Machine learning for atmospheric delay correction in geodesy: An advanced troposphere delay model based on Gaussian mixed long short-term memory network

Duo Wang[1], Lingke Wang[1], Hansjörg Kutterer[1]

1.Geodetic Institute, Karlsruhe Institute of Technology

Introduction & Motivation

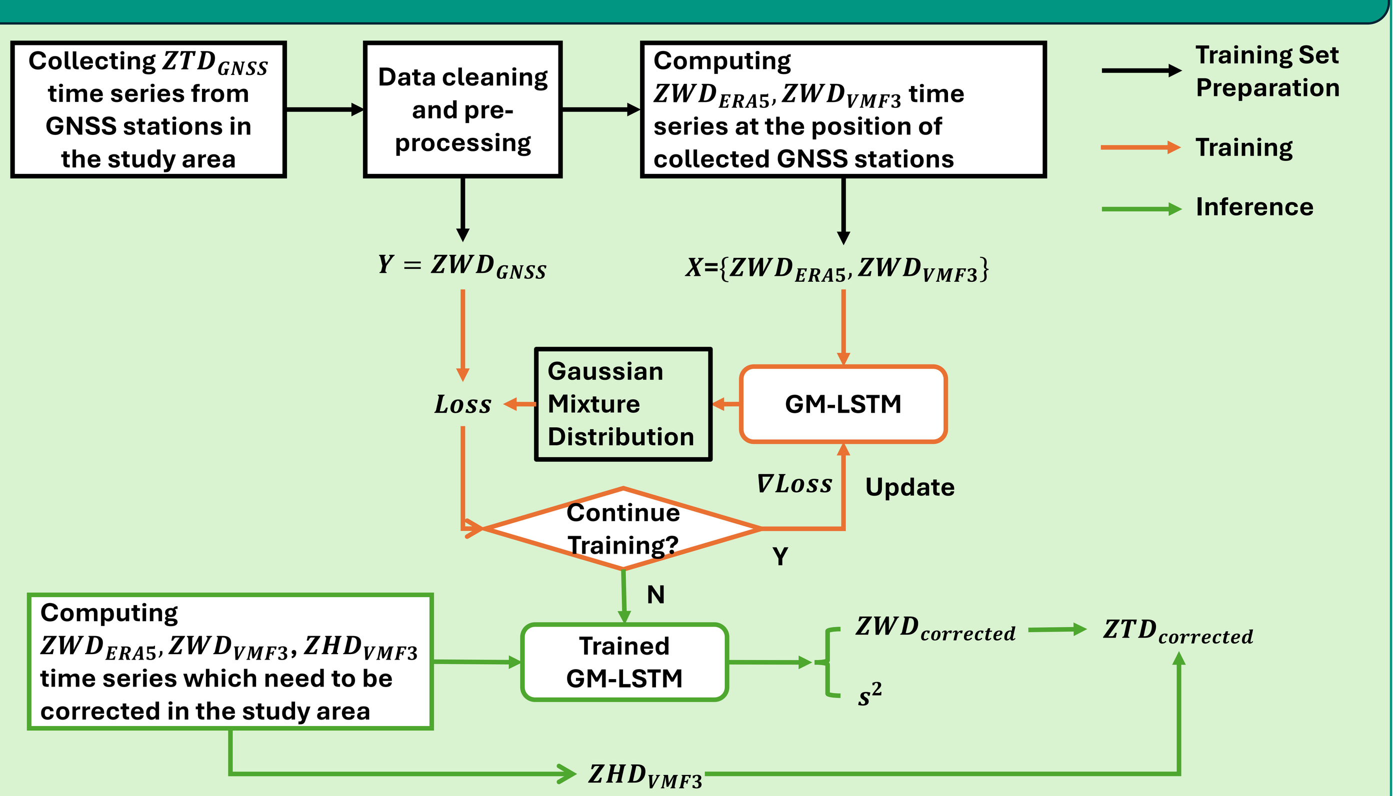
Tropospheric delay constitutes one of the most significant error sources in the fields of geodesy and spaceborne remote sensing.

While Global Navigation Satellite Systems (GNSS) can yield reliable Zenith Total Delay (ZTD) products as troposphere delay estimates by simultaneously observing multiple satellites, its spatial resolution remains limited due to the sparse distribution of GNSS stations. On the contrary, meteorological models (ERA5) or empirical models (VMF3) can provide estimated gridded ZTD data with 0.25° (ERA5) or 1° (VMF3) resolution. However, compared to ZTD retrieved by GNSS, the biases of ZTD retrieved from ERA5 or VMF3 are nearly 10-15 mm, which limits their application in fields like ground motion monitoring, real-time positioning, and GNSS meteorology.

To overcome these problems, we provide a Gaussian mixed long short-term memory network (GM-LSTM) to learn the feature of mapping $\{ZTD_{ERA5}, ZTD_{VMF3}\} \rightarrow ZTD_{GNSS}$ for this purpose. Then the precise ZTD at any position within the study area can be estimated through the trained model.

This study deals with deep learning to aggregate GNSS observation, combined with meteorological model data to generate precise ZTD estimation at any location. The outcome of this study can be directly used in the atmospheric correction of InSAR and GNSS to improve the accuracy of applications like deformation monitoring, positioning, and Integrated water vapor retrieval.

Methodology



- This approach uses LSTM to capture time-dependent features of Zenith Wet Delay (ZWD).
- ZWD is treated as a series of Gaussian Mixture Distributions rather than a sequence of numerical values.
- The variance of the learned distribution can reflect the uncertainty of water vapor activity.
- The model task is not regression, but the estimation of distribution parameters.

Methodology

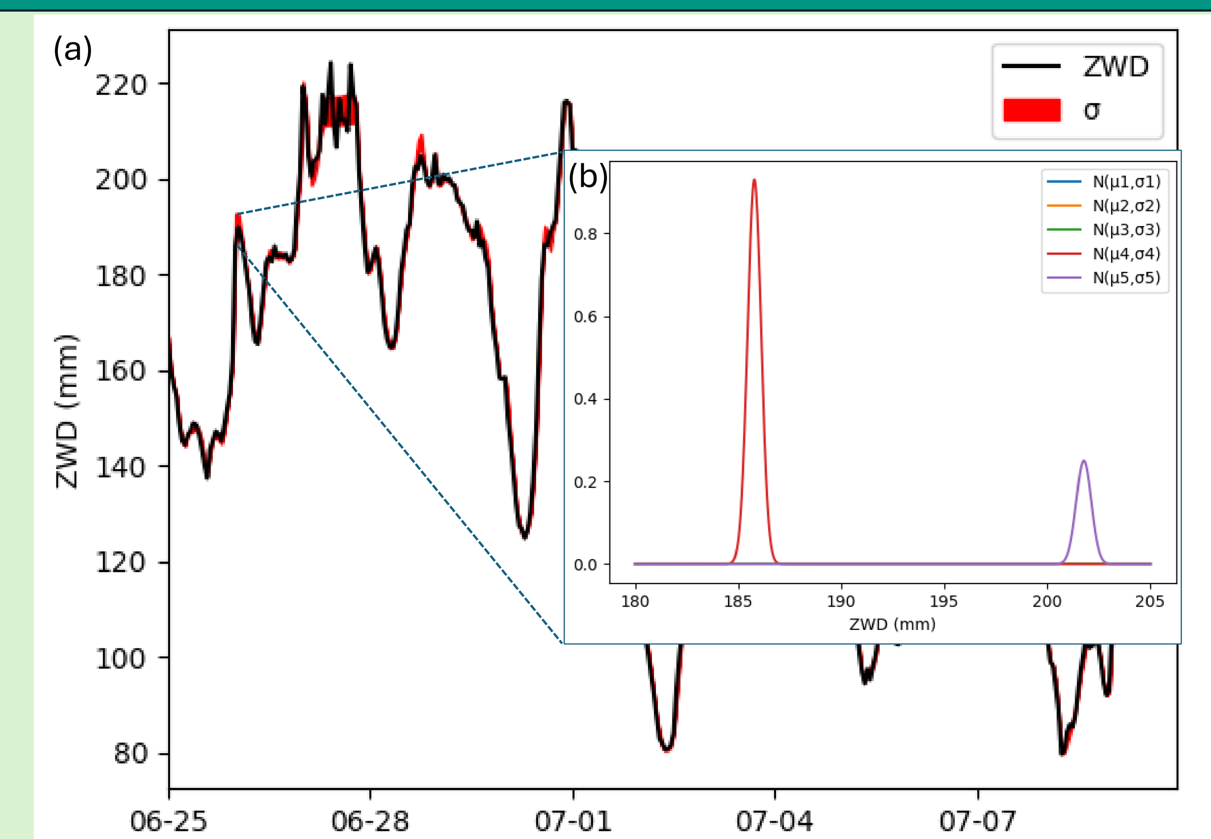


Fig.1 (a) The $ZWD_{GM-LSTM}$ estimation with variance for D400, Tübingen. (b) The Gaussian Mixture Distribution at 2022.06.26 01:00. The two peaks indicate increased uncertainty of estimation due to water vapor activity at that time.

Data & Experiments

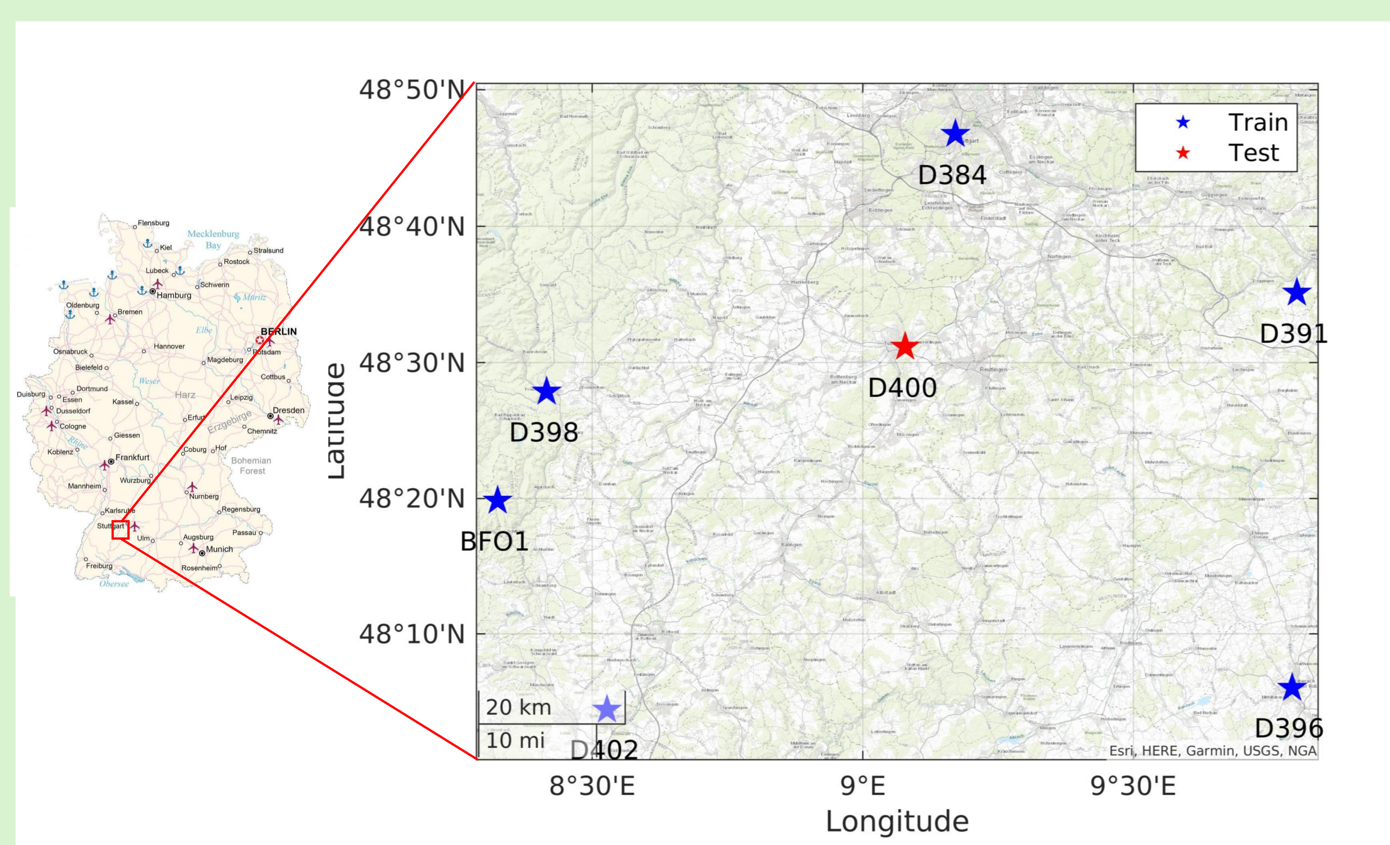


Fig.2 The location of GNSS stations for training and testing.

In order to evaluate the performance of the ZTD predicted by GM-LSTM at any location in the study area, we selected 6 GNSS stations near Tübingen, Baden-Württemberg, Germany for training and the station named D400 in Tübingen for testing. The test station is invisible to the model, allowing it to evaluate the performance at any location within the region.

Considering the seasonal variability of tropospheric delay, we tested the model's performance in winter (2021-11-26 to 2021-12-10) and summer (2022-06-25 to 2022-07-09), respectively.

For more details, please follow our github page: https://github.com/hgwxx1945/DeepZTD

Results

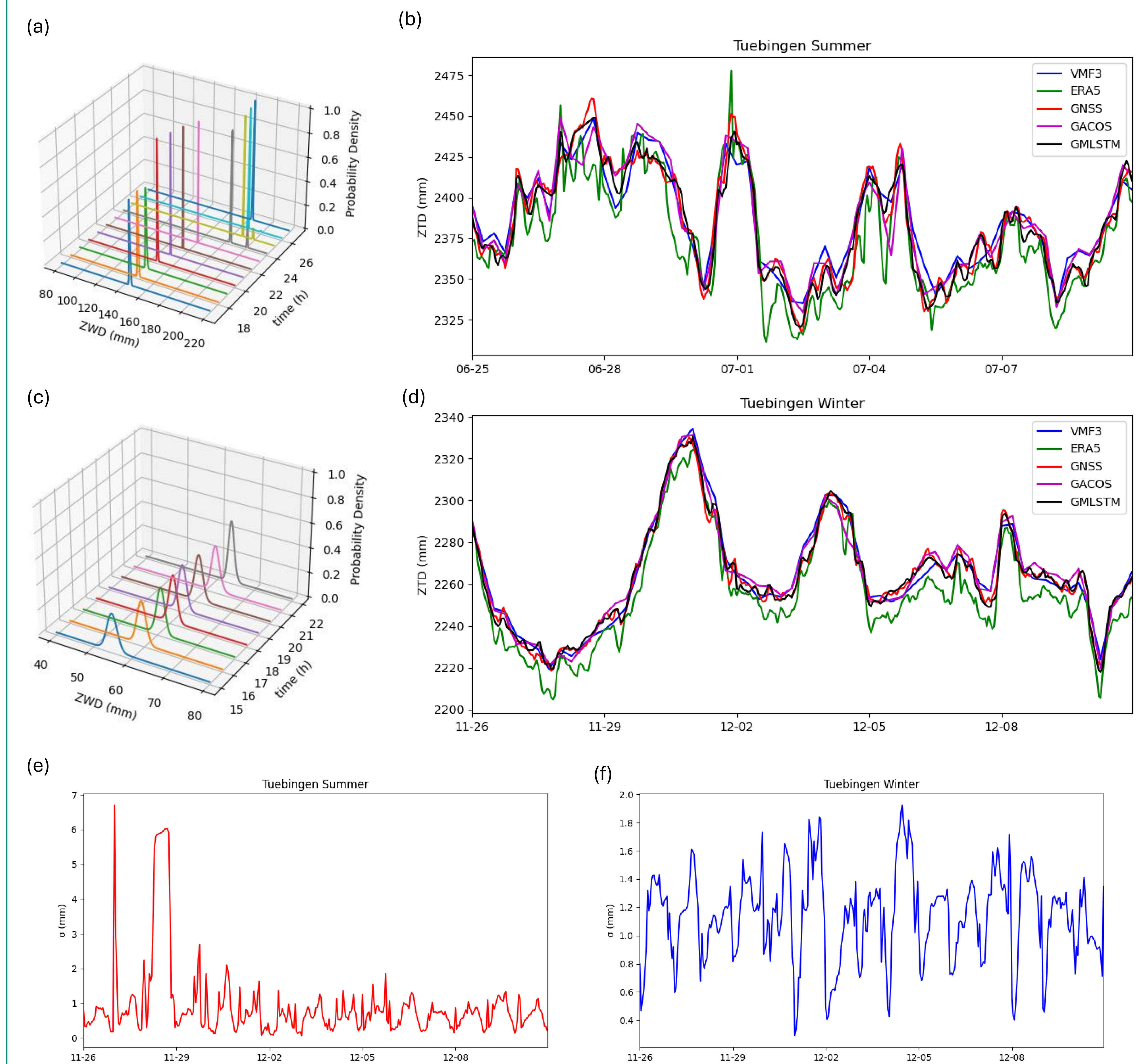


Fig.3 (a) The Gaussian Mixture Distributions from 2022.06.25 17:00 to 2022.06.26 02:00 . (b) The comparison of ZWD retrieved by VMF3, ERA5, GNSS, GACOS and GM-LSTM at 2022-06-25 to 2022-07-09, D400, Tübingen. (c) The Gaussian Mixture Distributions from 2021.11.26 15:00-22:00 . (d) The comparison of ZWD retrieved by VMF3, ERA5, GNSS, GACOS and GM-LSTM at 2021-11-26 to 2021-12-10, D400, Tübingen. (e) The variance, representing the uncertainty caused by water vapor activity, at 2022-06-25 to 2022-07-09, D400, Tübingen. (f) The variance, representing the uncertainty caused by water vapor activity, at 2021-11-26 to 2021-12-10, D400, Tübingen.

SEASON	GM-LSTM		ERA5		VMF3		GACOS	
	RMSE (mm)	MBE (mm)	RMSE (mm)	MBE (mm)	RMSE (mm)	MBE (mm)	RMSE (mm)	MBE (mm)
WINTER	2.37	0.06	9.22	-7.80	4.66	1.05	4.54	1.52
SUMMER	7.21	-0.93	16.43	-9.66	10.45	2.08	10.12	2.27

Conclusion & Outlook

- The GM-LSTM model demonstrates proficient capability in capturing ZWD time series features from GNSS tropospheric products. As a result, it achieves a state-of-the-art level of ZTD and ZWD estimation. In the absence of prior knowledge, modeling the ZWD as a probability distribution—rather than a deterministic single value—can enhance both the robustness and accuracy of the model. Furthermore, this approach effectively captures the inherent uncertainty associated with water vapor activity.
- In our future work, a processing chain for InSAR atmosphere correction by using GM-LSTM will be developed. It has the potential to reveal periodic complex deformations obscured by atmospheric delays. Meanwhile, a high-resolution water vapor field product will also be developed through using GM-LSTM, which will provide a data source for applications such as atmospheric water circulation, evapotranspiration and irrigation.