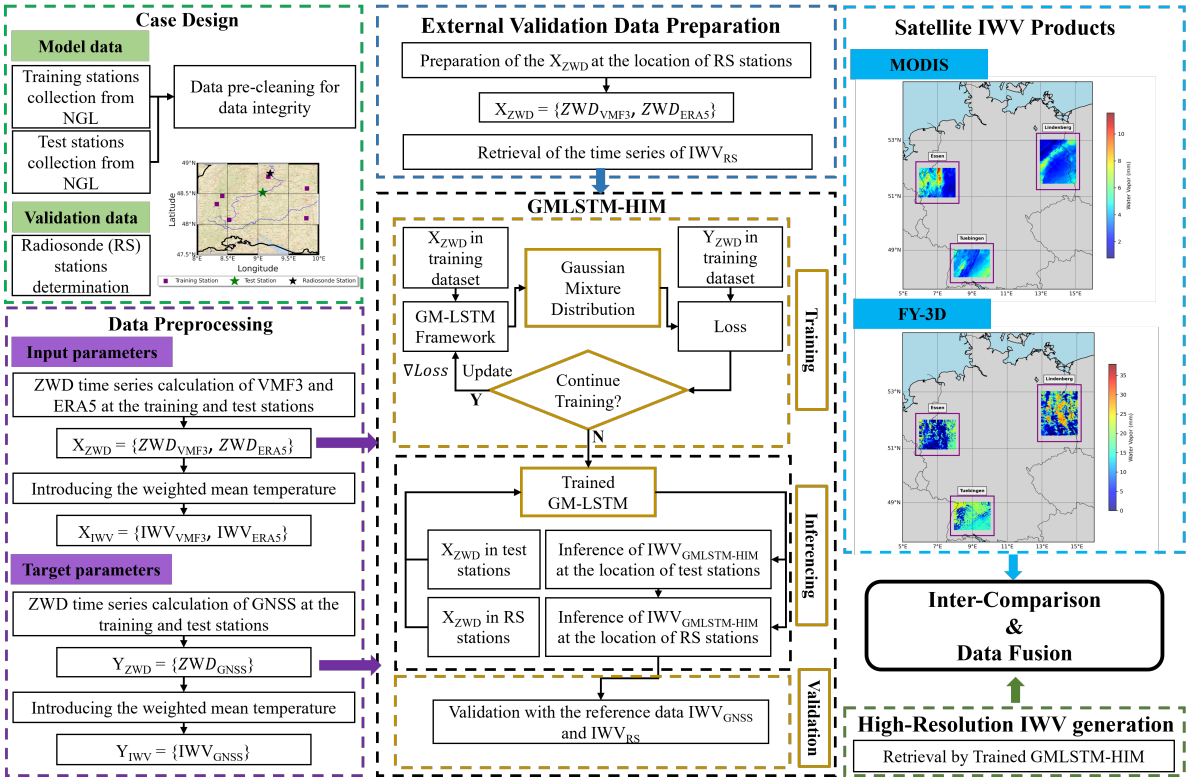

Fig. 2. Overall flowchart of the IWV retrieval method based on GMLSTM-HIM.

temporal resolution and a finer spatial resolution of 0.25°. Reference datasets, such as GNSS and radiosonde station observations, are inherently discrete and irregularly distributed. To reconcile these discrepancies, a temporal alignment was achieved through linear interpolation, while spatial interpolation was performed using the inverse distance weighting (IDW) method. Following the harmonization of temporal and spatial scales, the IWV estimates generated by the GMLSTM-HIM were systematically compared to reference IWV measurements. The accuracy assessment was conducted using key statistical metrics, including the correlation coefficient (R), root mean square error (RMSE), and mean bias (MB). Additionally, time series analyses were employed to evaluate the temporal performance of the model, offering a rigorous framework for assessing the capability of GMLSTM-HIM in accurately capturing IWV variability.

In detail, the R, computed using (13), represents the strength and direction of the linear relationship between the paired datasets, where n denotes the number of data pairs, $IWV_{m,i}$ is the IWV estimates by GMLSTM-HIM and $IWV_{r,i}$ is the reference IWV data. $\overline{IWV_m}$ and $\overline{IWV_r}$ indicate the average values. The RMSE, defined in (14), quantifies the magnitude of differences between the datasets. The MB, as given in (15), indicates the tendency of overestimation or underestimation in the dataset.

$$\mathrm{R} = \frac{\sum_{i=1}^{n}\left(IWV_{m,i} - \overline{IWV_m}\right)\left(IWV_{r,i} - \overline{IWV_r}\right)}{\sqrt{\sum_{i=1}^{n}\left(IWV_{m,i} - \overline{IWV_m}\right)^2 \sum_{i=1}^{n}\left(IWV_{r,i} - \overline{IWV_r}\right)^2}} \quad (13)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(IWV_{m,i} - IWV_{r,i}\right)^2}{n}} \quad (14)$$

$$\mathrm{MB} = \frac{1}{n}\sum_{i=1}^{n}\left(IWV_{m,i} - IWV_{r,i}\right) \quad (15)$$

E. High Spatiotemporal Resolution Regional Water Vapor Retrieval

Once trained and validated, the GMLSTM-HIM is employed to produce high-resolution regional IWV estimates across the study domain. The inference process involves feeding the model with a spatial grid of input points characterized by their latitude, longitude, and elevation, with elevation data derived from a Digital Elevation Model (DEM). This setup enables the model to generate IWV fields at arbitrary spatial resolutions, thereby meeting the requirements of both scientific research and operational applications.

To improve computational efficiency and ensure scalability for large-scale deployments, the inference workflow is parallelized using multi-processor computing techniques. This parallelization significantly accelerates the generation of IWV fields over extensive spatial domains, making the approach suitable for high-resolution and near-real-time applications.

Moreover, in addition to IWV estimation, the model concurrently provides uncertainty quantification at each grid point, expressed as predicted error. This uncertainty information offers valuable insights into the confidence and reliability of the generated IWV fields. By jointly considering both IWV estimates and their associated uncertainties, the proposed framework delivers a more comprehensive and robust characterization of atmospheric water vapor distributions, thereby enhancing the potential for downstream applications in climate and weather analysis.

F. IWV Inter-Comparison and Data-Fusion with Satellite Products

The validated model, GMLSTM-HIM, known for generating highly accurate and reliable IWV products, was used to inter-compare with MODIS and FY-3D satellite-derived IWV products. By aligning the $IWV_{GMLSTM\text{-}HIM}$ with the spatial resolution of the satellite data, a robust comparison was conducted using the relative difference as the evaluation metric. At each epoch, a relative difference is calculated for every pixel, defined as:

$$\delta_{i,j} = \frac{IWV_{i,j}^{SAT} - IWV_{i,j}^{GMLSTM-HIM}}{IWV_{i,j}^{GMLSTM-HIM}} \tag{16}$$

where the indices i and j denoted the location of the pixels and IWV is the IWV estimates by the satellite (SAT) and the GMLSTM-HIM. For each epoch, multiple pixels exhibit relative differences across the regions. To evaluate these differences, the pseudomedian and the interquartile range (IQR) were employed. The pseudomedian is calculated using the Wilcoxon signed-rank test with continuity correction [50]. It is defined as the median of all midpoints of observation pairs and coincides with the median when the dataset is symmetric. The pseudomedian of the relative differences reflects the accuracy of the satellite data, while the IQR provides insight into their precision. By conducting a direct intercomparison, this study investigates the differences between datasets and evaluates the suitability and limitations of satellite products under varying atmospheric and surface conditions. This comparative analysis provides valuable insights into the performance of satellite-based IWV retrievals and informs their appropriate application in future meteorological and climate-related studies.

To further enhance IWV coverage and quality, a data fusion strategy was implemented by integrating FY-3D or MODIS satellite-derived IWV with GMLSTM-HIM model estimates. First, a linear least squares regression was performed between the IWV values retrieved from the satellite and those generated by GMLSTM-HIM at corresponding locations. To reduce the impact of outliers in the satellite data, a 3-sigma filtering method was applied prior to regression, excluding observations that deviated by more than three standard deviations from the GMLSTM-HIM estimates. Using the resulting regression model, satellite IWV values were corrected and subsequently resampled onto a standardized 10-meter grid using nearest-neighbor interpolation. This resampling was achieved through a k-d tree search algorithm, where the search radius was defined as the average nearest-neighbor distance among observation points; In cases where no valid satellite observations were found within the radius, GMLSTM-HIM IWV estimates were used to fill the gaps. The final fused IWV product was generated by averaging the GMLSTM-HIM IWV estimates with the corrected and resampled satellite IWV values. This fusion approach combines the spatial coverage of satellite data with the accuracy of model outputs, resulting in a more robust and comprehensive IWV dataset suitable for further meteorological and climatological applications.

III. Study Area and Datasets

Germany, situated at the geographical center of Europe, offers a complex and heterogeneous atmospheric environment that is particularly well-suited for evaluating the performance and generalizability of high-resolution IWV estimation models. The country spans a broad latitudinal range (approximately 47°N to 55°N), encompassing diverse climatic zones—from temperate maritime conditions in the north to more continental regimes in the south and east. This variation is accompanied by significant topographic diversity, ranging from the high-elevation Alpine region in the south to the low-lying North German Plain. The considerable elevation differences across the country give rise to pronounced orographic effects, which strongly influence the spatial distribution of atmospheric moisture and regional precipitation dynamics. According to the German Weather Service (DWD) and prior studies [51], seasonal and spatial IWV variability is equally prominent: during summer months, IWV values can exceed 30 mm in low-elevation areas, while in winter they frequently fall below 10 mm in high-altitude zones. Additionally, Germany experiences a wide range of seasonal temperature gradients and convective activity, which further modulate atmospheric stratification and water vapor distribution. These multidimensional characteristics—spanning variations in latitude, elevation, temperature, and precipitation—render Germany an ideal natural case for systematically testing the accuracy, robustness, and transferability of the proposed GMLSTM-HIM model under complex and variable meteorological conditions.

To systematically evaluate model performance under varying geographic and climatic conditions, four representative regions across Germany were selected: Essen, Lindenberg, Meiningen, and Tuebingen. These regions were chosen based on two primary criteria: (1) the availability of a dense network of GNSS stations, which ensures sufficient spatial information for model training; and (2) the co-location or proximity of radiosonde stations, which serve as independent references for external validation. In each region, ZWD observations were sourced from the NGL dataset. A leave-one-out validation strategy was employed, wherein one GNSS station—typically centrally located—was reserved as the test site, and all other stations were used for model training. The test stations were completely excluded from the training set to guarantee unbiased performance evaluation. To incorporate seasonal variability in IWV estimation, two distinct 360-hour periods were analyzed for each region: a winter period (November 26 to December 11, 2021) and a summer period (June 25 to July 10, 2022). For the Meiningen region, due to incomplete or poor-quality GNSS data during the default periods, alternative intervals (November 18 to December 3, 2021, for winter; and June 9 to June 24, 2022, for summer) were adopted to ensure data integrity. The selected timeframes not only reflect seasonal contrasts in IWV magnitude but also provide a diverse set of atmospheric conditions for robust model evaluation. Detailed metadata for the test stations are provided in TABLE IV, and their spatial locations are illustrated in Fig. 3.

The external validation data, comprising radiosonde measurements, were obtained from the sounding database maintained by the Department of Atmospheric Science at the University of Wyoming. Detailed information about the radiosonde stations used in this study is provided in TABLE V. For the generation of the regional IWV, DEM data was utilized to provide surface height information as input for the inference

process. In this study, the DEM data was obtained from the U.S. Geological Survey (USGS) and is accessible at https://www.usgs.gov/.

To conduct a comprehensive analysis of IWV across specific spatial and temporal intervals, we obtained IWV data products from the MODIS and FY-3D satellite platforms. The spatial and temporal ranges were carefully aligned with the experiment's requirements, ensuring consistency across datasets for meaningful comparisons. All data were processed in unit with mm for uniformity during subsequent analyses. To ensure the reliability of the comparisons, we conducted a thorough quality assessment of each dataset. For MODIS data, the quality flags provided within the product were evaluated, excluding data points marked as low-quality or affected by cloud contamination. For FY-3D, we examined the consistency of the data with ground-based GNSS IWV measurements and removed outliers based on a predefined threshold (e.g., deviations beyond 3 standard deviations). Due to the limited winter availability of FY-3D observations, regions without valid data (e.g., Lindenberg and Essen) were excluded from the seasonal analysis to avoid interpolation-induced uncertainties.

TABLE IV
THE INFORMATION OF GNSS TEST STATIONS

Region	Station ID	(Latitude, Longitude) (in °)	Altitude (in m)
Lindenberg	D019	(52.2635°, 13.9425°)	92.3 m
Essen	D582	(51.4585°, 7.0161°)	217.3 m
Meiningen	D198	(50.5932°, 10.4166°)	408.7 m
Tuebingen	D400	(48.5195°, 9.0784°)	382.2 m

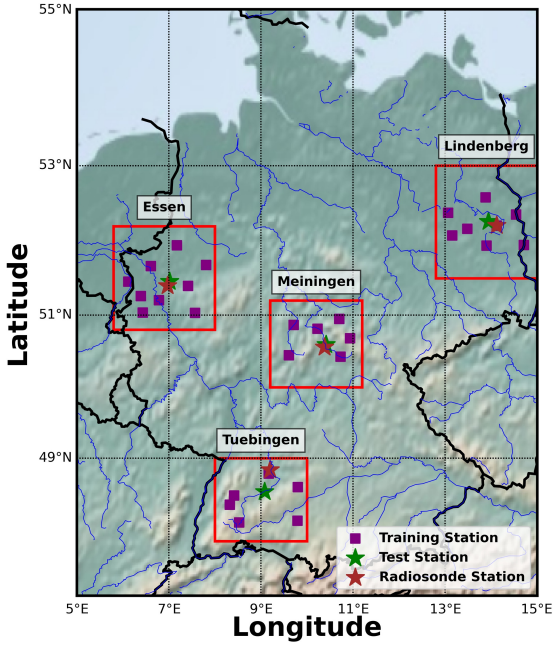

Fig. 3. The overview of selected GNSS stations from NGL.

TABLE V
THE INFORMATION OF RADIOSONDE STATIONS

Region	Station ID	(Latitude, Longitude) (in °)	Altitude (in m)
Lindenberg	RSLI	(52.2100°, 14.1200°)	112.0 m
Essen	RSES	(51.4000°, 6.9700°)	147.0 m
Meiningen	RSME	(50.5600°, 10.3800°)	450.0 m
Tuebingen	RSTU	(48.8331°,9.2000°)	314.0 m

IV. RESULTS AND ANALYSIS

A. Model performance and validation

While the methodology section has highlighted the significance of GNSS station availability, it is important to further examine how GNSS station density impacts model training and overall performance. In regions where GNSS stations are sparsely distributed, the model may face challenges in capturing fine-scale spatial variability in water vapor, due to limited representation of heterogeneous atmospheric conditions. Given the rapid temporal fluctuations and highly localized nature of IWV dynamics, insufficient station density may hinder the model's ability to generalize, particularly in areas with complex terrain or pronounced mesoscale variability. To mitigate this, we recommend expanding the spatial extent of the training region to include more GNSS stations, especially in data-sparse regions. Prior studies have demonstrated that when GNSS stations are available within 50 km of the target region, the proposed GMLSTM-HIM framework can achieve good performance improvements—up to approximately 80% over conventional tropospheric models. As the distance between the training stations and the target area increases from 50 km to 150 km, this performance advantage gradually diminishes, reflecting a decline in spatial representativeness [29]. Nevertheless, our experimental results demonstrate that the GMLSTM-HIM framework maintains satisfactory predictive accuracy when GNSS stations are located within a 200 km radius of the study area. This effective spatial range is consistent with findings from our previous work [29], in which a detailed quantitative analysis (see Section V-B, Table VI, and Figures 5–6) confirmed that the model performance remains stable within approximately 200 km. This indicates the model's robustness under moderate GNSS data sparsity, as long as a sufficient number of surrounding stations are available to capture regional-scale atmospheric variability. While GNSS station density varies globally, the network provided by NGL—often supplemented by national or regional permanent GNSS infrastructures—generally ensures sufficient coverage within the 200 km spatial buffer required by the proposed framework. These characteristics make the GMLSTM-HIM model operationally feasible across most regions worldwide. However, certain areas—such as northern Canada and Russia, parts of Central Asia and western China, as well as central and northern Africa—currently suffer from limited or absent GNSS station deployments. In these regions, the scarcity of ground-based observations poses challenges for model training and implementation, potentially constraining the accuracy and generalizability of IWV retrievals. However, a clear limitation of the current framework is its reliance on land-based GNSS infrastructure, which restricts its applicability to terrestrial regions. As such, the model is currently not suitable for

estimating IWV over oceanic areas, where GNSS station coverage is inherently absent. The following subsections evaluate model performance under these practical constraints, using case studies conducted in Germany.

The GMLSTM-HIM framework is computationally efficient and can be implemented on both CPU- and GPU-based systems. While model training and inference are feasible on conventional CPUs, the use of a GPU platform with at least 8 GB of dedicated memory is recommend to ensure significantly faster computation. In our experiments, training and inference were conducted on a deep learning workstation equipped with a single NVIDIA RTX 3090 GPU and an Intel(R) Core i7-10700 CPU (8 threads). For each study region, based on 14 days of data, model training required approximately 24 hours. Once trained, the model achieved rapid inference performance, with prediction times of less than one minute per region.

After training the GMLSTM-HIM, $IWV_{GMLSTM\text{-}HIM}$ at the test stations were inferred and validated against GNSS-derived IWV (IWV_{GNSS}). The results are presented in TABLE VI and Fig. 4. The mean RMSE and MB across both periods (winter and summer) are 0.83 mm and -0.3 mm for GMLSTM-HIM, compared to 2.63 mm and 2.08 mm for ERA5, and 1.30 mm and 0.36 mm for VMF3. These findings highlight that the GMLSTM-HIM achieves an improvement rate in RMSE of 68.44% relative to ERA5 and 36.15% relative to VMF3, demonstrating its capability to generate reliable IWV products. A comparison of the average RMSE during the winter and summer periods across the four cases reveals that the RMSE in winter is lower due to the reduced variability of IWV during this season. Although some MB values in Meiningen, are not improved compared to VMF3 or ERA5, the overall reduction in RMSE affirms the model's strong performance and indicates that the model's performance in capturing IWV estimates has been enhanced. According to the mean values in TABLE VI, the MB for winter cases is not improved compared to VMF3. This can be attributed to the fact that the VMF3 model, based on atmospheric physics, is designed to estimate IWV. This residual bias is likely related to local meteorological and topographic influences, as the surrounding elevated terrain may cause small-scale humidity variations that are not fully captured by the model training data. Although this model may result in larger errors in certain cases, it generally maintains a higher level of stability across a variety of weather conditions, which leads to a smaller MB. In contrast, the GMLSTM-HIM, with its more complex atmospheric dynamics modeling, tends to produce more systematic errors over the long term or under specific conditions, causing a larger MB.

Nevertheless, the GMLSTM-HIM shows a smaller overall RMSE and the absolute values of MB suggest that the magnitude of bias remains at a low level. This demonstrates that while the model exhibits a slight underestimation trend, its predictions remain reliable and robust. Hence, the improved RMSE, combined with the MB magnitude, validates the model's capability and reliability for generating accurate IWV estimates. When we introduce a regression analysis using multiple datasets, as shown in Fig. 5, the figure displays the specific correlation coefficients. It is evident that, overall, the R between $IWV_{GMLSTM\text{-}HIM}$ and IWV_{GNSS} are improved to around 0.99 and demonstrates the superior performance of GMLSTM-HIM for IWV estimation.

In addition, this study evaluated the model's performance at radiosonde stations by comparing the IWV estimates generated by GMLSTM-HIM with the IWV values directly observed by radiosondes. This comparison provides an additional assessment of the model's accuracy and reliability. The results are presented in TABLE VII and Fig. 6, highlighting the strong consistency between $IWV_{GMLSTM\text{-}HIM}$ and the radiosonde observations (IWV_{RS}). The results are shown in Fig. 6. All correlation coefficients exceed 0.9, demonstrating a highly similar variation trend and confirming the model's reliable performance.

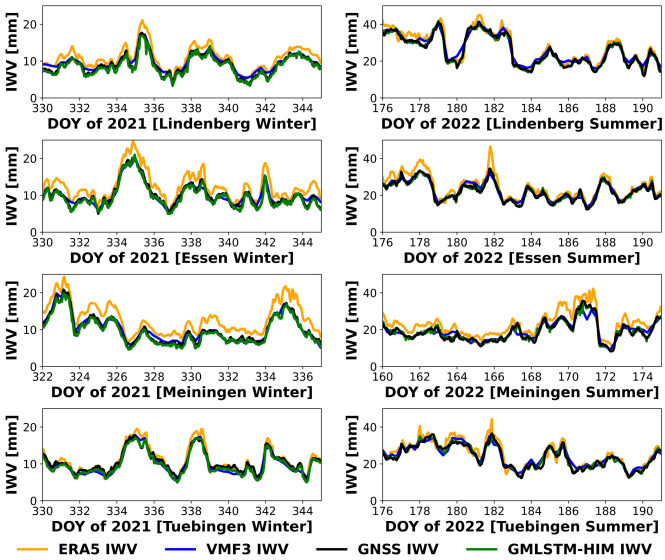

Fig. 4. The results of IWV estimates from ERA5, VMF3, GNSS, and GMLSTM-HIM for the four cases. The left column represents the winter period, while the right column corresponds to the summer period.

Moreover, Fig. 7 presents a comparative analysis of IWV retrieval RMSEs across four radiosonde stations and two seasons using GMLSTM-HIM, VMF3, and ERA5. Although GMLSTM-HIM does not consistently yield the lowest RMSE in every individual case—showing particularly similar performance to VMF3—it clearly outperforms ERA5 in most scenarios. To quantitatively assess the performance differences, a paired t-test was conducted based on RMSE values across all cases. The results show no statistically significant difference between GMLSTM-HIM and VMF3 ($t = 0.421$, $p = 0.68663$), while a significant improvement is observed when comparing GMLSTM-HIM with ERA5 ($t = -3.140$, $p = 0.01639$), indicating that GMLSTM-HIM achieves more accurate IWV estimates at a 95% confidence level.

This performance advantage is further supported by the known characteristics of VMF3-based IWV estimates. Compared to ERA5, IWV values derived from VMF3-based ZWD tend to exhibit higher agreement with radiosonde observations, particularly in regions with complex terrain or sharp moisture gradients. This improvement arises from the finer spatial resolution of GNSS-based retrievals, the utilization of high-resolution meteorological inputs, and the localized representation of atmospheric conditions. In contrast, ERA5—while offering temporally and spatially continuous reanalysis data—relies on grid-based averaging, which can obscure

TABLE VI

THE RMSES AND MBS AT FOUR TEST STATIONS (UNIT: MM) AND THE **BOLDED** DATA REPRESENT THE MINIMUM ABSOLUTE VALUES.

Regions	Station ID	Period	GMLSTM-HIM		ERA5		VMF3	
			RMSE	MB	RMSE	MB	RMSE	MB
Lindenberg	D019	Winter	**0.62**	**-0.45**	1.89	1.76	0.84	0.46
		Summer	**0.72**	**-0.20**	2.35	1.53	1.93	0.79
Essen	D582	Winter	**0.63**	-0.45	2.58	2.28	0.88	**0.03**
		Summer	**0.71**	**-0.06**	3.00	2.35	1.34	0.34
Meiningen	D198	Winter	**0.62**	-0.49	3.11	2.97	0.77	**0.20**
		Summer	**0.99**	-0.08	4.43	3.94	1.93	**0.01**
Tuebingen	D400	Winter	**0.74**	-0.56	1.05	**0.37**	0.92	-0.53
		Summer	**1.59**	**-0.09**	2.62	1.40	1.75	0.51
Mean		Winter	**0.65**	-0.49	2.16	1.85	0.85	**0.04**
		Summer	**1.00**	**-0.11**	3.10	2.31	1.74	0.41

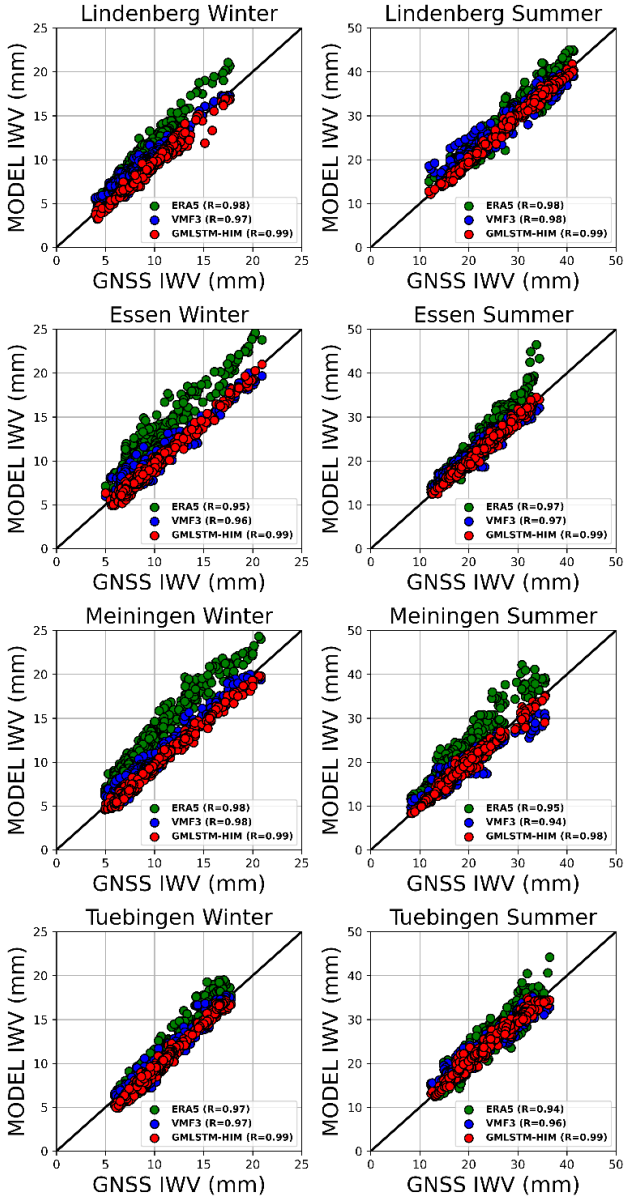

Fig. 5. The correlations between the IWV estimates from ERA5, VMF3, GMLSTM-HIM, and GNSS IWV for the four cases in both winter and summer.

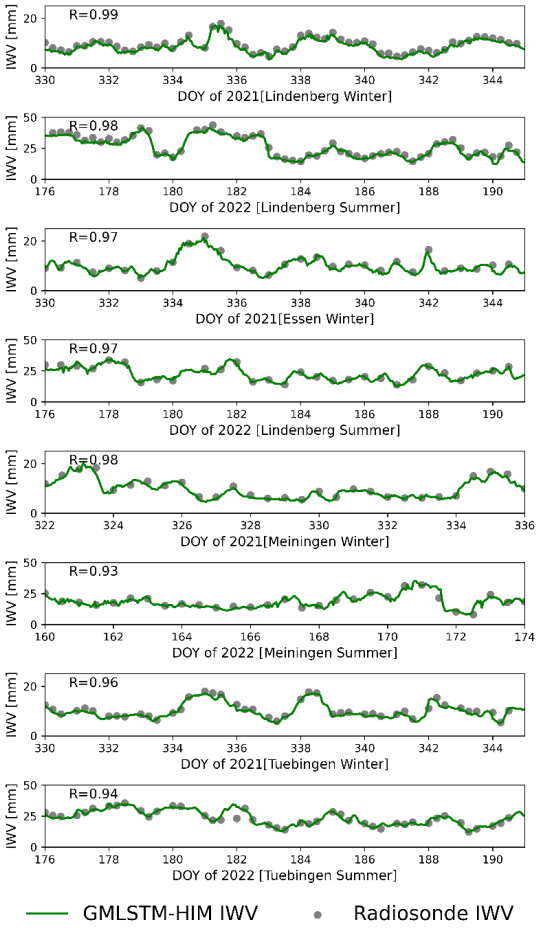

Fig. 6. Comparison of IWV estimates from GMLSTM-HIM and radiosonde data. The green line represents the IWV estimates from GMLSTM-HIM, while the gray dots correspond to radiosonde IWV measurements.

localized humidity structures, resulting in systematic underestimation or overestimation of IWV near radiosonde stations.

While GMLSTM-HIM may not consistently surpass VMF3 in every case, its performance remains robust across diverse seasonal and regional contexts. More importantly, the proposed deep learning framework offers critical advantages over both ERA5 and VMF3, including flexible spatial-temporal

resolution, dynamic adaptation to non-linear atmospheric patterns, and the capability to estimate predictive uncertainties. These attributes enhance its practical applicability in IWV retrieval tasks, particularly in data-scarce areas or under rapidly changing meteorological conditions, supporting both scientific research and operational weather monitoring. Furthermore, building upon the foundational work introduced by Reference [29], the GMLSTM-HIM framework was further applied to eight geographically diverse regions across Europe—Portugal (Beja), Spain (Burgos), France (Paris and Brive-la-Gaillarde), Germany (Tübingen), the Netherlands (Groningen), and Sweden (Sala and Svappavaara). Detailed validation and discussion of these regional experiments were reported in our previous study [29]; therefore, only a brief reference is provided here to avoid redundancy. These regions encompass a wide range of latitudinal positions, topographic elevations, climatic conditions, and distances from coastlines, thereby providing a comprehensive and diverse framework for assessing the spatial generalizability and adaptability of the proposed model. By maintaining consistent input data and experimental configurations, this extended validation effort enabled a comprehensive assessment of the model's sensitivity to seasonal variation and topographic complexity. Under winter conditions, the GMLSTM-HIM model achieved an average RMSE of 0.50 mm across the eight test areas, significantly outperforming ERA5 and VMF3, which exhibited average RMSEs of 1.36 mm and 1.19 mm, respectively. In the summer season, the model maintained superior predictive performance with an average RMSE of 0.93 mm, while ERA5 and VMF3 yielded higher errors of 3.12 mm and 1.66 mm, respectively. These findings are consistent with the patterns observed in the Germany-based case studies presented in this study (see Table 6), further reinforcing the robustness of the model across varying climatic conditions and diverse geographic settings.

To illustrate the model's performance in regions beyond the primary focus of this study, the Paris case was selected as a representative example. The summer and winter evaluation periods were set to July 9 to 23, 2021, and November 26 to December 10, 2021, respectively. Fig. 8(a) illustrates the spatial distribution of GNSS stations used in the region. Despite the limited availability of radiosonde observations during these periods, comparisons with GMLSTM-HIM estimates (Fig. 8b) revealed strong correlations—0.90 in summer and 0.97 in winter—demonstrating the model's ability to capture temporal IWV variability effectively. Furthermore, at the designated test station in Paris (Fig. 8c), GMLSTM-HIM outperformed both ERA5 and VMF3 in both seasons. In winter, RMSEs were 0.92 mm (GMLSTM-HIM), 1.03 mm (VMF3), and 1.52 mm (ERA5); in summer, the respective RMSEs were 1.45 mm, 1.54 mm, and 2.64 mm. The multi-regional validation confirms the model's robustness and adaptability across diverse environmental settings. Given sufficient GNSS coverage—readily achievable via existing national and global networks—the framework demonstrates strong potential for global operational deployment, offering a scalable and reliable solution for enhanced atmospheric water vapor monitoring.

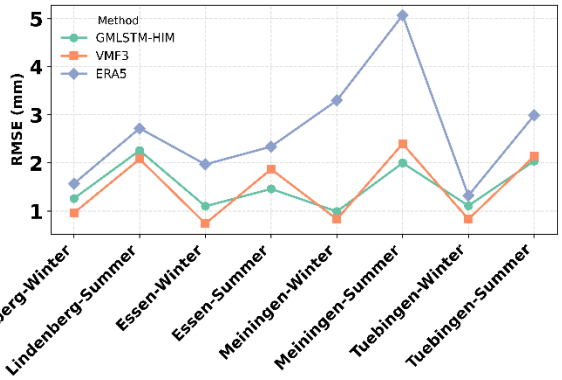

Fig. 7. RMSE of IWV estimates from GMLSTM-HIM, VMF3, and ERA5.

B. Inter-comparison Between GMLSTM-HIM and Satellite-Derived IWV Products

Following the validation of GMLSTM-HIM at both the test site and selected radiosonde stations, the results confirm the applicability and superiority of the proposed model. Building on this validation, the present section focuses on the generation of two-dimensional (2D) IWV fields over the study regions, rather than point-based IWV values at individual stations. To enable a more detailed spatial analysis, four representative regions were selected: Lindenberg (13°E–15°E, 51.5°N–53°N), Essen (6°E–8°E, 51°N–52°N), Meiningen (9.5°E–11°E, 50°N–51°N), and Tuebingen (8°E–10°E, 48°N–49°N). Each region is associated with corresponding satellite-derived IWV data, as shown in the right column of Fig. 2. Four cases were selected within the study period based on the availability of satellite IWV products. The satellite datasets employed include MODIS-Aqua (1 km and 5 km resolutions), MODIS-Terra (1 km and 5 km resolutions), and FY-3D (1km resolution). After collecting all pixels from the satellite products, the geographic locations of each pixel were fixed. Elevation information was subsequently incorporated from DEM. Based on this combined information, 2D IWV distributions were inferred using the trained GMLSTM-HIM. The analysis focuses on relative differences (δ), quantified through pseudomedian (pM) and interquartile range (IQR) statistics, to evaluate systematic biases and spatial consistency. Statistical metrics, including the pseudomedian (pM), interquartile range (IQR), and the number of observations (N), were computed for each satellite product and stratified by region and season (winter/summer), as summarized in TABLE VII. Notably, the FY-3D satellite lacked data coverage for the Lindenberg and Essen sites during winter.

The comparative analysis reveals several important insights into the performance of the proposed GMLSTM-HIM model when benchmarked against satellite-derived IWV estimates. Predominantly negative pM values across various regions indicate a systematic underestimation of IWV by the satellite instruments, suggesting the presence of persistent bias. This underestimation is modulated by seasonal variability. For example, the MODIS-Aqua-1km product shows a pM of −0.59 in Lindenberg during winter, which improves to −0.29 in summer, implying that atmospheric stability and enhanced satellite retrieval accuracy in warmer months reduce discrepancies. Resolution of IWV products also plays a critical

role: higher-resolution satellite products (1 km) typically result in larger IQRs—such as 0.33 for MODIS-Terra-1km in Lindenberg—than their lower-resolution (5 km) counterparts, which show smaller IQRs (e.g., 0.12 for MODIS-Terra-5km), highlighting the increased variability associated with finer spatial scales and the advantage of spatial averaging at coarser resolutions. Despite its finer spatial resolution, the MODIS 1 km IWV product is generally more susceptible to retrieval noise and cloud contamination than the 5 km product. This increased sensitivity primarily stems from the smaller footprint size, which limits spatial averaging and enhances the influence of sub-pixel heterogeneity, surface emissivity variations, and local atmospheric anomalies. Furthermore, at finer spatial scales, each pixel captures less radiative energy, reducing the signal-to-noise ratio and making the measurements more vulnerable to thermal noise intrinsic to the sensor detectors. These sensor-level thermal fluctuations can further degrade retrieval accuracy, particularly under low signal conditions or in regions with strong spatial variability. In contrast, the 5 km product benefits from greater spatial integration, which effectively suppresses random noise, including sensor-level

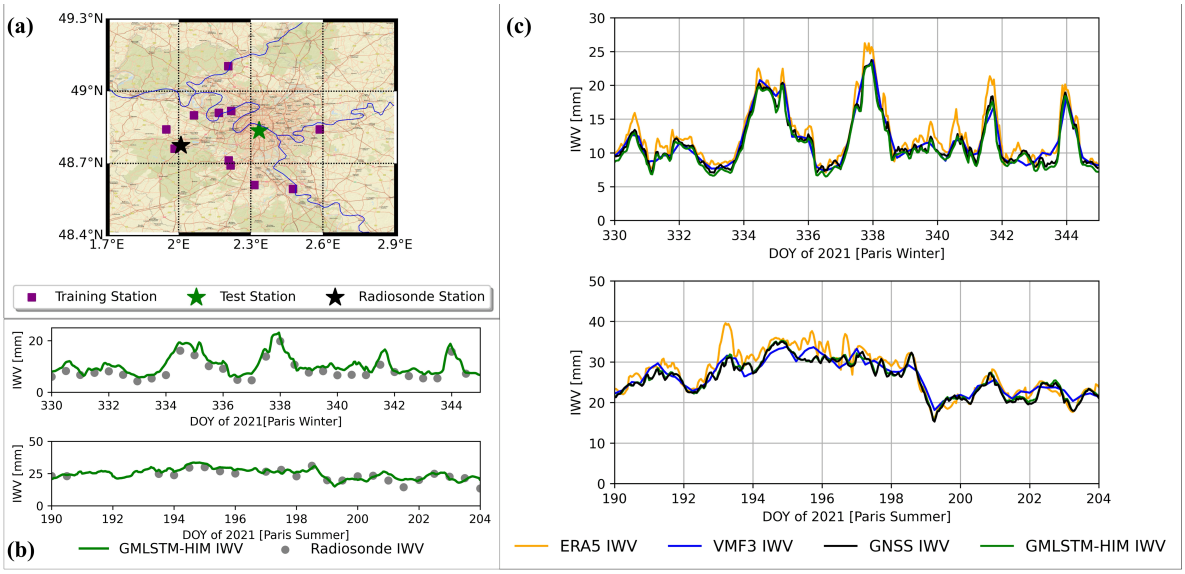

Fig. 8. Validation results for the Paris case study. (a) Spatial distribution of all stations in the selected region; (b) Comparison between GMLSTM-HIM retrieved IWV and radiosonde observations during winter and summer periods, respectively; (c) IWV retrieval performance at the test station, highlighting differences among GMLSTM-HIM, VMF3, GNSS and ERA5 estimates.

TABLE VII

STATISTICAL ANALYSIS OF THE RELATIVE DIFFERENCES (Δ) BETWEEN SATELLITE OBSERVATIONS AND GMLSTM-HIM ESTIMATES.

Satellite	Lindenberg						Essen					
	Winter			Summer			Winter			Summer		
	pM	IQR	N	pM	IQR	N	pM	IQR	N	pM	IQR	N
MODIS-Aqua-1km	-0.59	0.27	259154	-0.29	0.45	251756	-0.43	0.26	178921	-0.32	0.39	161305
MODIS-Aqua-5km	0.01	0.27	1314	-0.01	0.10	5442	0.06	0.23	1645	0.03	0.11	4331
MODIS-Terra-1km	-0.57	0.33	257939	-0.24	0.37	249778	-0.40	0.25	172660	-0.28	0.28	170151
MODIS-Terra-5km	0.02	0.12	2321	0.03	0.11	5769	-0.01	0.16	1675	0.06	0.11	3181
FY-3D	-	-	-	-0.63	0.34	230154	-	-	-	-0.56	0.40	186968

Satellite	Meiningen						Tuebingen					
	Winter			Summer			Winter			Summer		
	pM	IQR	N	pM	IQR	N	pM	IQR	N	pM	IQR	N
MODIS-Aqua-1km	-0.43	0.21	136585	-0.09	0.29	126757	-0.64	0.18	191212	-0.29	0.30	171960

MODIS-Aqua-5km	-0.08	0.18	824	0.08	0.15	5828	-0.09	0.18	885	-0.07	0.15	7925
MODIS-Terra-1km	-0.43	0.22	138198	-0.04	0.26	126156	-0.61	0.21	178551	-0.25	0.30	184562
MODIS-Terra-5km	-0.03	0.11	844	0.15	0.18	5305	-0.11	0.21	847	-0.13	0.15	5247
FY-3D	-0.64	0.12	4768	-0.38	0.26	146838	-0.57	0.18	78722	-0.53	0.25	215732

thermal noise, and enhances overall signal stability. The averaging over larger footprints reduces the impact of pixel-level radiometric fluctuations, thereby improving the signal-to-noise ratio. Moreover, cloud masking and water vapor retrieval algorithms typically perform more robustly at coarser resolutions, where atmospheric and surface conditions exhibit less spatial variability, further contributing to more reliable retrievals. Consequently, although the 1 km product offers greater spatial detail suitable for high-resolution studies, the 5 km product tends to yield more robust and consistent PWV estimates at the regional scale, making it better suited for model validation and broader-scale analyses—a pattern that aligns well with the findings of our inter-comparison.

Spatial heterogeneity in model performance is also evident. Tuebingen, located in a topographically complex area, exhibits the most pronounced biases (pM $\leqslant -0.64$ for MODIS-Aqua-1km in winter), likely due to the challenges in representing moisture gradients over rugged terrain. In addition, the accuracy of the MODIS cloud mask can degrade in mountainous regions, where complex surface reflectance, terrain shadowing, and variable viewing geometry often lead to cloud misclassification or undetected thin clouds. These issues can introduce further uncertainties in water vapor retrievals. In contrast, relatively flat regions such as Essen show much milder deviations (pM $\geqslant -0.43$), underscoring the sensitivity of model performance to orographic effects. The FY-3D product generally exhibits the highest deviations (e.g., −0.63 in Lindenberg during summer), which may be attributed to sensor-specific characteristics, differences in overpass times, or retrieval algorithm inconsistencies when compared to MODIS. Furthermore, the statistical robustness of these comparisons varies with sample size. Observations range from several hundred (e.g., 824 matchups for MODIS-Aqua-5km in Meiningen winter) to over 250,000 (e.g., 259,154 for MODIS-Aqua-1km in Lindenberg winter), which strengthens confidence in large-sample comparisons but may also amplify spatially correlated errors.

The inter-comparison between GMLSTM-HIM and satellite-derived IWV products reveals distinct characteristics for MODIS and FY-3D datasets in Germany. MODIS products, particularly at 5 km resolution, demonstrate better agreement with the model (pM: -0.13 to 0.15) and lower variability (IQR: 0.10-0.27) compared to their 1 km counterparts, which show higher spatial detail but larger biases (pM: -0.64 to -0.04). FY-3D exhibits consistently higher deviations (pM: -0.63 to -0.38) across regions. While MODIS proves more reliable for operational use due to its balanced resolution and temporal coverage, FY-3D can serve as supplementary data. Both satellites perform better in flat regions like Essen than in complex terrain such as Tuebingen, highlighting the need for improved quality control and correction of satellite products in mountainous areas. To optimize IWV estimation performance, a combined use of MODIS products is recommended—using 1 km resolution data for high-precision research applications and 5 km resolution data for operational monitoring—complemented by FY-3D observations for cross-validation purposes, particularly during periods of MODIS data unavailability. While the 1 km MODIS IWV product offers finer spatial detail, it is more susceptible to retrieval noise, cloud contamination, and surface heterogeneity, often leading to increased variability and less stable agreement with ground truth. In contrast, the 5 km product, though coarser in resolution, demonstrates more consistent performance and higher correlation with model-based estimates. This highlights a practical trade-off between spatial resolution and retrieval accuracy: finer-resolution products may capture localized features but are prone to higher uncertainty, whereas moderate-resolution products provide more robust and reliable IWV fields, especially in near-real-time contexts.

Despite these discrepancies, the results validate GMLSTM-HIM's fundamental capacity to reconstruct spatially continuous IWV fields with flexibility in spatial and temporal resolution. Nonetheless, several factors likely contribute to the observed biases: input data limitations stemming from ERA5's grid-based averaging, topographic complexities that are insufficiently resolved in DEMs, and inter-sensor discrepancies among satellite retrievals. Due to space constraints, a subset of representative inter-comparison results is selectively presented in Figs. 9 to 11.

Meanwhile, the uncertainty associated with GMLSTM-HIM in IWV inference, quantified as the predictive variance defined in (12), is illustrated in Figs. 12 to 14 to demonstrate the model's confidence levels across different spatial and temporal domains. Due to space constraints, only selected representative cases are presented. Overall, the results demonstrate that the uncertainty during the winter period is generally lower than in the summer, which can be attributed to the more dynamic and variable atmospheric conditions typically observed during the summer months. Nevertheless, the overall uncertainty remains at a relatively low level across most epochs, suggesting that the model demonstrates robust and reliable performance.

However, certain time periods exhibit notably higher uncertainty. A clear example is observed on DOY 170 in 2022, as highlighted in Fig. 12. On this day, two specific epochs—approximately at around 11:00 and 13:00 UTC—display elevated uncertainty levels. Referring to the test station results at the Meiningen summer station (see Fig. 4), it is evident that during this period, the ERA5-derived IWV is substantially

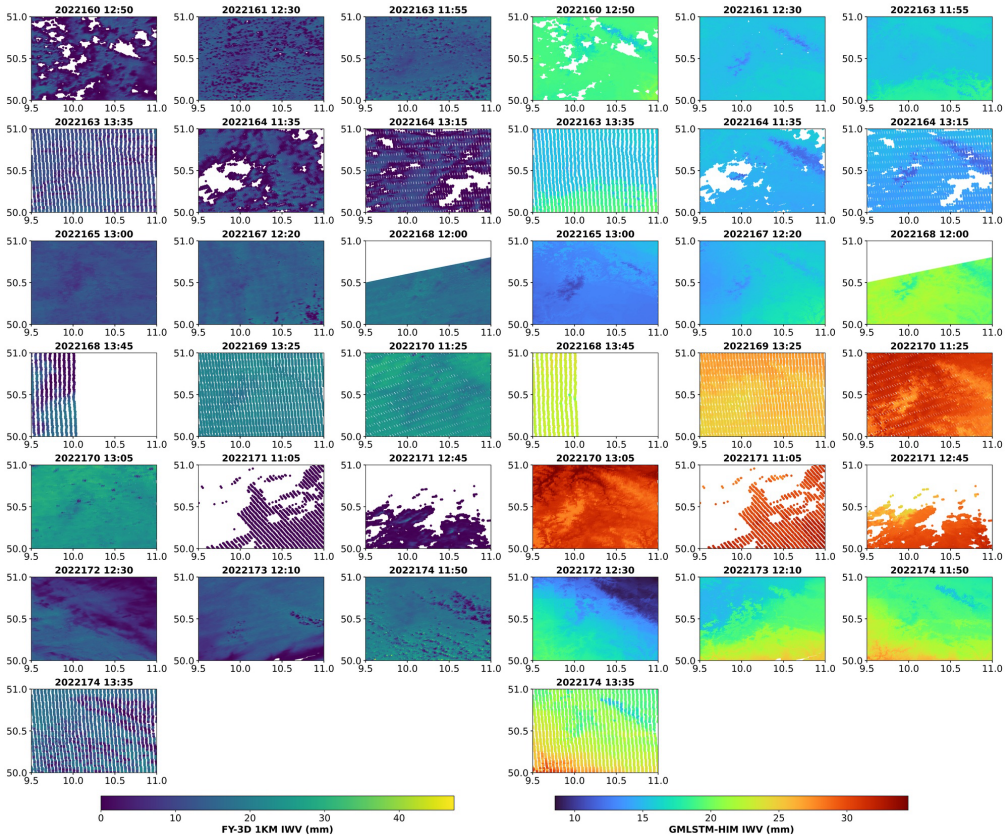

Fig. 9. Inter-comparison between FY-3D-1km observations and GMLSTM-HIM estimates during the summer period in Meiningen.

higher than the corresponding GNSS-derived IWV, while the IWV inferred from the VMF3 model is significantly lower. This discrepancy reveals a substantial systematic bias in the input ZWD values obtained from both ERA5 and VMF3 datasets. Such inconsistency in the input time series introduces non-physical discontinuities, which pose challenges for the model in capturing consistent patterns during training and inference, thereby leading to elevated uncertainty in the predicted IWV. This phenomenon can be explained by the prevailing meteorological conditions on DOY 170. During this time, Meiningen experienced strong convective activity and a heatwave event, characterized by intense surface evaporation and rapid development of the atmospheric boundary layer. ERA5, being a reanalysis model, tends to overestimate near-surface humidity under such conditions, particularly within the boundary layer, resulting in an overestimation of IWV. Conversely, the VMF3 product, which is based on numerical weather prediction outputs with relatively low temporal resolution (typically 6-hour intervals), may fail to capture the rapid moisture build-up occurring around midday. Consequently, the VMF3-derived ZWD underestimates the actual atmospheric water vapor content, leading to a systematic underestimation of IWV. These contrasting biases in the input data sources collectively introduce substantial uncertainty into the model's inference process during this period.

C. Data-Fusion Between GMLSTM-HIM and Satellite-Derived IWV Products

To enhance the spatial consistency and reliability of IWV estimates, a structured data fusion methodology was implemented. This approach involved correcting biases in satellite IWV products through linear regression against GMLSTM-HIM outputs, followed by spatial resampling to a common grid. The corrected satellite data were then averaged with the model-derived IWV to produce a fused product that integrates both observational accuracy and model-based continuity.

To evaluate the effectiveness of the proposed data fusion strategy, two representative cases were analyzed and visualized in Fig. 15. Subfigures (a–c) correspond to the FY-3D 1 km product on day 164 of 2022 at 11:35 UTC, while subfigures (d–f) correspond to the MODIS 5 km product on day 335 of 2021 at 21:55 UTC. To quantitatively demonstrate the benefit of the data fusion, the Interpolated RMSE was computed after interpolating the model-derived IWV field onto the satellite observation grid to ensure spatial consistency. In the 1 km case, fusion reduced Interpolated RMSE from 0.7009 mm to 0.3504 mm. For the 5 km case, Interpolated RMSE decreased from 0.2870 mm to 0.1435 mm. These improvements indicate that the fusion introduces additional satellite-derived details, enhancing data consistency. Direct comparison of the satellite-derived IWV products with the GMLSTM-HIM outputs revealed noticeable white noise in their difference fields. This noise is likely attributed to inherent uncertainties in satellite

sensors, data retrieval algorithms, or potential radiometric calibration errors onboard the satellites. To address this issue, a data fusion approach was employed, which involved two key steps: 1) bias correction of the satellite IWV using linear regression, and 2) spatial resampling to align the satellite data with the GMLSTM-HIM model's grid using nearest-neighbor interpolation, implemented through a k-d tree structure.

As shown in subfigures (b) and (c), and (e) and (f), the fused IWV fields retain the spatial continuity and low-IWV sensitivity characteristic of the GMLSTM-HIM model, while also incorporating the satellite-derived observational features. The fusion process notably improves regions where the model tends to underestimate IWV, particularly under low-humidity conditions. In these areas, the satellite data enrich the model's predictions, enhancing the realism of the field. Conversely, in regions with sparse or noisy satellite data, the GMLSTM-HIM estimates preserve the robustness of the final product. This complementary interaction between satellite data and model outputs underscores the strength of the fusion strategy in balancing the model's predictive reliability and the satellite's observational coverage, ultimately leading to a more accurate and spatially consistent IWV representation.

In addition to improving general IWV estimation, the fusion strategy effectively addresses the problem of extreme outliers in satellite data. It is common for satellite-derived IWV products to exhibit large or small outliers, often caused by uncertainties in the satellite sensors, errors in the retrieval algorithms, or radiometric calibration issues. These extreme values can significantly degrade the quality of the satellite data and introduce scatter in the IWV field. However, the Gaussian Mixture Model-based nature of the GMLSTM-HIM output is inherently designed to smooth the data, preventing such extreme outliers from adversely affecting the final result.

By fusing the satellite data with the GMLSTM-HIM model outputs, the extreme and unreliable values in the satellite data are effectively corrected. The smoothing effect of the model complements the noise in the satellite data, resulting in a more continuous and consistent final IWV field. This fusion process excels not only in correcting low level IWV regions, particularly under low-humidity conditions, but also in mitigating the unreliable effects of extreme outliers in the satellite products. Consequently, this approach significantly enhances the overall accuracy and reliability of the final IWV product, offering a more robust and spatially consistent representation suitable for applications in meteorology and climate monitoring.

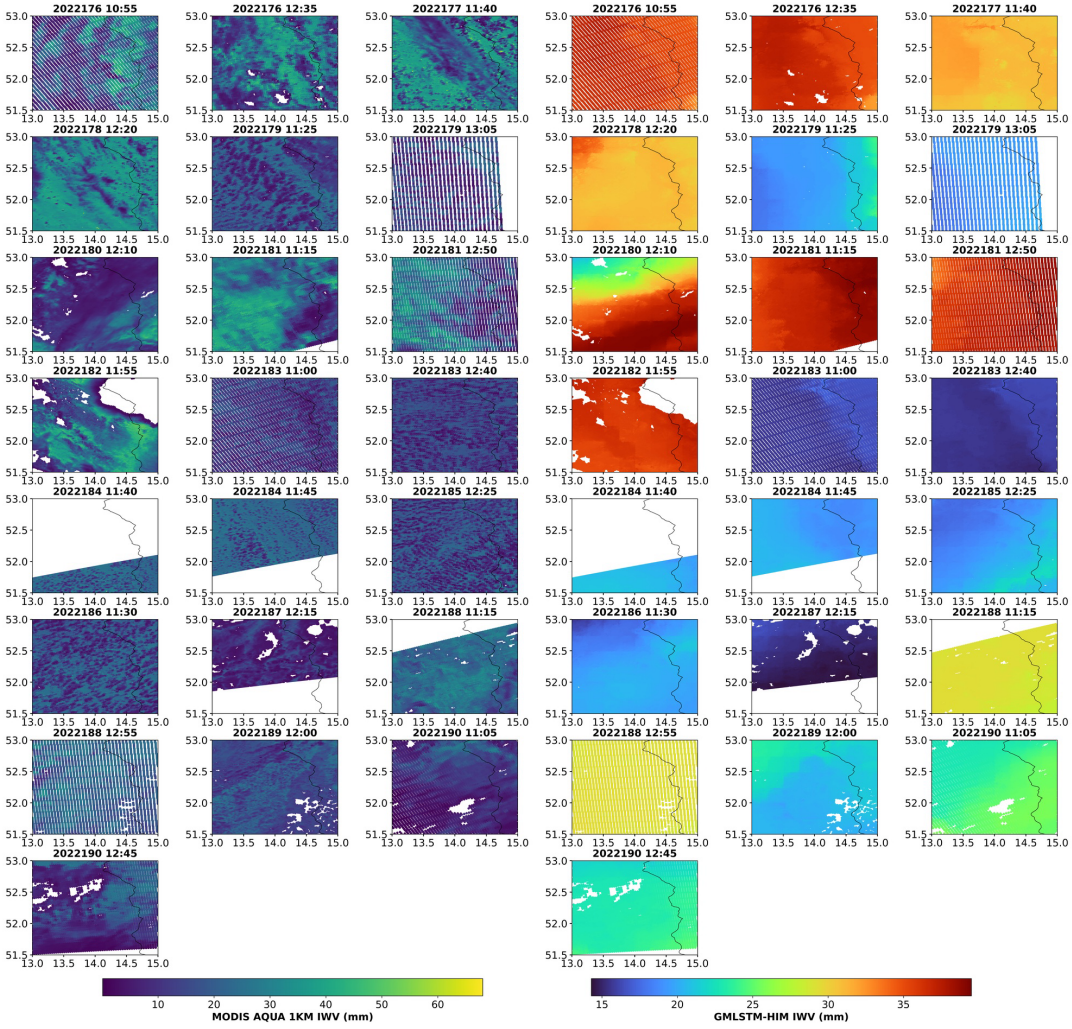

Fig. 10. Inter-comparison between MODIS-Aqua-1km observations and GMLSTM-HIM estimates during the summer period in Lindenberg.

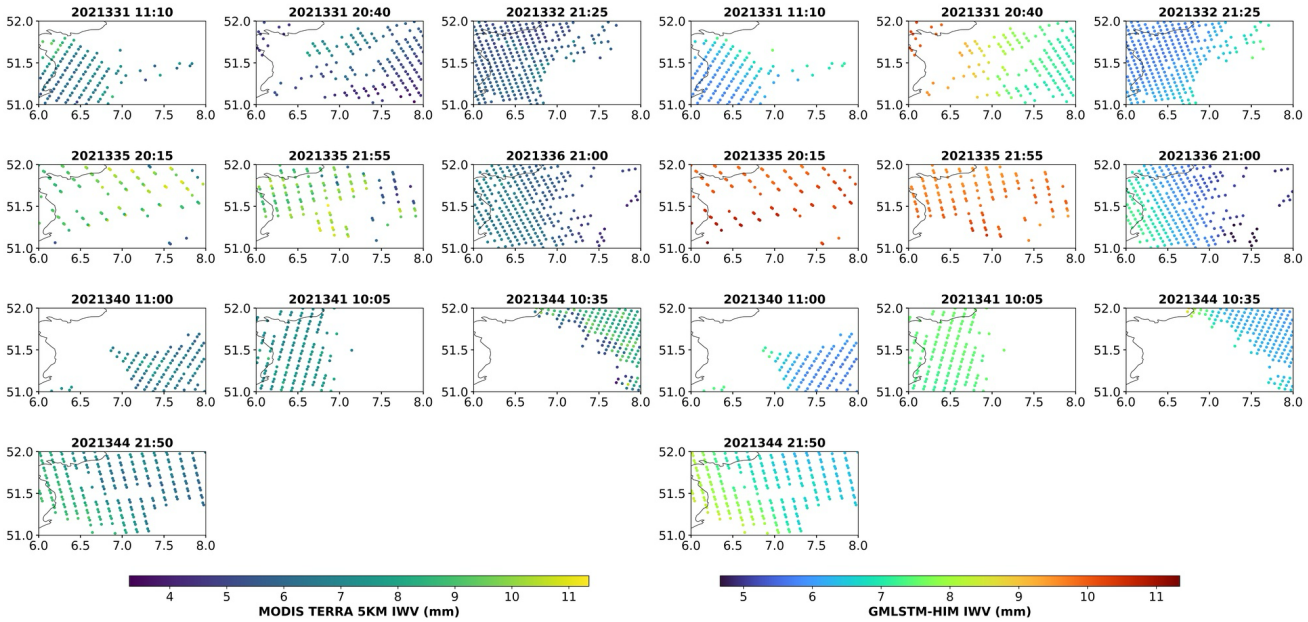

Fig. 11. Inter-comparison between MODIS-Terra-5km observations and GMLSTM-HIM estimates during the winter period in Essen.

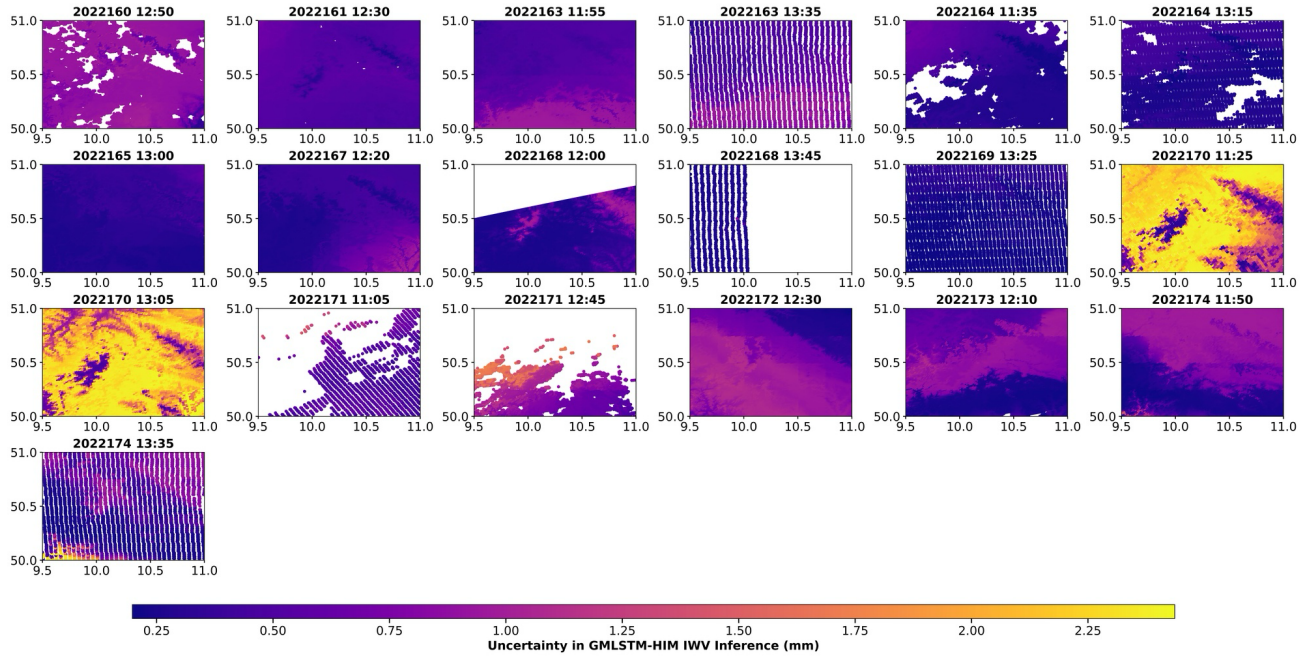

Fig. 12. Uncertainty in GMLSTM-HIM inference during the summer period in Meiningen, corresponding to the inference results shown in Fig. 8.

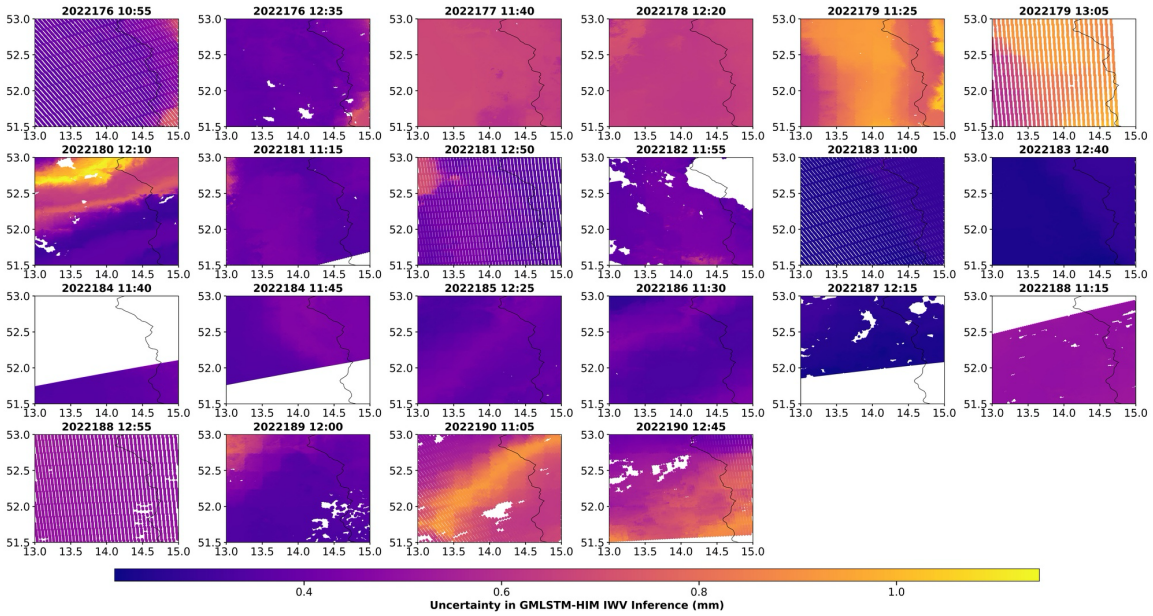

Fig. 13. Uncertainty in GMLSTM-HIM inference during the summer period in Lindenberg, corresponding to the inference results shown in Fig. 9.

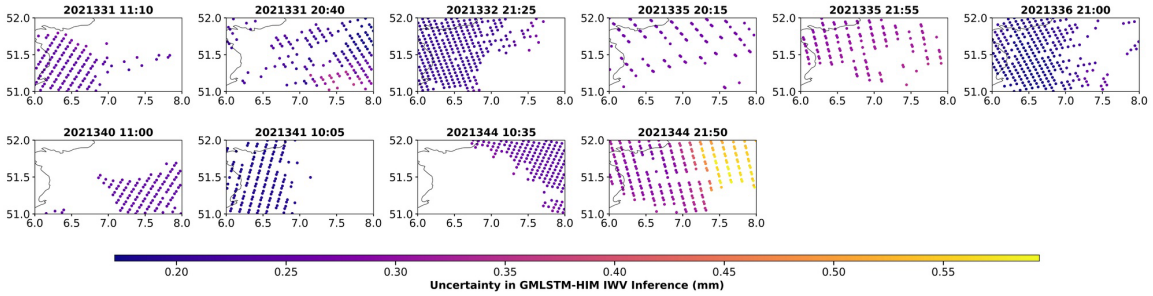

Fig. 14. Uncertainty in GMLSTM-HIM inference during the winter period in Essen, corresponding to the inference results shown in Fig. 10.

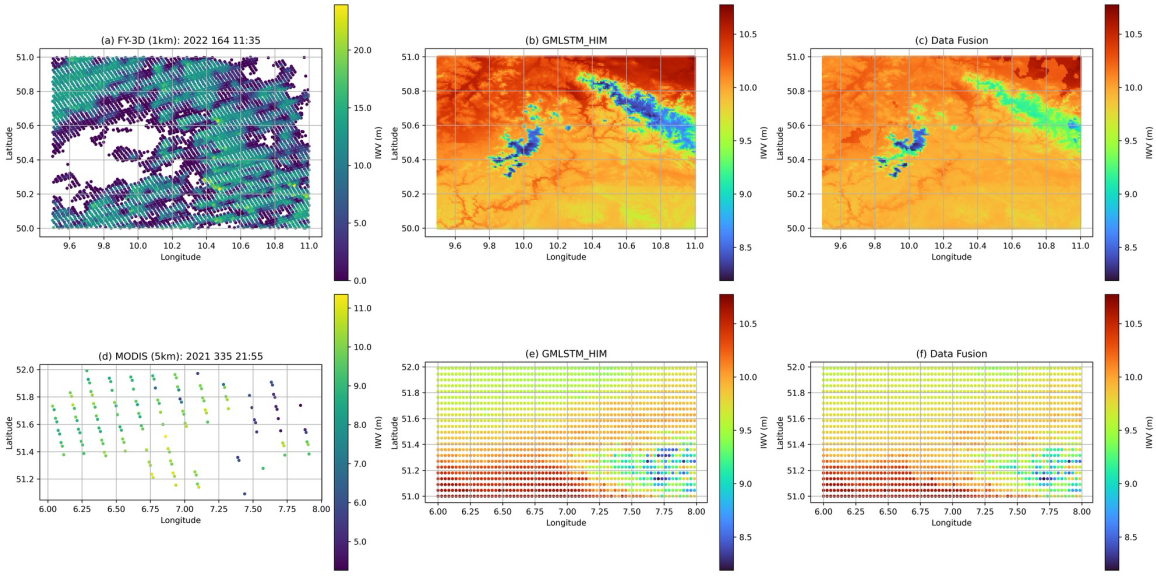

Fig. 15. Data fusion of satellite-derived IWV with GMLSTM-HIM outputs. Subfigures (a), (b), and (c) present results from the FY-3D 1 km case on day 164 of 2022 at 11:35 UTC, while subfigures (d), (e), and (f) correspond to the MODIS 5 km case on day 335 of 2021 at 21:55 UTC. Due to differences in IWV magnitude, the GMLSTM-HIM and fused results share a common color bar scale to facilitate direct comparison.

V. Conclusion

This study proposes a novel deep learning framework GMLSTM-HIM for high-resolution IWV estimation based on GNSS -derived ZWD, NWM-derived ZWD and auxiliary geophysical features. The model captures spatial correlations and temporal evolution of atmospheric moisture with improved accuracy.

Validation at GNSS test stations shows that GMLSTM-HIM achieves a mean RMSE of 0.83 mm and MB of −0.3 mm, outperforming ERA5 (RMSE: 2.63 mm, MB: 2.08 mm) and VMF3 (RMSE: 1.30 mm, MB: 0.36 mm). These results correspond to RMSE reductions of 68.44% over ERA5 and 36.15% over VMF3. Although minor MB deviations remain at stations like Meiningen, especially in winter, the overall magnitude remains low, confirming reliable model performance. Correlation coefficients with GNSS-derived IWV reach ~0.99, indicating strong agreement. Cross-validation with radiosonde IWV further demonstrates robustness, with all correlation coefficients above 0.9. RMSE comparisons show GMLSTM-HIM performs similarly to VMF3 and better than ERA5 in most cases. A paired t-test confirms that the improvement over ERA5 is statistically significant ($p = 0.01639$), while differences with VMF3 are not ($p = 0.68663$), indicating competitive accuracy.

Beyond point-based evaluation, the model reconstructs 2D IWV fields across four regions. Comparisons with MODIS and FY-3D satellite IWV reveal consistent spatial patterns and highlight the influence of resolution, season, and topography. MODIS-5 km products show better agreement (e.g., pM: −0.13 to 0.15; IQR: 0.10–0.27), while 1 km products display higher variability. FY-3D shows larger deviations (pM: −0.63 to −0.38), likely due to sensor-specific differences.

Uncertainty estimation results indicate low overall predictive uncertainty, particularly in winter. Notable spikes occur during convective or unstable conditions, such as DOY 170 in 2022. On this day, strong discrepancies between input sources (ERA5 overestimation, VMF3 underestimation) under convective activity increased inference uncertainty. This highlights the model's sensitivity to input quality during extreme weather events.

During the data fusion process between satellite observations and GMLSTM-HIM outputs, high-frequency random fluctuations—characteristic of white noise—were identified within the satellite IWV datasets. These residuals remained evident in the discrepancies between satellite and GMLSTM-HIM estimates, thereby complicating systematic bias correction. This phenomenon underscores the inherent sensitivity of satellite retrievals to noise, cloud contamination, and surface heterogeneity, highlighting the necessity of incorporating denoising strategies in the future development of satellite-derived IWV products.

In summary, GMLSTM-HIM offers accurate, spatiotemporally flexible, and robust IWV estimation. It consistently outperforms ERA5 and shows performance comparable to VMF3, with added benefits such as uncertainty quantification and adaptability to complex atmospheric dynamics. These advantages support its application in both research and operational IWV retrieval, especially in data-sparse or rapidly changing environments. Future work will focus on extending the proposed framework to oceanic regions and integrating BeiDou B2b-PPP tropospheric parameters to further enhance its retrieval capability and spatial applicability.

References

[1] Y.-A. Liou, Y.-T. Teng, T. Van Hove, and J. C. Liljegren, "Comparison of Precipitable Water Observations in the Near Tropics by GPS, Microwave Radiometer, and Radiosondes," J. Appl. Meteor., vol. 40, no. 1, pp. 5–15, Jan. 2001, doi: 10.1175/1520-0450(2001)040<0005:copwoi>2.0.co;2.

[2] H. Li, S. Choy, X. Wang, H. Liang, and K. Zhang, "Monitoring the Migration of Water Vapor Using Ground-Based GNSS Tropospheric Products," IEEE Geosci. Remote Sensing Lett., vol. 20, pp. 1–5, 2023, doi: 10.1109/lgrs.2023.3282362.

[3] A. E. Dessler, Z. Zhang, and P. Yang, "Water-vapor climate feedback inferred from climate fluctuations, 2003–2008," Geophysical Research Letters, vol. 35, no. 20, Oct. 2008, doi: 10.1029/2008gl035333.

[4] F. Shi., "The first validation of the precipitable water vapor of multisensor satellites over the typical regions in China," Remote Sensing of Environment, vol. 206, pp. 107–122, Mar. 2018, doi: 10.1016/j.rse.2017.12.022.

[5] J. Vaquero-Martínez et al., "Inter-comparison of integrated water vapor from satellite instruments using reference GPS data at the Iberian Peninsula," Remote Sensing of Environment, vol. 204, pp. 729–740, Jan. 2018, doi: 10.1016/j.rse.2017.09.028.

[6] Q. Zhao et al., "General method of precipitable water vapor retrieval from remote sensing satellite near-infrared data," Remote Sensing of Environment, vol. 308, p. 114180, July 2024, doi: 10.1016/j.rse.2024.114180.

[7] C. Ichoku et al., "Analysis of the performance characteristics of the five-channel Microtops II Sun photometer for measuring aerosol optical thickness and precipitable water vapor," J.-Geophys.-Res., vol. 107, no. D13, July 2002, doi: 10.1029/2001jd001302.

[8] D. D. Turner, S. A. Clough, J. C. Liljegren, E. E. Clothiaux, K. E. Cady-Pereira, and K. L. Gaustad, "Retrieving Liquid Wat0er Path and Precipitable Water Vapor From the Atmospheric Radiation Measurement (ARM) Microwave Radiometers," IEEE Trans. Geosci. Remote Sensing, vol. 45, no. 11, pp. 3680–3690, Nov. 2007, doi: 10.1109/tgrs.2007.903703.

[9] D. Pérez-Ramírez, F. Navas-Guzmán, H. Lyamani, J. Fernández-Gálvez, F. J. Olmo, and L. Alados-Arboledas, "Retrievals of precipitable water vapor using star photometry: Assessment with Raman lidar and link to sun photometry," J. Geophys. Res., vol. 117, no. D5, Mar. 2012, doi: 10.1029/2011jd016450.

[10] A. Barreto, E. Cuevas, B. Damiri, P. M. Romero, and F. Almansa, "Column water vapor determination in night period with a lunar photometer prototype," Atmos. Meas. Tech., vol. 6, no. 8, pp. 2159–2167, Aug. 2013, doi: 10.5194/amt-6-2159-2013.

[11] X. Li et al., "Multi-GNSS Meteorology: Real-Time Retrieving of Atmospheric Water Vapor From BeiDou, Galileo, GLONASS, and GPS Observations," IEEE Trans. Geosci. Remote Sensing, vol. 53, no. 12, pp. 6385–6393, Dec. 2015, doi: 10.1109/tgrs.2015.2438395.

[12] M. Migliaccio and A. Gambardella, "Microwave radiometer spatial resolution enhancement," IEEE Trans. Geosci. Remote Sensing, vol. 43, no. 5, pp. 1159–1169, May 2005, doi: 10.1109/tgrs.2005.844099.

[13] D. J. Seidel, B. Sun, M. Pettey, and A. Reale, "Global radiosonde balloon drift statistics," J. Geophys. Res., vol. 116, no. D7, Apr. 2011, doi: 10.1029/2010jd014891.

[14] A. Jones et al., "Evolution of stratospheric ozone and water vapour time series studied with satellite measurements," Atmos. Chem. Phys., vol. 9, no. 16, pp. 6055–6075, Aug. 2009, doi: 10.5194/acp-9-6055-2009.

[15] Y. S. Bennouna, B. Torres, V. E. Cachorro, J. P. Ortiz de Galisteo, and C. Toledano, "The evaluation of the integrated water vapour annual cycle over the Iberian Peninsula from EOS-MODIS against different ground-based techniques," Quart J Royal Meteoro Soc, vol. 139, no. 676, pp. 1935–1956, Jan. 2013, doi: 10.1002/qj.2080.

[16] H. Wang, X. Liu, K. Chance, G. González Abad, and C. Chan Miller, "Water vapor retrieval from OMI visible spectra," Atmos. Meas. Tech., vol. 7, no. 6, pp. 1901–1913, June 2014, doi: 10.5194/amt-7-1901-2014.

[17] M. Grossi et al., "Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B," Atmos. Meas. Tech., vol. 8, no. 3, pp. 1111–1133, Mar. 2015, doi: 10.5194/amt-8-1111-2015.

[18] R. Román et al., “Comparison of total water vapor column from GOME-2 on MetOp-A against ground-based GPS measurements at the Iberian Peninsula,” Science of The Total Environment, vol. 533, pp. 317–328, Nov. 2015, doi: 10.1016/j.scitotenv.2015.06.124.
[19] M. Bevis et al., “GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water,” J. Appl. Meteor., vol. 33, no. 3, pp. 379–386, Mar. 1994, doi: 10.1175/1520-0450(1994)033<0379:gmmzwd>2.0.co;2.
[20] K. Wilgan, T. Hadas, P. Hordyniec, and J. Bosy, “Real-time precise point positioning augmented with high-resolution numerical weather prediction model,” GPS Solut, vol. 21, no. 3, pp. 1341–1353, Mar. 2017, doi: 10.1007/s10291-017-0617-6.
[21] R. C. Ssenyunzi et al., “Performance of ERA5 data in retrieving Precipitable Water Vapour over East African tropical region,” Advances in Space Research, vol. 65, no. 8, pp. 1877–1893, Apr. 2020, doi: 10.1016/j.asr.2020.02.003.
[22] P. Yuan et al., “Feasibility of ERA5 integrated water vapor trends for climate change analysis in continental Europe: An evaluation with GPS (1994–2019) by considering statistical significance,” Remote Sensing of Environment, vol. 260, p. 112416, July 2021, doi: 10.1016/j.rse.2021.112416.
[23] L. Huang et al., “High-precision GNSS PWV retrieval using dense GNSS sites and in-situ meteorological observations for the evaluation of MERRA-2 and ERA5 reanalysis products over China,” Atmospheric Research, vol. 276, p. 106247, Oct. 2022, doi: 10.1016/j.atmosres.2022.106247.
[24] L. Wang and H. Kutterer, “An advanced regional integrated water vapor estimation model utilizing least squares support vector machine for the upper Rhine graben region,” GPS Solut, vol. 29, no. 1, Jan. 2025, doi: 10.1007/s10291-024-01796-3.
[25] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” Neural Networks, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.
[26] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning An MIT Press Book*, vol. 29, no. 7553. 2016.
[28] C. Lu et al., “TropNet: a deep spatiotemporal neural network for tropospheric delay modeling and forecasting,” J Geod, vol. 97, no. 4, Mar. 2023, doi: 10.1007/s00190-023-01722-4.
[29] D. Wang, L. Wang, and H. Kutterer, “An Advanced Tropospheric Delay Model Based on Gaussian Mixed Long Short-Term Memory Network,” IEEE Trans. Geosci. Remote Sensing, vol. 63, pp. 1–14, 2025, doi: 10.1109/tgrs.2025.3549124.
[30] A. Klos, J. Bogusz, R. Pacione, V. Humphrey, and H. Dobslaw, “Investigating temporal and spatial patterns in the stochastic component of ZTD time series over Europe,” GPS Solut, vol. 27, no. 1, Nov. 2022, doi: 10.1007/s10291-022-01351-y.
[31] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” Neural Networks, vol. 18, no. 5–6, pp. 602–610, July 2005, doi: 10.1016/j.neunet.2005.06.042.
[32] T. Nilsson, J. Böhm, D. D. Wijaya, A. Tresch, V. Nafisi, and H. Schuh, “Path Delays in the Neutral Atmosphere,” Springer Atmospheric Sciences. Springer Berlin Heidelberg, pp. 73–136, 2013. doi: 10.1007/978-3-642-36932-2_3.
[33] C. Yu, N. T. Penna, and Z. Li, "Generation of real-time mode high-resolution water vapor fields from GPS observations," JGR Atmospheres, vol. 122, no. 3, pp. 2008–2025, Feb. 2017, doi: 10.1002/2016jd025753.
[34] J. Kestin, J. V. Sengers, B. Kamgar-Parsi, and J. M. H. L. Sengers, “Thermophysical Properties of Fluid H2O,” Journal of Physical and Chemical Reference Data, vol. 13, no. 1, pp. 175–183, Jan. 1984, doi: 10.1063/1.555707.
[35] M. Bevis, S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, “GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system,” J. Geophys. Res., vol. 97, no. D14, pp. 15787–15801, Oct. 1992, doi: 10.1029/92jd01517.
[36] H. Namaoui, S. Kahlouche, and A. H. Belbachir, “Evaluation of MODIS Water Vapour Products over ALGERIA using Radiosonde Data,” Anu. Inst. Geoc., vol. 44, Oct. 2021, doi: 10.11137/1982-3908_2021_44_40110.
[37] C. Jiang et al., “Comparison of ZTD derived from CARRA, ERA5 and ERA5-Land over the Greenland based on GNSS,” Advances in Space Research, vol. 72, no. 11, pp. 4692–4706, Dec. 2023, doi: 10.1016/j.asr.2023.09.002.
[38] S. Jade and M. S. M. Vijayan, "GPS-based atmospheric precipitable water vapor estimation using meteorological parameters interpolated from NCEP global reanalysis data," J. Geophys. Res., vol. 113, no. D3, Feb. 2008, doi: 10.1029/2007jd008758.
[39] D. Landskron and J. Böhm, “VMF3/GPT3: refined discrete and empirical troposphere mapping functions,” J Geod, vol. 92, no. 4, pp. 349–360, Sept. 2017, doi: 10.1007/s00190-017-1066-2.
[40] J. Kouba, “Testing of global pressure/temperature (GPT) model and global mapping function (GMF) in GPS analyses,” J Geod, vol. 83, no. 3–4, pp. 199–208, Feb. 2009, doi: 10.1007/s00190-008-0229-6.
[41] J. Kouba, “Implementation and testing of the gridded Vienna Mapping Function 1 (VMF1),” J Geod, vol. 82, no. 4–5, pp. 193–205, June 2007, doi: 10.1007/s00190-007-0170-0.
[42] M. D. King, Y. J. Kaufman, W. P. Menzel, and D. Tanre, “Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS),” IEEE Trans. Geosci. Remote Sensing, vol. 30, no. 1, pp. 2–27, 1992, doi: 10.1109/36.124212.
[43] B.-C. Gao and R.-R. Li, “The Time Series of Terra and Aqua MODIS Near-IR Water Vapor Products,” IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium. IEEE, p. III-186-III–189, 2008. doi: 10.1109/igarss.2008.4779314.
[44] S. W. Seemann, J. Li, W. P. Menzel, and L. E. Gumley, “Operational Retrieval of Atmospheric Temperature, Moisture, and Ozone from MODIS Infrared Radiances,” J. Appl. Meteor., vol. 42, no. 8, pp. 1072–1091, Aug. 2003, doi: 10.1175/1520-0450(2003)042<1072:oroatm>2.0.co;2.
[45] C. Dong et al., “An Overview of a New Chinese Weather Satellite FY-3A,” Bull. Amer. Meteor. Soc., vol. 90, no. 10, pp. 1531–1544, Oct. 2009, doi: 10.1175/2009bams2798.1.
[46] J. Yang, P. Zhang, N. Lu, Z. Yang, J. Shi, and C. Dong, “Improvements on global meteorological observations from the current Fengyun 3 satellites and beyond,” International Journal of Digital Earth, vol. 5, no. 3, pp. 251–265, May 2012, doi: 10.1080/17538947.2012.658666.
[47] R. M. Parinussa et al., “Global surface soil moisture from the Microwave Radiation Imager onboard the Fengyun-3B satellite,” International Journal of Remote Sensing, vol. 35, no. 19, pp. 7007–7029, Oct. 2014, doi: 10.1080/01431161.2014.960622.
[48] X. Li et al., “LEO–BDS–GPS integrated precise orbit modeling using FengYun-3D, FengYun-3C onboard and ground observations,” GPS Solut, vol. 24, no. 2, Feb. 2020, doi: 10.1007/s10291-020-0962-8.
[49] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
[50] F. Wilcoxon, “Individual Comparisons by Ranking Methods,” Springer Series in Statistics. Springer New York, pp. 196–202, 1992. doi: 10.1007/978-1-4612-4380-9_16.
[51] G. Mutabdzija, *Regional Geography of Europe*, Columbia, SC, 2018. (No DOI available)

Lingke Wang is currently a Ph.D. candidate at Karlsruhe Institute of Technology (KIT), Germany. He received his master’s degree from University of Stuttgart, Germany in 2021 and received his bachelor’s degree from Wuhan University, China in 2019. His research interests include GNSS data processing, tropospheric water vapor model and multipath effect mitigation model.

Duo Wang received the B.Sc. degree in information and computing science and M.Sc. degree in computer science from the Beijing University of Technology, China, in 2015 and 2020. He earned his Ph.D. in Geodesy and Geoinformatics from the Karlsruhe Institute of Technology (KIT, Germany) under the supervision of Prof. Hansjörg Kutterer in 2025. Currently, he is a postdoctoral researcher at the Department of Land Surveying and Geo-Informatics (LSGI), The Hong Kong Polytechnic

University. His research focuses on deep learning in Geodesy, particularly in the field of InSAR, tropospheric modeling, and GNSS-based climate and environmental studies.

Joseph L. Awange is University Professor for Environmental Geoinformatics at the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, since 2024. He received his B.Sc. and M.Sc. degrees in Surveying from the University of Nairobi, Kenya, and later earned a second M.Sc. and Ph.D. in Geodesy from the University of Stuttgart, Germany, under the supervision of Prof. Erik W. Grafarend. From 2002 to 2004, he was a JSPS Fellow at Kyoto University, Japan. He held academic positions at Curtin University, Australia, where he served as Professor of Geospatial Science until 2023. He has been actively involved in international scientific publishing and currently serves as Acquisition Manager for Springer Nature in Earth Sciences, Geography, and Environment. His research interests include geospatial and AI-based techniques for environmental change detection, hydroclimate variability, GNSS and InSAR applications, and mathematical geosciences. He has authored over 25 scholarly books and more than 250 peer-reviewed journal articles.

Hansjörg Kutterer is University Professor for Geodetic Earth System Science at the Karlsruhe Institute of Technology, Germany, since 2018. He received his Dipl.-Ing., Dr.-Ing. and Dr.-Ing. habil. degrees at the University of Karlsruhe in 1990, 1993 and 2001, respectively. From 2004 to 2011 he was University Professor at Leibniz University Hannover and from 2011 to 2018 he was Director General of the Federal Agency of Cartography and Geodesy in Frankfurt/Main, Germany. He has been holding various high-level positions in national and international administrative and scientific committees. At present, e.g., he is Assistant Editor-in-Chief of the Journal of Geodesy. His research interests refer to satellite-based geodetic observation techniques and systems such as GNSS and InSAR, to geodetic modelling and data analysis and their relation to Earth system research.