https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

HiMIC-Daily: A high-resolution (daily and 1 km) multi-indicator atmospheric moisture collection over China, 2003–2020

Zhiying Su[1], Hui Zhang[2], Sijia Wu[1], Zhaoliang Zeng[3], Tao Zhang[4], Ming Luo[1*]

1. Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510006, China.

2. Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China.

3. State Key Laboratory of Severe Weather Meteorological Science and Technology & Institute of Artificial Intelligence for Meteorology, Chinese Academy of Meteorological Sciences, Beijing 100081, China.

4. School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China.

Correspondence: luom38@mail.sysu.edu.cn.

Abstract

Near-surface atmospheric moisture is a fundamental component of the hydrological cycle and plays a key role in regulating land-atmosphere exchanges and surface energy partitioning. Reliable daily high-resolution moisture data are essential for regional climate analysis and fine-scale applications, particularly for capturing short-term variability and extreme moisture dynamics. With complex terrain and a dense population, China is highly vulnerable to extreme hydro-meteorological extremes, yet existing moisture products over China are largely constrained by coarse temporal resolution, insufficient spatial detail, and limited indicators. Here, we present HiMIC-Daily, a seamless daily 1-km-resolution near-surface atmospheric moisture dataset for China, 2003–2020. HiMIC-Daily provides a comprehensive suite of six widely used indicators that characterize atmospheric moisture from different perspectives: actual vapor pressure (AVP), dew point temperature (DPT), mixing ratio (MR), relative humidity (RH), specific humidity (SH), and vapor pressure deficit (VPD). This dataset is generated using the Light Gradient Boosting Machine (LightGBM) framework,

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

which integrates in-situ observations from 2419 meteorological stations with multiple environmental and temporal covariates, including ERA5-Land derived near-surface temperature and DPT, AVP, land surface temperature, topography, and day of year. Validation against observations shows that HiMIC-Daily achieves robust performance across all six indicators, with R^2 values ranging from 0.877 to 0.989. The strongest performance is obtained for AVP, DPT, MR, and SH, with R^2 values exceeding 0.985, and error metrics remain within acceptable ranges for all indicators (e.g., mean absolute error of 0.677 hPa and a root mean square error of 0.933 hPa for AVP). Compared with two existing coarse resolution products, HiMIC-Daily provides finer spatial detail, higher accuracy, and more realistic temporal variability across different climatic regions. These capabilities support spatially explicit studies of climate variability and environmental processes. The HiMIC-Daily dataset is publicly available at https://doi.org/10.11888/Atmos.tpdc.303449.

1. Introduction

Near-surface atmospheric moisture is a key component of the atmospheric branch of the hydrological cycle, directly regulating the exchange of heat and energy between the surface of the Earth and the lower atmosphere (Baker et al., 2015; Sherwood et al., 2018; Angert et al., 2008), and exerts profound effects on hydrological processes, weather systems, ecosystems, agriculture, and human living environments (Bengtsson, 2010; Wood et al., 2019; Harpold and Brooks, 2018; Yuan et al., 2019; Grossiord et al., 2020). Variations in near-surface atmospheric moisture can modulate the spatial and temporal patterns of precipitation and droughts (Laskar et al., 2014; Fujita and Sato, 2017; Angert et al., 2008), regulate the initiation and evolution of convective weather systems (Wang, 2003; Tomassini, 2020), and alter evapotranspiration and soil moisture conditions across heterogeneous landscapes (Lengfeld and Ament, 2012; Rosolem et al., 2013). Atmospheric moisture also interacts with temperature and wind to regulate human thermal comfort, heat stress, and building energy consumption (Mekis et al., 2015; Davis et al., 2016; Cui et al., 2024; Fischer and Knutti, 2012). Therefore,

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

accurately quantifying and characterizing near-surface atmospheric moisture is essential for understanding climate change, diagnosing land-atmosphere feedbacks, assessing environmental and public health risks, and supporting climate adaptation strategies.

The spatial and temporal variability of atmospheric moisture and its associated impacts has attracted growing attention in recent years (Zhou et al., 2026; Wood et al., 2019; Wang et al., 2021; Sobolewski et al., 2021; Chia and Lim, 2022; Jain et al., 2021). Existing studies largely rely on meteorological station observations or coarse-resolution gridded datasets. For example, Wood et al. (2019) explored moisture variability using station networks in complex mountainous regions. Sobolewski et al. (2021) assessed human thermal adaptation under hot and humid conditions based on meteorological observations. In addition, Jain et al. (2021) further investigated global trends in extreme fire weather using the ERA5 reanalysis dataset at a spatial resolution of 0.25°×0.25°. Although these data sources have substantially advanced our understanding of atmospheric moisture variability and its environmental impacts, they remain insufficient for resolving fine-scale daily moisture dynamics. Meteorological stations provide direct observations but are often sparsely and unevenly distributed, particularly in mountainous regions, limiting their ability to capture spatially continuous moisture patterns. (Peng et al., 2019). Station measurements may also be influenced by local topography, land cover, and surface conditions, introducing uncertainties in spatially explicit analyses (Wood et al., 2019). In contrast, gridded reanalysis products provide spatially complete moisture fields, but their coarse spatial resolutions tend to smooth local variability associated with terrain, surface heterogeneity, and regional hydroclimatic contrasts. These limitations hinder their applicability in studies such as urban climate analysis, heat-humidity risk assessments, and wet-dry abrupt transitions, which require high-resolution daily moisture information.

Although a variety of atmospheric moisture datasets have been developed in recent decades, most of them remain limited in spatial resolution, temporal resolution, and indicator coverage (Table 1). For example, ERA5 (Hersbach et al., 2020) provides global hourly moisture variables at a spatial resolution of 0.25° × 0.25° from 1940

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

onward. Other widely used global reanalysis products, such as NCEP/NCAR (Kalnay and Kanamitsu, 2018) and MERRA-2 (Gelaro et al., 2017) , provide moisture indicators at spatial resolutions of 1.875° × 1.875° and 0.5° × 0.625°, respectively. Regional datasets have also been developed for China. For instance, the China Meteorological Forcing Dataset (CMFD) (He Jie, 2024) provides 3-hourly moisture variables at a 0.1° × 0.1° resolution since 1951, while the China Daily Meteorological Dataset (CDMet) (Zhang et al., 2024b) offers daily RH at a spatial resolution of 4 km during 2000–2020. However, these existing moisture datasets are limited by coarse spatial resolution or short temporal coverage. There is an urgent need to develop a daily near-surface atmospheric moisture dataset to support fine-scale climate and environmental studies.

Atmospheric moisture can be characterized in both relative and absolute terms. Commonly used indicators include relative humidity (RH), actual vapor pressure (AVP), vapor pressure deficit (VPD), dew point temperature (DPT), mixing ratio (MR), and specific humidity (SH) (Murray, 1996; Buck, 1981; Bolton, 1980; Wallace and Hobbs, 2006). These indicators are physically related but differ in their sensitivity to temperature, pressure, and water vapor content, and are therefore suited to distinct scientific and practical applications (Grossiord et al., 2020; Davis et al., 2016; Gamelin et al., 2022; Jain et al., 2021; Baldwin et al., 2023; Wood et al., 2019). For example, RH has been widely used in human comfort assessment and weather forecasting (Buzan et al., 2015; Guarnieri et al., 2023; Wolkoff and Kjaergaard, 2007). AVP provides a direct measure of atmospheric water vapor content and is essential for water vapor transport and budget analyses (Kimball et al., 1997; Qiu et al., 2021). VPD quantifies atmospheric drying demand and serves as a key indicator of assessing vegetation water stress, drought impacts, and fire risk (Sangines De Carcer et al., 2018; Gamelin et al., 2022; Silvestrini et al., 2011). DPT provides an intuitive measure of air moisture content and has been applied in frost prediction and evaporative cooler design (Lawrence, 2005; Choińska et al., 2017; Snyder and Melo-Abreu, 2005). MR describes the mass of water vapor relative to dry air and has been used for hydrograph separation and identification of groundwater recharge sources (Carrera et al., 2004). SH is a fundamental variable for atmospheric moisture diagnostics and supports assessments

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

of precipitable water and longwave radiation (Rangwala, 2012; Ruckstuhl et al., 2007; Wood et al., 2019). Since no single indicator fully captures all aspects of atmospheric moisture and most existing datasets cover only one or a few indicators, a high-resolution dataset integrating multiple moisture indicators is urgently needed, particularly for regions with complex climates and topography like China.

Changes in atmospheric moisture across China have received increasing attention in recent years (Peng et al., 2024; Chen et al., 2022; Zhou et al., 2026). Peng et al. (2024) noted that increased water vapor may exacerbate the risk of extreme winter cold events, particularly in northern China. Chen et al. (2022) demonstrated that rising atmospheric moisture can intensify heat stress and increase the frequency of extreme compound heat-humidity events, most prominently in eastern China and the Sichuan Basin. Beyond characterizing mean moisture conditions, daily high-resolution moisture information is also essential for abrupt wet-dry transitions and urban moisture variability. Therefore, accurate, fine-scale atmospheric moisture data are essential for advancing understanding of regional climate extremes, assessing public health risks, and maintaining regional ecological and environmental stability. However, high-resolution atmospheric moisture datasets that integrate multiple moisture indicators remain scarce over China.

To address this gap, we develop HiMIC-Daily—a daily 1 km multi-indicator atmospheric moisture dataset for China covering 2003–2020. HiMIC-Daily is generated using the LightGBM algorithm by integrating meteorological station observations with multiple high-resolution environmental and temporal covariates, including land surface temperature (MOD11A1/MYD11A1, Zhang et al. (2022)), land cover (MCD12Q1, Friedl and Sulla-Menashe (2019)), topographic variables (MERIT DEM, Yamazaki et al. (2017)), column water vapor (MCD19A2, Lyapustin and Wang (2022)), vapor pressure (TerraClimate, Abatzoglou et al. (2018)), 2-m near-surface air temperature and dewpoint temperature (ERA5-Land, Muñoz-Sabater et al. (2021)), and population density (WorldPop, Gaughan et al. (2013)). LightGBM is selected for its higher efficiency and robustness compared to other tree-based machine-learning algorithms, such as XGBoost and CatBoost (Guo et al., 2023; Daoud, 2019; Su et al.,

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

2021). By providing spatially continuous daily estimates of six atmospheric moisture indicators, HiMIC-Daily offers a reliable data foundation for research in meteorology, ecology, environmental science, and human health. The aims of this study are: (1) to construct a seamless daily high-resolution (~1 km) multi-indicator atmospheric moisture dataset over China (HiMIC-Daily) during 2003–2020, covering six indicators (RH, AVP, VPD, DPT, MR, and SH); (2) to evaluate the accuracy and spatial variation of HiMIC-Daily through validation against in situ measurements; (3) to interpret the dataset and validate its reliability across geographical regions and temporal periods. The remainder of this paper is structured as follows. Section 2 describes the data sources and the methodology. Section 3 presents a comprehensive analysis of the accuracy and spatial variation of the atmospheric moisture indicators. Section 4 discusses covariate importance, dataset comparison, the temporal dynamics of HiMIC-Daily, and potential limitations. Data availability is given in Sect. 5, and the main findings are summarized in Sect. 6.

Table 1. Existing datasets of near-surface atmospheric moisture.

Datasets (Ref.)	Coverage	Spatial resolution	Temporal resolution	Temporal coverage	Variables
GSOD (Huang et al., 2017)	Global	Site	Daily	1929–present	DPT
ERA5 (Hersbach et al., 2020)	Global	0.25° × 0.25°	Hourly	1940–present	DPT
ERA5-Land (Muñoz-Sabater et al., 2021)	Global	0.1° × 0.1°	Hourly	1950–present	DPT
NCEP/NCAR (Kalnay and Kanamitsu, 2018)	Global	1.875° × 1.875°	6-hour	1948–present	RH, SH
MERRA-2 (Gelaro et al., 2017)	Global	0.625° × 0.5°	Hourly	1980–present	SH,
JRA-3Q (Kosaka et al., 2024)	Global	0.375° × 0.375°	3-hour	1947–present	RH, DPT, SH
HadISDH (Willett et al., 2014)	Global	5° × 5°	Monthly	1973–2024	RH, DPT, SH
GLDAS (Rodell et al., 2004)	Global	0.25° × 0.25°	3-hour	2000–present	SH
NLDAS-2 (Xia et al.,	North	0.125° × 0.125°	Hourly	1979–present	SH

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

2012)	America				
CMFD (He Jie, 2024)	China	0.1° × 0.1°	3-hour	1951–2024	RH, SH
CDMet (Zhang et al., 2024b)	China	4km × 4km	Daily	2000–2020	RH

2. Data and methods

We develop the high-resolution atmospheric moisture dataset by integrating meteorological station observations and multiple remote sensing and reanalysis products. The flowchart for dataset construction based on the LightGBM model is illustrated in Fig. 1, which consists of three major procedures.

First, all gridded datasets are unified at a 1 km spatial resolution and a consistent coordinate reference system (i.e., WGS84), and spatiotemporally matched across China. We then extract grid values at meteorological station locations to construct candidate samples, which contain both observed moisture variables and corresponding spatially consistent environmental covariates. Second, these daily samples are randomly split into training and validation subsets, which account for 80% and 20% of the samples, respectively. Annual LightGBM models (one for each moisture indicator in a year) with optimized parameters are trained separately for each moisture indicator, and model performance is assessed using the validation subset. Third, the trained annual models are applied spatially across China to generate daily estimates of six daily near-surface atmospheric moisture indicators for 2003–2020.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

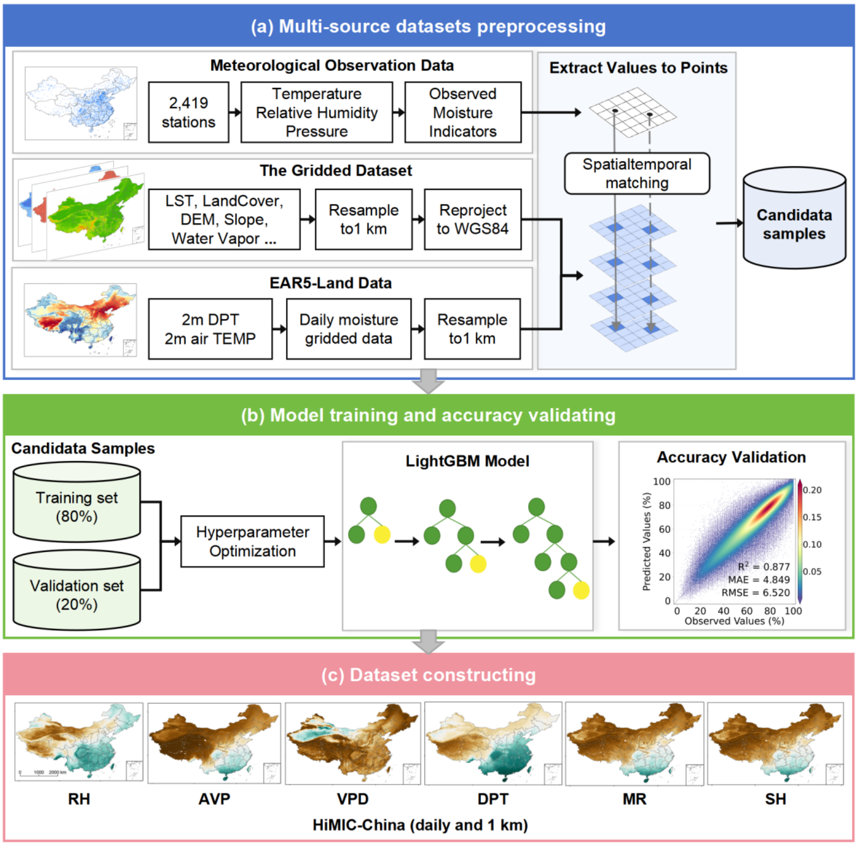

Figure 1. Flowchart of the HiMIC-Daily dataset construction. LST: Land surface temperature; TEMP: Temperature; LightGBM: Light Gradient Boosting Machine.

2.1 Data and preprocessing

2.1.1 Station Observations

Daily meteorological data from 2419 stations across mainland China (Fig. 2) during 2003–2020 are collected from the China Meteorological Data Service Center (CMDSC, http://data.cma.cn/) of the China Meteorological Administration (CMA). The station dataset includes daily mean near-surface air temperature (SAT), RH, and surface pressure (PRS). All raw observations are strictly quality-controlled and verified by the CMA. In line with CMA's data usage terms, the station data used in this research cannot be redistributed, and direct access to the data requires a formal application via CMA's official platform.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

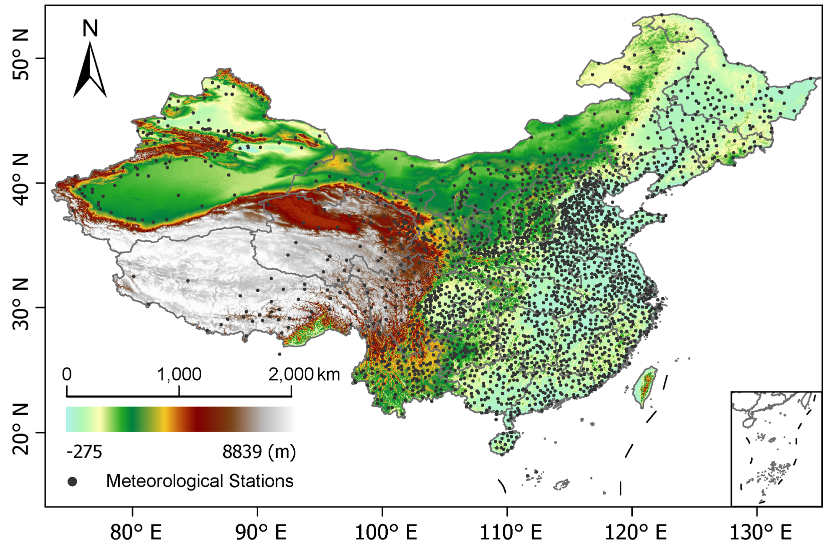

Figure 2. Spatial distribution of meteorological stations (dots) and elevation (shading) in China.

2.1.2 Covariates

The spatiotemporal variability of near-surface atmospheric moisture is jointly influenced by land surface properties, topography, meteorological conditions, human activities, and temporal factors (Taheri et al., 2022; Ige et al., 2017; Antonoglou et al., 2022; Dyer, 2012; Sterling et al., 2012). Accordingly, we select a set of environmental and temporal covariates to support the construction of atmospheric moisture, including land surface temperature (LST), DPT and SAT from ERA5-Land, landcover, elevation and slope, column water vapor, vapor pressure, population density, day of year, month of year, and day of month (Table 2).

LST is closely linked to surface water status and surface energy balance, and therefore serves as a critical boundary condition for land-atmosphere interactions, making it a valuable indicator for predicting atmospheric moisture (Taheri et al., 2022; Chen and Liu, 2020). As LST reflects thermal conditions associated with evapotranspiration, energy partitioning, and water vapor transport, it has been widely utilized in remote sensing-based estimation studies to characterize the spatiotemporal dynamics of atmospheric moisture (Kalma et al., 2008; Duffour et al., 2016; Zhang et

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

al., 2024a). In this study, we use a daily 1 km × 1 km LST dataset derived from the MODIS LST products (MOD11A1 and MYD11A1), covering both daytime and nighttime observations (Zhang et al., 2022). Spatiotemporal gap-filling and quality control procedures have been applied to improve data completeness and reliability (Zhang et al., 2022). Comparisons with observations show good agreement, with root mean squared errors (RMSE) of 1.80–1.88°C and 1.23–1.45°C for daytime and nighttime LSTs, respectively (Zhang et al., 2022). Daily mean LST was calculated by averaging the daytime and nighttime values to represent daily surface thermal conditions. The 2-m DPT and SAT are obtained from the ERA5-Land dataset (Muñoz-Sabater et al., 2021) and are used to generate daily gridded data for six near-surface moisture indicators. The detailed calculation procedures are described in Section 2.2.1. An "index-matching" strategy is adopted during model development, in which the ERA5-Land derived RH is used as a covariate when predicting the corresponding 1-km RH.

Land cover is included to represent surface characteristics that regulate water and energy exchanges between the land surface and the lower atmosphere (Kalma et al., 2008; Ige et al., 2017). We derive land cover information from the MODIS Land Cover Type product (MCD12Q1.006; Friedl and Sulla-Menashe (2019)). Topography can also influence atmospheric conditions, particularly in regions with complex terrain (Antonoglou et al., 2022; Castino et al., 2020). Elevation data are thus fetched from the Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT DEM; Yamazaki et al. (2017)), from which slope was subsequently derived. To characterize atmospheric water availability (Reber and Swope, 1972; Dyer, 2012), column water vapor from the MCD19A2.061 product (Lyapustin and Wang, 2022) and vapor pressure from the TerraClimate dataset (Abatzoglou et al., 2018) are incorporated as moisture-related predictors. Population density from the WorldPop dataset (Gaughan et al., 2013) is included to account for potential anthropogenic influences on near-surface moisture conditions (Liu et al., 2008; Sterling et al., 2012). In addition, temporal variables including day of year, month of year, and day of month are incorporated to capture the seasonal and interannual variability in atmospheric moisture. All covariate

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

datasets are accessed through Google Earth Engine (GEE). They are then harmonized to a consistent spatial extent, coordinate reference system (WGS84), and spatial resolution (1 km) using mosaicking, reprojection, resampling, clipping, and aggregation.

Table 2. Gridded datasets and covariates used in this study.

Category	Spatial resolution	Temporal resolution	DOI
Land surface temperature	1 km	daily	https://doi.org/10.5067/MODIS/MYD11A1.061
Landcover	500 m	yearly	https://doi.org/10.5067/MODIS/MCD12Q1.061
Elevation and slope	3arc-second	/	https://doi.org/10.1002/2017GL072874
Column water vapor	1 km	daily	https://doi.org/10.5067/MODIS/MCD19A2.061
Vapor pressure	4 km	monthly	https://doi.org/10.1038/sdata.2017.191
2-m dewpoint temperature	11132 m	daily	https://doi.org/10.24381/cds.e9c9c792
2-m temperature	11132 m	daily	https://doi.org/10.24381/cds.e9c9c792
Population density	100 m	yearly	https://doi.org/10.5258/SOTON/WP00645
Day of year, month of year, day of month	/	/	/

2.2 Methods

2.2.1 Calculation of atmospheric moisture indicators

Six daily near-surface atmospheric moisture indicators including RH (%), AVP (hPa), VPD (hPa), DPT (°C), MR (g kg^{-1}), and SH (g kg^{-1}) are produced in this study. The corresponding calculation procedures are summarized in Table 3. For the meteorological station observations, daily mean SAT (°C), RH, and PRS (hPa) are directly obtained from the station measurements. SVP (hPa) is calculated from SAT using the Magnus-Tetens equation (Eq. 1); AVP is derived from RH and SVP; VPD is

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

calculated as the difference between SVP and AVP; and DPT is estimated from AVP. The pressure-dependent indicators, MR and SH, are further calculated using AVP and PRS observations:

$$SVP = 6.112 \times \exp^{\frac{17.67 \times SAT}{SAT+243.5}} \quad (1)$$

Based on the ERA5-Land dataset, AVP is derived from the 2-m dew point temperature (T_d, °C), while SVP is calculated from the 2-m air temperature (Eq. 1). The resulting AVP and SVP values are subsequently used to derive the six atmospheric moisture indicators (Table 3).

Table 3. Equations of six near-surface atmospheric moisture indicators.

Indicators (Ref.)	Equation	Unit
RH (Murray, 1996)	$RH = 100 \times AVP \div SVP$	%
AVP (Murray, 1996)	$AVP = RH \times SVP \div 100$ $AVP = 6.112 \times \exp^{\frac{17.67 \times T_d}{T_d+243.5}}$	hPa
VPD (Buck, 1981)	$VPD = SVP - AVP$	hPa
DPT (Bolton, 1980)	$DPT = \frac{\log(AVP/6.112) \times 243.5}{17.67 - \log(AVP/6.112)}$	°C
MR (Wallace and Hobbs, 2006)	$MR = \frac{0.62197 \times AVP}{PRS - AVP} \times 1000$	g kg^{-1}
SH (Wallace and Hobbs, 2006)	$SH = \frac{MR}{1+MR} \times 1000$	g kg^{-1}

2.2.2 Prediction of atmospheric moisture indicators

The Light Gradient Boosting Machine (LightGBM) algorithm (Ke et al., 2017) is employed to predict atmospheric moisture indicators. LightGBM is an efficient gradient boosting decision tree (GBDT) framework that has been widely applied meteorological variable simulation, air quality prediction, and hydrological modeling (Li et al., 2023; Tian et al., 2021; Fan et al., 2019; Ju et al., 2019), due to its high computational efficiency and strong predictive performance. Compared with other commonly used tree-based machine learning algorithms, such as eXtreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Random Forest,

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

LightGBM improves training efficiency and predictive capability through several optimized learning strategies (Daoud, 2019; Guo et al., 2023), including a leaf-wise tree growth, Gradient-based One-Side Sampling (GOSS), and Exclusive Feature Bundling (EFB) (Ke et al., 2017). These strategies enable the LightGBM model to capture complex nonlinear relationships between atmospheric moisture and environmental variables while maintaining relatively fast training speed (Su et al., 2021). In addition, LightGBM provides feature importance metrics that can be used to assess the relative contributions of different predictors to the reconstruction of atmospheric moisture fields (Ke et al., 2017).

The LightGBM model is implemented using the Python LightGBM library (https://lightgbm.readthedocs.io). To ensure robust model training and evaluation, the observed atmospheric moisture indicators are randomly divided into training (80%) and validation (20%) subsets. Hyperparameter tuning is performed using a grid search combined with 5-fold cross-validation within the training subset, and the final hyperparameter combination is selected by minimizing RMSE (Sarker, 2021). To account for possible year-to-year variations in the relationships between atmospheric moisture and its predictors, a separate LightGBM model is developed for each moisture indicator and each year. The trained models are then applied to the daily gridded datasets to generate atmospheric moisture predictions (Fig. 1).

2.2.3 Assessment of accuracy

The accuracy of the trained LightGBM models is evaluated by comparing model predictions with the corresponding station-derived moisture indicators in the validation subset. Three standard statistical metrics are used, including the coefficient of determination (R^2), mean absolute error (MAE), and RMSE. These metrics have been widely used in previous studies to assess the predictive performance of environmental and atmospheric variables (Zhou et al., 2026; Zhang et al., 2023; Wu et al., 2023; Peng et al., 2019). R^2 measures the proportion of observed variance explained by the model and reflects the overall goodness of fit. MAE quantifies the average magnitude of

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

prediction errors, providing an intuitive measure of model accuracy. RMSE assigns greater weight to larger errors and is therefore sensitive to extreme deviations, making it useful for evaluating prediction reliability. The three metrics are computed as follows (Eqs. 2–4):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{oi} - \hat{y}_{pi})^2}{\sum_{i=1}^{n}(y_{oi} - \overline{y}_o)^2} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{oi} - \hat{y}_{pi})^2} \tag{3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{oi} - \hat{y}_{pi}\right| \tag{4}$$

where n is the total number of samples, y_{oi} and $\hat{y}_{pi}$ are the i_{th} observed and predicted moisture index values, respectively; and $\overline{y}_o$ is the mean of the observed moisture index values.

3. Results

3.1 Evaluation of atmospheric moisture indicators

3.1.1 Overall accuracy

The overall predictive accuracy of HiMIC-Daily is evaluated using the validation subset, in which model predictions are compared with the corresponding station-derived moisture indicators. All six indicators achieve high accuracy, with R^2 values exceeding 0.877 (Table 4). In particular, AVP, DPT, MR, and SH exhibit R^2 values greater than 0.985. As shown in Fig. 3, the predicted and observed atmospheric moisture values are closely distributed around the 1:1 line for all indicators, demonstrating strong agreement between the LightGBM model predictions and station measurements. MAE and RMSE for each indicator are within acceptable ranges, with RH recording an MAE of 4.849% and an RMSE of 6.520%, followed by AVP with an MAE of 0.677 hPa and an RMSE of 0.933 hPa, VPD with an MAE of 0.983 hPa and an RMSE of 1.411 hPa, DPT with an MAE of 1.120°C and an RMSE of 1.608°C, MR with an MAE of 0.459 g kg^{-1} and an RMSE of 0.633 g kg^{-1}, and SH with an MAE of

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

0.451 g kg^{-1} and an RMSE of 0.620 g kg^{-1} (Table 4). These evaluation metrics collectively confirm the model's reliability in estimating atmospheric moisture conditions across China.

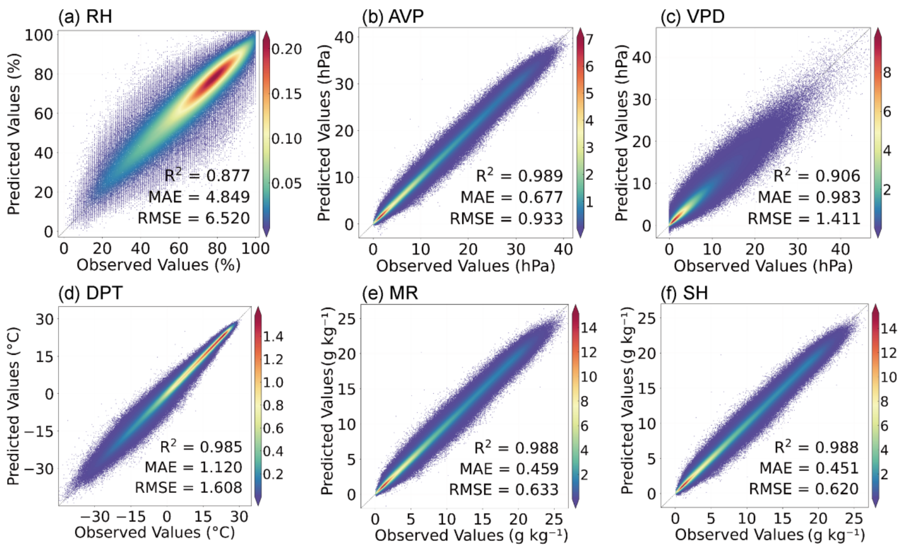

Figure 3. Scatter plots of predicted and observed atmospheric moisture indicators over China during 2003–2020: **(a)** RH, **(b)** AVP, **(c)** VPD, **(d)** DPT, **(e)** MR, and **(f)** SH. Colors indicate the density of data points, with red representing higher density and blue representing lower density. The black dashed line is the 1:1 line.

3.1.2 Validation on yearly, monthly, and daily scales

We further evaluate the predictive performance of the LightGBM model at annual, monthly, and daily scales (Fig. 4). At the annual scale, all six moisture indicators maintain consistently high prediction accuracy throughout all years from 2003 to 2020. AVP, DPT, MR, and SH exhibit particularly strong performance, with annual R^2 values consistently exceeding 0.980 across all years. RH and VPD also demonstrate reliable accuracy, with annual R^2 values ranging from 0.850 to 0.900 for RH and from 0.890 to 0.910 for VPD (Fig. 4a). The annual MAE and RMSE values for all indicators remained relatively stable throughout the study period (Fig. 4d, g). Slight interannual variations are observed, with relatively higher accuracy in 2017 for most indicators and lower accuracy in 2004 for AVP, DPT, MR, and SH, and in 2018 and 2020 for VPD.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

At the monthly scale, prediction accuracy shows clear seasonal variations (Fig. 4b, e, and h). AVP, DPT, MR, and SH maintain robust performance throughout the year, with monthly R^2 values consistently above 0.950 and reaching their highest levels in autumn and winter. In contrast, RH and VPD exhibit more pronounced seasonal variations. RH reaches its highest accuracy in summer, with an R^2 of 0.910 in June, and declines to its lowest in winter. VPD shows a similar pattern, with values increasing from winter to summer. The monthly error metrics further confirm these seasonal patterns. For RH and DPT, MAE and RMSE values are substantially lower in summer than in winter, whereas AVP, VPD, MR, and SH show the opposite seasonal pattern, with lower errors in winter and higher in summer. These seasonal variations in prediction accuracy suggest seasonal variability in climatic conditions and their influence on atmospheric moisture dynamics across China.

At the daily scale, the LightGBM model maintains robust predictive performance across all six atmospheric moisture indicators (Fig. 4c, f and i). AVP, DPT, MR, and SH show consistently high accuracy, with daily R^2 values exceeding 0.977 and minimal day-to-day fluctuations. RH reaches its highest accuracy around day 30 and its lowest around days 26–27. VPD performs best in early-month to mid-month. The daily MAE and RMSE values are consistent with the R^2 results, showing lower prediction errors for AVP, DPT, MR, and SH, and slightly larger variability for RH and VPD. These results further demonstrate the model's capability to capture atmospheric moisture variations at daily resolution.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

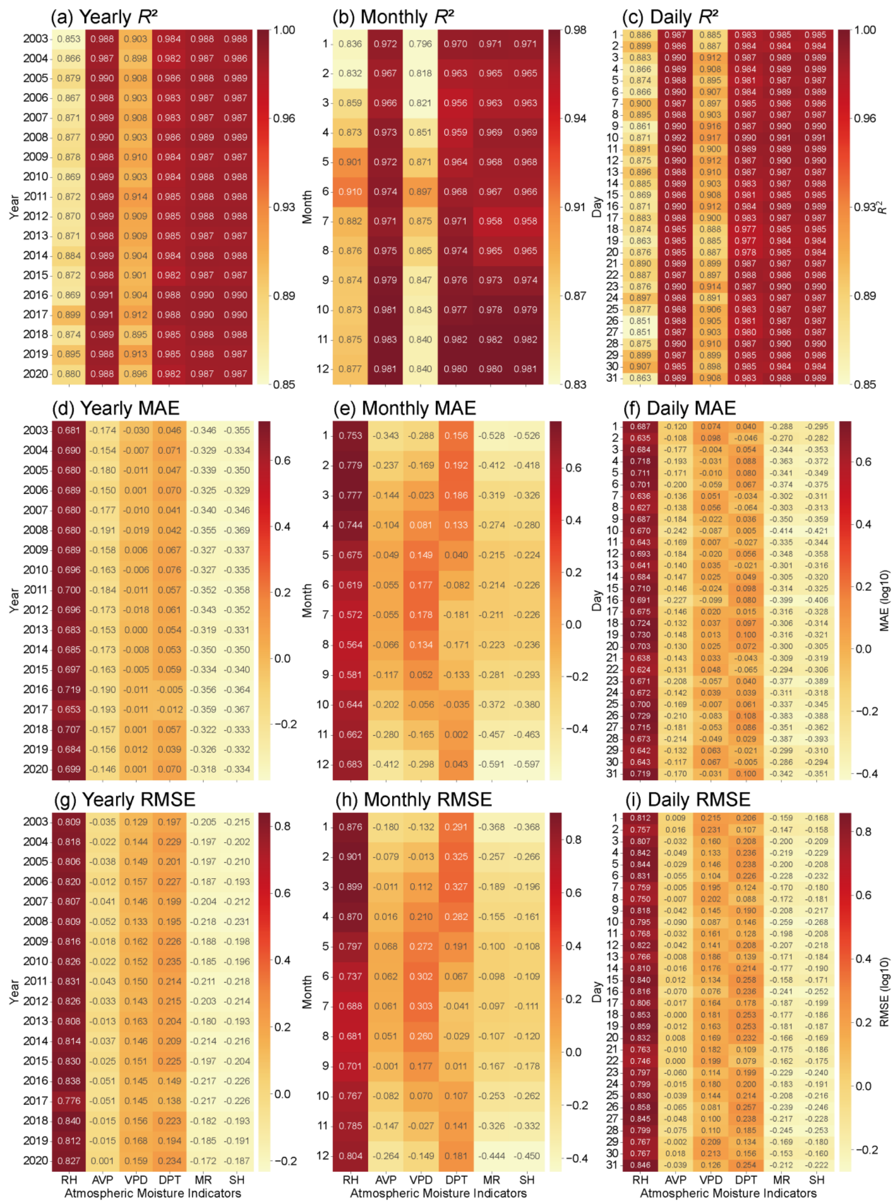

Figure 4. Annual, monthly, and daily R^2, MAE, and RMSE of six atmospheric moisture indicators over China during 2003–2020. Colors represent values, with red for higher values and yellow for lower. MAE and RMSE values are shown on a logarithmic scale, i.e., log (10).

3.1.3 Spatial variation in prediction accuracy

To further assess the predictive performance of the LightGBM model at the spatial scale, we analyze the spatial distributions of R^2, MAE, and RMSE for the six moisture

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

indicators across individual stations (Fig. 5). The results show that the model reproduces the moisture variability well at most stations, although the accuracy varies among indicators and regions. Regarding the spatial distribution of R^2, all indicators demonstrate strong explanatory power across most stations (Fig. 5a–f). Specifically, AVP, DPT, MR, and SH exhibit a similar spatial pattern, with R^2 values exceeding 0.900 at most stations. Higher prediction accuracy is concentrated in northern China, the North China Plain (NCP) and the Yangtze River Plain, whereas relatively lower accuracy appears in parts of southwestern China. In contrast, RH and VPD exhibit high R^2 values in the NCP, with lower values in southwestern China. In terms of error metrics, MAE and RMSE exhibit consistent spatial distributions within each indicator (Figure 5g–r). RH and DPT exhibit similar spatial patterns, with lower errors concentrated south of the Yangtze River Plain and higher errors in northern China. In contrast, AVP, VPD, MR, and SH display the opposite pattern, characterized by lower errors in northern China and relatively higher errors in southern China.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

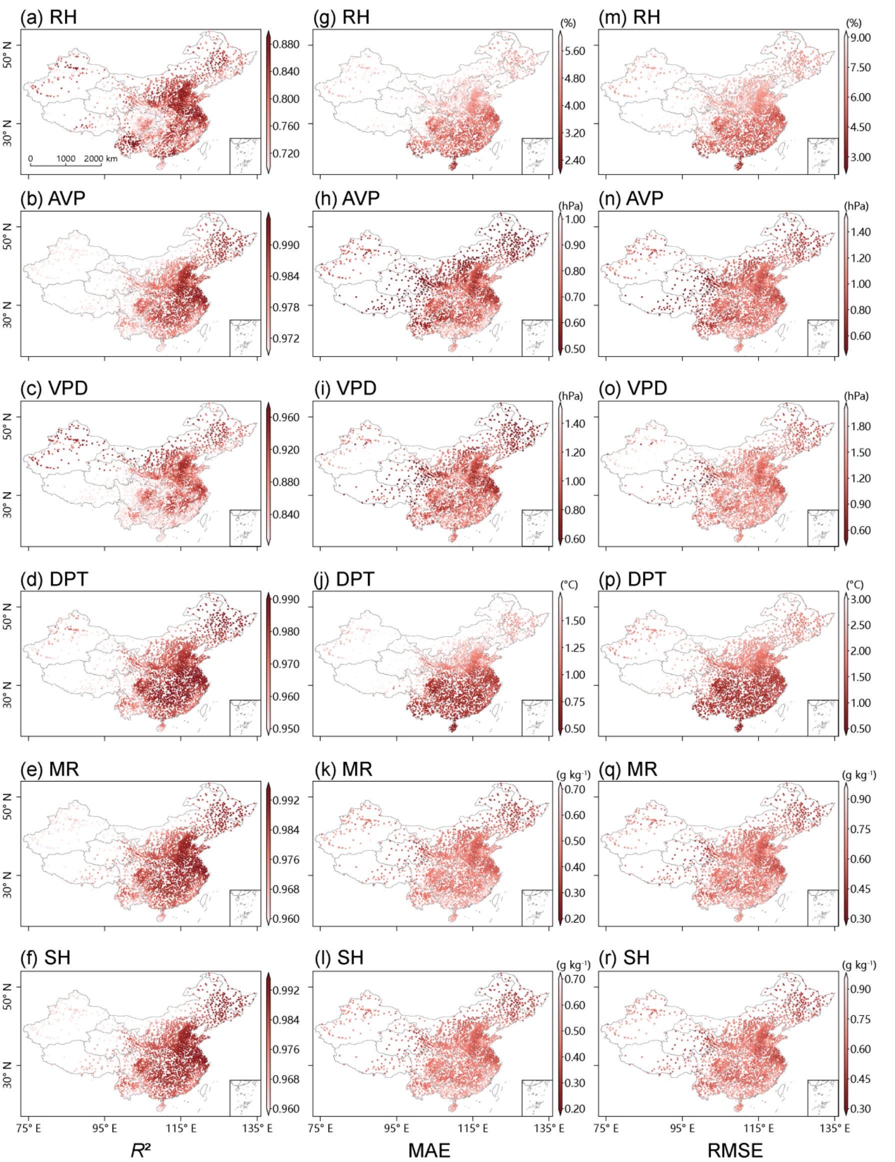

Figure 5. Spatial distribution of R^2, MAE, and RMSE values for six atmospheric moisture indicators at meteorological stations over China during 2003–2020: Each row corresponds to one moisture indicator (from top to bottom: RH, AVP, VPD, DPT, MR, and SH), while **(a-f)**, **(g-l)**, and **(m-r)** correspond to R^2, MAE, and RMSE, respectively. Deeper red indicates larger values, reflecting better model performance.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

3.2 Spatial variations of the atmospheric moisture indicators

The validation results above demonstrate that the LightGBM model provides reliable predictions of near-surface atmospheric moisture across multiple spatial and temporal scales. Based on this robust performance, a high-resolution atmospheric moisture dataset for China is generated at a daily and 1-km resolution from 2003 to 2020. To illustrate the spatial characteristics and event-scale applicability of the dataset, we examine the 2003–2020 climatological mean conditions and the spatial patterns on two representative extreme days: an extreme cold event on 3 January 2008 (Zhou et al. (2011)) and an extreme heat event on 13 August 2013 (Sun et al. (2014)).

The long-term mean shows consistent spatial patterns across five atmospheric moisture indicators excluding VPD (Fig. 6). RH is highest in southeastern China and decreases gradually toward northwestern China, reflecting the transition from humid to arid climates and the influence of complex terrain (Fig. 6a–f). VPD shows an opposite spatial pattern, with high values in arid inland basins, such as the Turpan and Tarim Basins, and low values in humid and high-altitude regions. AVP, DPT, MR, and SH show similar spatial structures, with higher values over warm and humid low-latitude regions, such as the Pearl River Delta, and lower values over the Tibetan Plateau and regions north of the Qinling Mountains-Huaihe River line. These patterns reflect the combined effects of temperature, moisture availability, elevation, and large-scale climatic conditions on near-surface atmospheric moisture.

The two representative extreme days further demonstrate HiMIC-Daily's ability to capture event-scale spatial variability (Fig. 6g–r). During the cold event on 3 January 2008, most moisture indicators show lower values across large parts of China, with particularly evident reductions in DPT (Fig. 6g–l). During the 2013 heat event on 13 August, AVP, MR, and SH increased markedly across eastern China (Fig. 6m–r), indicating enhanced atmospheric water vapor content under hot and humid conditions. Overall, the HiMIC-Daily dataset reveals fine-scale spatial variability that is difficult to detect using coarser products. The enhanced spatial detail advances our

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

understanding of hydrological cycles, atmospheric dynamics, ecological systems, agricultural activities, and human settlements.

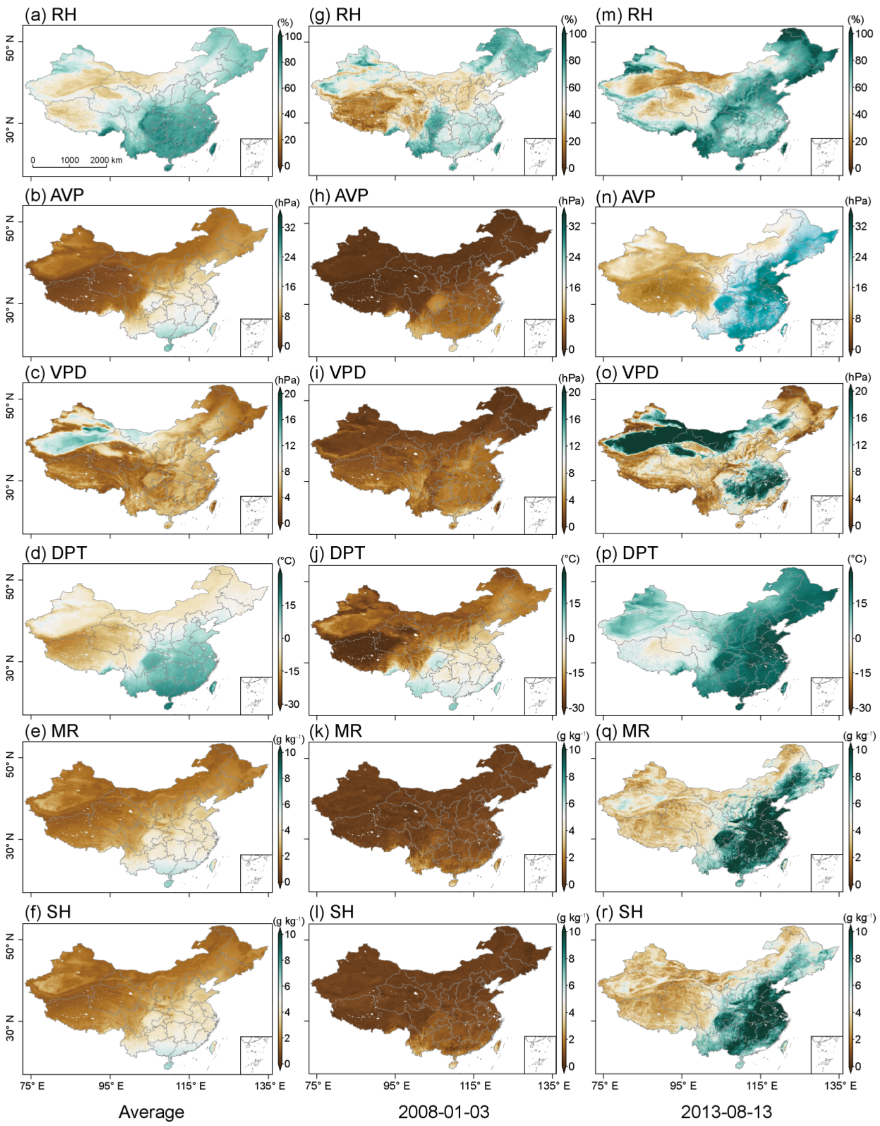

Figure 6. Spatial patterns of six moisture indicators over China during 2003–2020: Each row corresponds to one moisture index (from top to bottom: RH, AVP, VPD, DPT, MR, and SH). **(a-f)** annual mean values during 2003–2020, **(g-l)** distributions on 3 January, 2008, and **(m-r)** distributions on 13 August, 2013.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

4. Discussion

4.1 Covariate importance

To examine the relative contribution of individual covariates to atmospheric moisture prediction, feature importance is quantified for the six moisture indicator models based on the mean information gain derived from LightGBM ensemble trees. Across all predictor variables, moisture indicators derived from ERA5-Land emerge as the most influential predictors (Fig. 7). This result is expected because these predictors are physically closest to the moisture indicators and retain large-scale atmospheric moisture variability derived from ERA5-Land 2-m SAT and DPT (Muñoz-Sabater et al., 2021). Vapor pressure serves as the second most influential feature in most models except for VPD, given that it directly represents the actual water vapor content in the atmosphere and underpins the calculation of most moisture indicators (Dyer, 2012). LST also contributes to model performance, as variations in LST influence evaporation, surface energy balance, and land-atmosphere moisture exchange (Wang et al., 2017). DEM further improves the model by modulating local moisture conditions via moisture convergence, cold air drainage, and evapotranspiration, thereby enabling effective capture of spatial heterogeneity in moisture (Dobrowski, 2011).

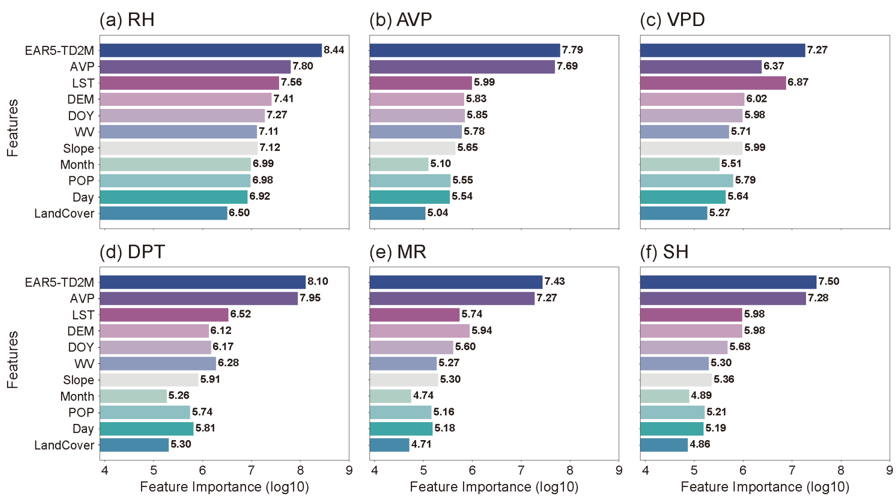

Figure 7. Relative importance of eleven covariates in predicting moisture indicators: **(a)** RH, **(b)** AVP, **(c)** VPD, **(d)** DPT, **(e)** MR, **(f)** SH. The covariates include ERA5-TD2M (near-surface temperature and DPT from ERA5-Land), AVP (water vapor pressure), LST (land surface

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

temperature), DEM (elevation), DOY (day of year), WV (column water vapor), Slope, POP (population density), LandCover, Month of year, and Day of month. The feature importance values are shown on a logarithmic scale, i.e., log (10).

4.2 Comparison with existing atmospheric moisture datasets

We compare our HiMIC-Daily with two existing atmospheric moisture datasets: the China Meteorological Forcing Dataset (CMFD) developed by He et al. (2024), and the 4-km daily gridded meteorological dataset for China (CDMet) produced by Zhang et al. (2024b). CMFD and CDMet have coarser spatial resolutions of 0.1° × 0.1° and 4 km × 4 km, respectively (Table 1). Since RH is available in all three datasets, the comparison focuses on daily RH over China on July 28, 2013, with particular attention to three representative urban agglomerations: Beijing-Tianjin-Hebei (BTH), the middle Yangtze River Valley (mYRV), and the Pearl River Delta (PRD) (Fig. 8). This date is selected because it is available across all three datasets and represents a prolonged 2013 extreme heat event when moisture conditions exhibit substantial variability.

All three datasets capture the broad national-scale RH variability, with lower values in northwestern China and higher values in eastern China (Fig. 8, left). However, clear differences emerge at finer spatial scales. Compared with CMFD and CDMet, HiMIC-Daily can capture more detailed RH variations associated with terrain, land-surface heterogeneity, and urban-rural contrasts, particularly within the BTH, mYRV, and PRD regions. These fine-scale informations are substantially smoothed in the coarser-resolution products. The comparison with station observations further demonstrates the improved performance of HiMIC-Daily (Fig. 8, right panels). CMFD and CDMet show larger deviations from the station measurements, with CMFD underestimating RH across most regions. In contrast, HiMIC-Daily shows closer agreement with observations, indicating that integrating station measurements, reanalysis data, and high-resolution environmental covariates improves the representation of local moisture variability. Our HiMIC-Daily provides finer spatial detail and better station-level consistency than existing coarse-resolution products, supporting its use in fine-scale applications such as intra-urban climate analysis, heat-humidity risk assessment, and localized moisture assessment.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

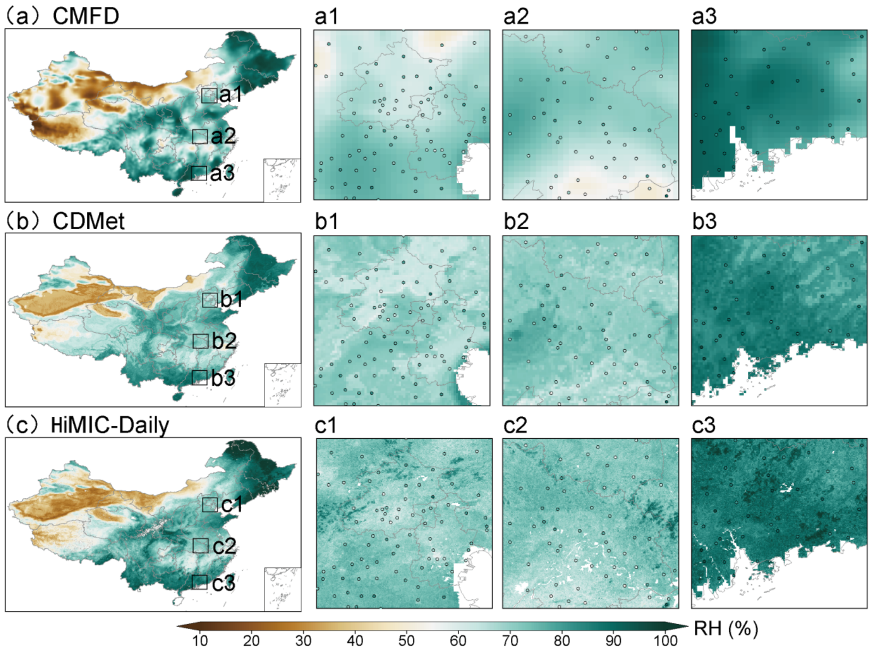

Figure 8. Comparison of spatial patterns among CMFD, CDMet, and HiMIC-Daily for RH over the mainland of China and its three largest urban agglomerations on July 28, 2013 (1: Beijing-Tianjin-Hebei, 2: middle Yangtze River Valley, and 3: Pearl River Delta). Station observations appear as colored circles, shaded according to measured RH (%).

To further evaluate the temporal performance of HiMIC-Daily, we compare daily RH time series from the three datasets with station observations in 2013 (Fig. 9a). HiMIC-Daily shows the closest agreement with the station observations, accurately capturing both the magnitude and day-to-day variability of RH. CMFD captures the seasonal evolution of RH but shows larger deviations in the magnitude, whereas CDMet exhibits clear seasonal biases, with overestimation in spring and underestimation in summer and autumn. A representative station from each urban agglomeration (BTH, mYRV, PRD) is selected for further detailed comparison (Fig. 9b). HiMIC-Daily maintains the best agreement with observed RH, while CMFD closely follows the observed temporal variations but with noticeable biases. CDMet shows larger seasonal deviations again. These results confirm the robustness of HiMIC-Daily in capturing

© Author(s) 2026. CC BY 4.0 License.

daily RH dynamics, highlighting its advantage for applications requiring temporally resolved near-surface moisture information.

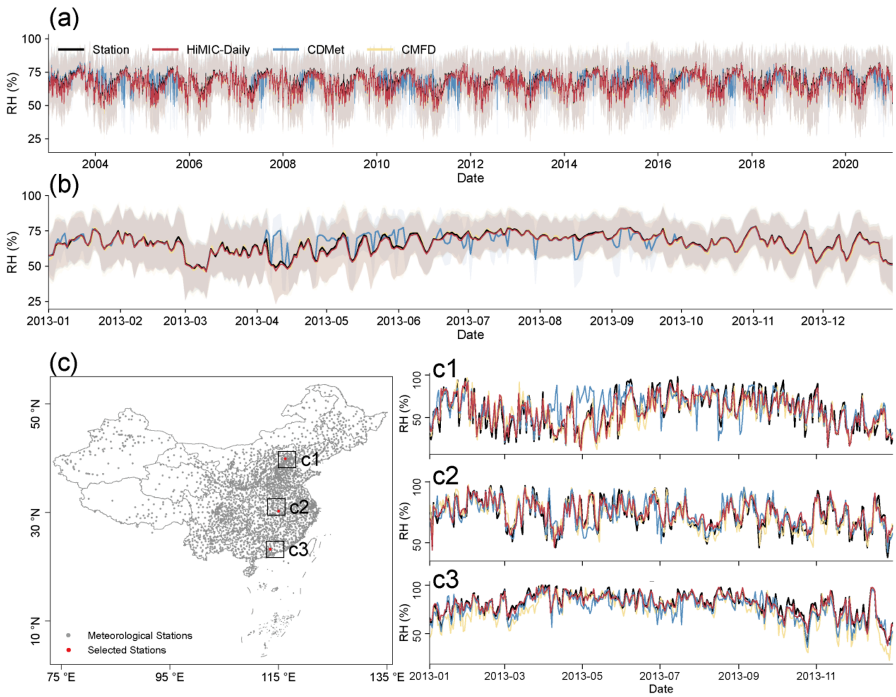

Figure 9. Comparison of daily RH (%) time series among HiMIC-Daily, CMFD, CDMet, and station observations: **(a)** Time series over China from 2003 to 2020, **(b)** Time series over China in 2013, **(c)** Representative station comparison in three major urban agglomerations of China (C1: Beijing-Tianjin-Hebei, C2: middle Yangtze River Valley, and C3: Pearl River Delta).

4.3 Temporal response of HiMIC-Daily to extreme events

We further assess the capability of HiMIC-Daily to represent temporal variations in atmospheric moisture response to extreme events. A representative cold extreme, the 2008 Chinese ice storm, which severely struck the densely populated and economically developed region of south-central China (Zhou et al., 2011). During the 2008 ice storm, Guiyang station is selected as one of the most severely affected cities, where radial ice thickness exceeded 80 mm (Zhou et al., 2011). Air temperature plunged below freezing on January 11 and remained subfreezing throughout the event. This rapid cooling reduced the saturation water vapor capacity of the atmosphere and substantially altered near-surface moisture conditions. HiMIC-Daily captures these abrupt changes well. RH

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

increases sharply within one day, showing close agreement with station observations, while AVP, VPD, DPT, MR, and SH all decrease rapidly and remain at low levels throughout the event. These responses are physically consistent with the combined effects of lower air temperature, reducing saturation vapor pressure, and weakening atmospheric water vapor content under persistent cold conditions.

Pudong New Area in Shanghai is selected to represent the July 2013 extreme heat event, where air temperature began to rise on June 27 and increased by ~10°C within one week (Sun et al., 2014). This rapid warming substantially enhanced the atmospheric demand for water vapor. During this heat event, RH fluctuates and declines, agreeing well with station observations, while AVP, VPD, DPT, MR, and SH all rise. The sharpest increase is seen in VPD, reflecting the strong increase in atmospheric drying demand under rapid warming. These two contrasting extreme events demonstrate that HiMIC-Daily effectively captures temporal moisture dynamics under both cold and heat extremes, showing rapid and physically consistent agreement with station observations for RH. The dataset maintains robust performance without degradation under extreme conditions, confirming its reliability for analysing event-scale moisture dynamics short-term hydroclimatic processes across China.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

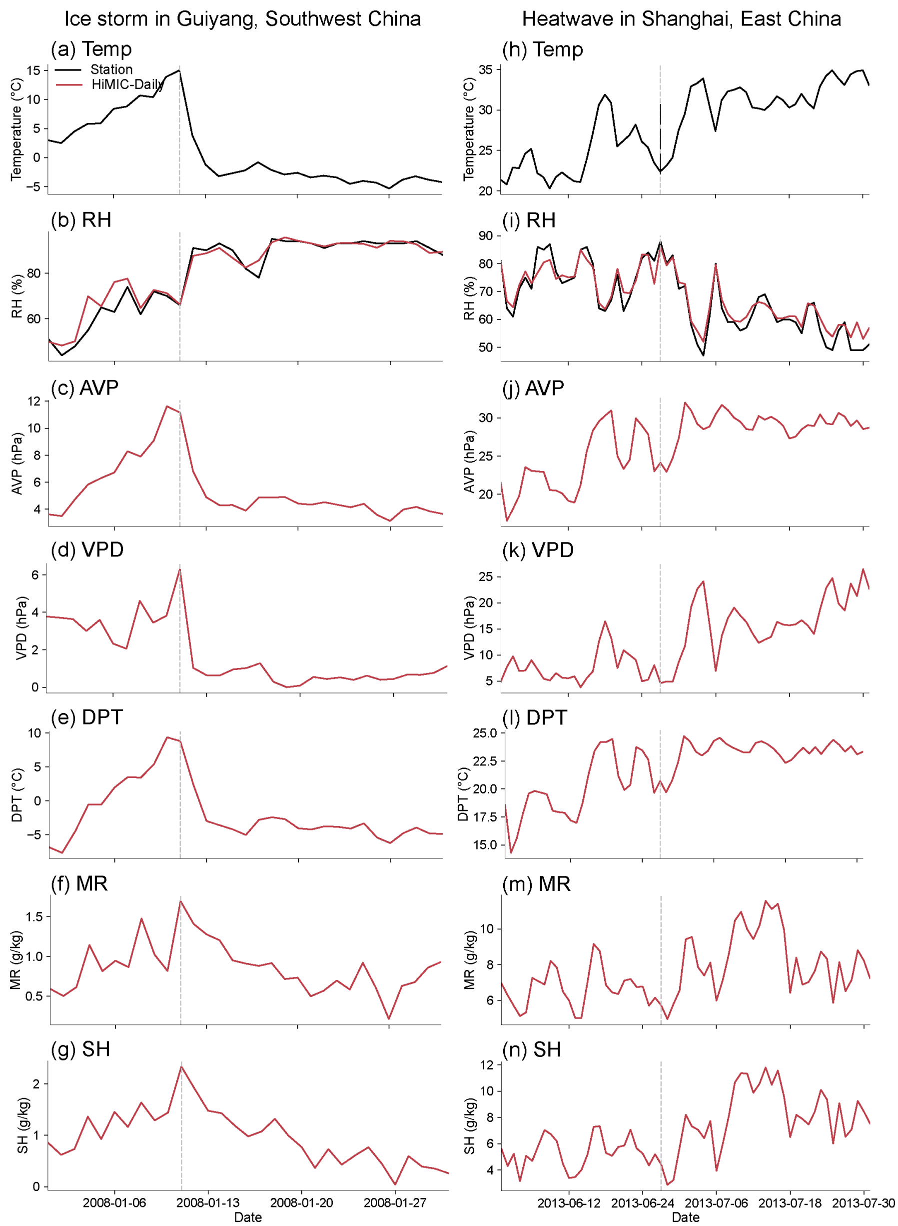

Figure 10. Temporal response of HiMIC-Daily to two extreme events: **(a-g)** the 2008 ice storm at the Guiyang station in Southwest China, **(h-n)** the 2013 summer heatwave in Pudong New Area, Shanghai in East China. Panels from top to bottom in each group show air temperature, RH (%), AVP (hPa), VPD (hPa), DPT (°C), MR (g/kg), and SH (g/kg). Gray dashed vertical lines mark the onset dates of the events.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

4.4 Limitations and future works

HiMIC-Daily offers new capabilities for studying rapidly evolving hydroclimatic processes at fine spatial and temporal scales. In particular, the dataset can advance research on wet-dry abrupt transitions, urban moisture island variability, and rapid moisture changes during heat extremes (He et al., 2025; Huang et al., 2026; Gu et al., 2025). Nevertheless, several limitations should be acknowledged. First, meteorological stations used for model training are unevenly distributed across China, with dense coverage in eastern China but sparse coverage in western mountainous regions such as the Tibetan Plateau. This spatial imbalance may affect model training and validation, particularly in areas with complex terrain and limited ground observations (Sonak and Patankar, 2015; Jafarigol and Trafalis, 2023). In regions with sparse station coverage, the model may rely more largely on remote sensing and reanalysis covariates, which could increase the potential uncertainty of the reconstructed moisture fields.

Second, our model"s performance is partly constrained by uncertainties in the input datasets. For instance, ERA5-Land 2-m SAT and DPT are subject to biases inherited from precipitation forcing and underestimated uncertainties from the land surface model (Muñoz-Sabater et al., 2021). These uncertainties may propagate through the machine-learning framework and affect the final estimates of atmospheric moisture indicators. Similarly, uncertainties in remote sensing products, topographic data, and ancillary covariates may affect local predictions. Future improvements in the quality, spatial resolution, and temporal consistency of input datasets would further enhance the reliability of HiMIC-Daily.

Third, the temporal coverage of HiMIC-Daily is partly constrained by the availability of a temporally consistent land cover product. In this study, MCD12Q1.006 is used as the land cover covariate spanning 2001–2020 (Friedl and Sulla-Menashe, 2019), which limits the current version of the dataset to this period. Future work can incorporate land cover datasets with longer temporal coverage to extend HiMIC-Daily to more recent years and earlier decades. In addition, the framework developed here

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

could be transferred to other regions and expanded toward a global-scale atmospheric moisture product.

5. Data availability

The high-resolution (daily and 1 km) multi-indicator atmospheric moisture dataset over China during 2003–2020 (HiMIC-Daily) is publicly available through the National Tibetan Plateau Data Center (TPDC) of China at https://doi.org/10.11888/Atmos.tpdc.303449 (last access: May 2026; (Su et al., 2026b)). HiMIC-Daily provides six atmospheric moisture indicators: relative humidity (RH, %), actual vapor pressure (AVP, hPa), vapor pressure deficit (VPD, hPa), dew point temperature (DPT, ℃), mixing ratio (MR, g kg^{-1}), and specific humidity (SH, g kg^{-1}). The dataset has a high spatial resolution of 1 km × 1 km and covers mainland China at a daily temporal resolution. All indicators demonstrate high predictive accuracy, with R^2 values exceeding 0.877 and MAE and RMSE values within acceptable ranges. The dataset is organized by moisture indicator and year. Six main folders correspond to the six moisture indicators, and each folder contains 18 annual files in ‘.zip’ format, covering 2003–2020. Each ‘.zip’ file contains one annual NetCDF file, in which daily data are stored along the time coordinate. To balance storage efficiency with numerical precision, all variables are scaled by a factor of 100 and saved as 16-bit integer (int16) values. Users can recover the original values by dividing the stored data by 100. The naming rule for each ‘.zip’ file follows “HiMIC_China_Daily_{Indicator}_{Year}.zip”, where “Indicator” denotes one of the six moisture variables (RH, AVP, VPD, DPT, MR, and SH) and “Year” ranges from 2003 to 2020. Further details can be found in the “README.pdf” file.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

6. Code availability

The codes for generating the HiMIC-Daily dataset are available from Zenodo at https://doi.org/10.5281/zenodo.20124067 (last access in May 2026; (Su et al., 2026a)) and are compatible with Python 3.10. The repository contains a folder named "Codes_HiMIC_Daily" that includes three Python scripts: "Input Data Preprocessing.py", "HiMIC-Daily Model.py", and "Result Visualization.py". In addition, "DataSamples_HiMIC_Daily" provides the sample data required to run the codes. The complete repository supports reproducible implementation and independent validation of the proposed framework.

7. Conclusion

Accurate atmospheric moisture data at daily and 1-km scales are essential for advancing the understanding of regional climate variability and hydro-meteorological extremes, assessing public health risks, and maintaining regional ecological and environmental stability. However, existing atmospheric moisture datasets over China remain limited in spatial and temporal resolutions and in indicator coverage. The commonly used products, such as CMFD and CDMet, suffer from coarse spatial resolution and limited coverage of moisture variables, which constrain fine-scale applications. In this study, we develop HiMIC-Daily, a seamless daily 1-km-resolution atmospheric moisture dataset for China spanning 2003–2020, covering six key moisture indicators: RH, AVP, VPD, DPT, MR, and SH.

HiMIC-Daily is generated using the LightGBM framework by integrating meteorological station observations with multiple high-resolution datasets, including land surface temperature (MOD11A1/MYD11A1), land cover (MCD12Q1), topographic variables (MERIT DEM), column water vapor (MCD19A2), vapor pressure (TerraClimate), 2-m SAT and DPT (ERA5-Land), and population density (WorldPop). The dataset demonstrates reliable performance, with R^2 values exceeding 0.877 and MAE and RMSE within acceptable ranges. Compared with CMFD and CDMet, HiMIC-Daily provides finer spatial detail and closer agreement with station

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

observations, thus better capturing variations in moisture indicators. The HiMIC-Daily dataset also captures temporal moisture responses during both cold and heat extremes, highlighting its value for event-scale analyses. The fine spatial and temporal detail of this dataset makes it suitable for a wide range of studies in meteorology, ecology, environmental science, and human health, supporting applications that require fine-scale atmospheric moisture information.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015, Sci Data, 5, 170191, https://doi.org/10.1038/sdata.2017.191, 2018.

Angert, A., Lee, J.-E., and Yakir, D.: Seasonal variations in the isotopic composition of near-surface water vapour in the eastern Mediterranean, Tellus B: Chemical and Physical Meteorology, 60, https://doi.org/10.1111/j.1600-0889.2008.00357.x, 2008.

Antonoglou, N., Balidakis, K., Wickert, J., Dick, G., de la Torre, A., and Bookhagen, B.: Water-Vapour Monitoring from Ground-Based GNSS Observations in Northwestern Argentina, Remote Sensing, 14, https://doi.org/10.3390/rs14215427, 2022.

Baker, L. H., Collins, W. J., Olivié, D. J. L., Cherian, R., Hodnebrog, amp, Oslash, Myhre, G., and Quaas, J.: Climate responses to anthropogenic emissions of short-lived climate pollutants, Atmospheric Chemistry and Physics, 15, 8201–8216, https://doi.org/10.5194/acp-15-8201-2015, 2015.

Baldwin, J. W., Benmarhnia, T., Ebi, K. L., Jay, O., Lutsko, N. J., and Vanos, J. K.: Humidity's Role in Heat-Related Health Outcomes: A Heated Debate, Environmental Health Perspectives, 131, 055001, https://doi.org/10.1289/EHP11807, 2023.

Bengtsson, L.: The global atmospheric water cycle, Environmental Research Letters, 5, https://doi.org/10.1088/1748-9326/5/2/025202, 2010.

Bolton, D.: The Computation of Equivalent Potential Temperature, Monthly weather review, 108, 1046–1053, https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2, 1980.

Buck, A. L.: New Equations for Computing Vapor Pressure and Enhancement Factor, Journal of Applied Meteorology, 20, 1527–1532, https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2, 1981.

Buzan, J. R., Oleson, K., and Huber, M.: Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5, Geoscientific Model Development, 8, 151–170, https://doi.org/10.5194/gmd-8-151-2015, 2015.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Carrera, J., Vázquez-Suñé, E., Castillo, O., and Sánchez-Vila, X.: A methodology to compute mixing ratios with uncertain end-members, Water Resources Research, 40, https://doi.org/10.1029/2003wr002263, 2004.

Castino, F., Bookhagen, B., and de la Torre, A.: Atmospheric dynamics of extreme discharge events from 1979 to 2016 in the southern Central Andes, Climate Dynamics, 55, 3485–3505, https://doi.org/10.1007/s00382-020-05458-1, 2020.

Chen, H., He, W., Sun, J., and Chen, L.: Increases of extreme heat-humidity days endanger future populations living in China, Environmental Research Letters, 17, https://doi.org/10.1088/1748-9326/ac69fc, 2022.

Chen, J. M. and Liu, J.: Evolution of evapotranspiration models using thermal and shortwave remote sensing data, Remote Sensing of Environment, 237, https://doi.org/10.1016/j.rse.2019.111594, 2020.

Chia, S. Y. and Lim, M. W.: A critical review on the influence of humidity for plant growth forecasting, IOP Conference Series: Materials Science and Engineering, 1257, https://doi.org/10.1088/1757-899x/1257/1/012001, 2022.

Choińska, A., Kaleta, A., Winiczenko, R., and Górnicki, K.: Evaluation of models for the dew point temperature determination, Technical Sciences, 3, 241–257, https://doi.org/10.31648/ts.5425, 2017.

Cui, P., Lu, J., Wu, Y., Tang, J., and Jiang, J.: Effect of urban morphology on microclimate and building cluster energy consumption in cold regions of China, Sustainable cities and society, 115, 105838, https://doi.org/10.1016/j.scs.2024.105838, 2024.

Daoud, E. A.: Comparison between XGBoost, LightGBM and CatBoost Using a Home Credit Dataset, https://doi.org/doi.org/10.5281/zenodo.3607805,

Davis, R. E., McGregor, G. R., and Enfield, K. B.: Humidity: A review and primer on atmospheric moisture and human health, Environmental Research, 144, 106–116, https://doi.org/10.1016/j.envres.2015.10.014, 2016.

Dobrowski, S. Z.: A climatic basis for microrefugia: the influence of terrain on climate, Global Change Biology, 17, 1022–1035, https://doi.org/10.1111/j.1365-2486.2010.02263.x, 2011.

Duffour, C., Lagouarde, J. P., and Roujean, J. L.: A two parameter model to simulate thermal infrared directional effects for remote sensing applications, Remote Sensing of Environment, 186, 250–261, https://doi.org/10.1016/j.rse.2016.08.012, 2016.

Dyer, I.: Measurement of humidity, Anaesthesia & Intensive Care Medicine, 13, 121–123, https://doi.org/10.1016/j.mpaic.2011.12.011, 2012.

Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X., and Zeng, W.: Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data, Agricultural Water Management, 225, https://doi.org/10.1016/j.agwat.2019.105758, 2019.

Fischer, E. M. and Knutti, R.: Robust projections of combined humidity and temperature extremes, Nature Climate Change, 3, 126–130, https://doi.org/10.1038/nclimate1682, 2012.

Friedl, M. and Sulla-Menashe, D.: MCD12Q1 MODISTerra+ aqua land cover type yearly L3 global 500m SIN grid V006, https://doi.org/10.5067/MODIS/MCD12Q1.061, 2019.
Fujita, M. and Sato, T.: Observed behaviours of precipitable water vapour and precipitation intensity in response to upper air profiles estimated from surface air temperature, Scientific reports, 7, 4233, https://doi.org/10.1038/s41598-017-04443-9, 2017.
Gamelin, B. L., Feinstein, J., Wang, J., Bessac, J., Yan, E., and Kotamarthi, V. R.: Projected U.S. drought extremes through the twenty-first century with vapor pressure deficit, Scientific Reports, 12, 8615, https://doi.org/10.1038/s41598-022-12516-7, 2022.
Gaughan, A. E., Stevens, F. R., Linard, C., Jia, P., and Tatem, A. J.: High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015, PLoS ONE, 8, e55882, https://doi.org/10.1371/journal.pone.0055882, 2013.
Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.
Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T. W., Sperry, J. S., and McDowell, N. G.: Plant responses to rising vapor pressure deficit, The New phytologist, 226, 1550–1566, https://doi.org/10.1111/nph.16485, 2020.
Gu, L., Schumacher, D. L., Fischer, E. M., Slater, L. J., Yin, J., Sippel, S., Chen, J., Liu, P., and Knutti, R.: Flash drought impacts on global ecosystems amplified by extreme heat, Nature Geoscience, 18, 709–715, https://doi.org/10.1038/s41561-025-01719-y10.1038/s41561-025-01719-y, 2025.
Guarnieri, G., Olivieri, B., Senna, G., and Vianello, A.: Relative Humidity and Its Impact on the Immune System and Infections, Int J Mol Sci, 24, https://doi.org/10.3390/ijms24119456, 2023.
Guo, X., Gui, X., Xiong, H., Hu, X., Li, Y., Cui, H., Qiu, Y., and Ma, C.: Critical role of climate factors for groundwater potential mapping in arid regions: Insights from random forest, XGBoost, and LightGBM algorithms, Journal of Hydrology, 621, 129599, https://doi.org/10.1016/j.jhydrol.2023.129599, 2023.
Harpold, A. A. and Brooks, P. D.: Humidity determines snowpack ablation under a warming climate, Proc Natl Acad Sci U S A, 115, 1215–1220, https://doi.org/10.1073/pnas.1716789115, 2018.
He Jie, Y. K. L. I. X.: China meteorological forcing dataset v2. 0 (1951-2020), National Tibetan Plateau Data Center, https://doi.org/10.11888/Atmos.tpdc.302088, 2024.
He, N., Yin, J., Liu, P., Fu, X., Zhang, Q., Cheng, L., Moulds, S., and Volchak, A.: Global Increases in Dry-Wet Abrupt Alternation Events Under Climate Change,

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Geophysical Research Letters, 52, e2025GL117322, https://doi.org/10.1029/2025GL117322, 2025.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, Journal of Climate, 30, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1, 2017.

Huang, S., Gan, Y., Wang, C., Wang, S., Chen, J., Zhang, X., Chen, Z., Gu, X., Xia, J., Chen, N., and Gong, J.: Urbanization-induced compression effect accelerates and amplifies dry-wet abrupt alternations in Chinese cities, Science Bulletin, 71, 1461–1473, https://doi.org/10.1016/j.scib.2026.01.078, 2026.

Ige, S. O., Ajayi, V. O., Adeyeri, O. E., and Oyekan, K. S. A.: Assessing remotely sensed temperature humidity index as human comfort indicator relative to landuse landcover change in Abuja, Nigeria, Spatial Information Research, 25, 523–533, https://doi.org/10.1007/s41324-017-0118-2, 2017.

Jafarigol, E. and Trafalis, T.: A distributed approach to meteorological predictions: addressing data imbalance in precipitation prediction models through federated learning and gans, arXiv preprint arXiv:2310.13161, https://doi.org/10.48550/arXiv.2310.13161, 2023.

Jain, P., Castellanos-Acuna, D., Coogan, S. C. P., Abatzoglou, J. T., and Flannigan, M. D.: Observed increases in extreme fire weather driven by atmospheric humidity and temperature, Nature Climate Change, 12, 63–70, https://doi.org/10.1038/s41558-021-01224-1, 2021.

Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., and Rehman, M. U.: A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-Short-Term Wind Power Forecasting, IEEE Access, 7, 28309–28318, https://doi.org/10.1109/ACCESS.2019.2901920, 2019.

Kalma, J. D., McVicar, T. R., and McCabe, M. F.: Estimating Land Surface Evaporation: A Review of Methods Using Remotely Sensed Surface Temperature Data, Surveys in Geophysics, 29, 421–469, https://doi.org/10.1007/s10712-008-9037-z, 2008.

Kalnay, E. and Kanamitsu, M.: The NCEPNCAR 40-year reanalysis project, in: Renewable energy, Routledge, Vol1_146–Vol141_194, 2018.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y.: Lightgbm: A highly efficient gradient boosting decision tree, Advances in Neural Information Processing Systems, 30, 2017.

Kimball, J. S., Running, S. W., and Nemani, R.: An improved method for estimating surface humidity from daily minimum temperature, Agricultural and Forest Meteorology, 85, 87–98, https://doi.org/10.1016/S0168-1923(96)02366-0, 1997.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhiro, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q Reanalysis, Journal of the Meteorological Society of Japan. Ser. II, 102, 49–109, https://doi.org/10.2151/jmsj.2024-004, 2024.

Laskar, A. H., Huang, J.-C., Hsu, S.-C., Bhattacharya, S. K., Wang, C.-H., and Liang, M.-C.: Stable isotopic composition of near surface atmospheric water vapor and rain–vapor interaction in Taipei, Taiwan, Journal of Hydrology, 519, 2091–2100, https://doi.org/10.1016/j.jhydrol.2014.10.017, 2014.

Lawrence, M. G.: The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications, Bulletin of the American Meteorological Society, 86, 225–234, https://doi.org/10.1175/bams-86-2-225, 2005.

Lengfeld, K. and Ament, F.: Observing Local-Scale Variability of Near-Surface Temperature and Humidity Using a Wireless Sensor Network, Journal of Applied Meteorology and Climatology, 51, 30–41, https://doi.org/10.1175/jamc-d-11-025.1, 2012.

Li, X., Luo, M., Zhao, Y., Zhang, H., Ge, E., Huang, Z., Wu, S., Wang, P., Wang, X., and Tang, Y.: A daily high-resolution (1 km) human thermal index collection over the North China Plain from 2003 to 2020, Scientific Data, 10, 634, https://doi.org/10.1038/s41597-023-02535-y, 2023.

Liu, X., Li, X., Shi, X., Wu, S., and Liu, T.: Simulating complex urban development using kernel-based non-linear cellular automata, Ecological Modelling, 211, 169–181, https://doi.org/10.1016/j.ecolmodel.2007.08.024, 2008.

Lyapustin, A. and Wang, Y.: MODISTerra+Aqua Land Aerosol Optical Depth Daily L2G Global 1km SIN Grid V061 2022, https://doi.org/10.5067/MODIS/MCD19A2.061, 2022.

Mekis, É., Vincent, L. A., Shephard, M. W., and Zhang, X.: Observed Trends in Severe Weather Conditions Based on Humidex, Wind Chill, and Heavy Rainfall Events in Canada for 1953–2012, Atmosphere-Ocean, 53, 383–397, https://doi.org/10.1080/07055900.2015.1086970, 2015.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth system science data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Murray, F. W.: On the computation of saturation vapor pressure, https://doi.org/10.1175/1520-0450(1967)006<0203:OTCOSV>2.0.CO;2, 1996.

Peng, S., Ding, Y., Liu, W., and Li, Z.: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, Earth System Science Data, 11, 1931–1946, https://doi.org/10.5194/essd-11-1931-2019, 2019.

Peng, Y., He, Y., Wang, X., Jia, J., Huang, B., Li, D., Ji, F., Wang, S., and Guan, X.: The influence of water vapour on winter humidity-related extreme cold events in China, International Journal of Climatology, 44, 3464–3482, https://doi.org/10.1002/joc.8533, 2024.

Qiu, R., Li, L., Kang, S., Liu, C., Wang, Z., Cajucom, E. P., Zhang, B., and Agathokleous, E.: An improved method to estimate actual vapor pressure without relative humidity data, Agricultural and Forest Meteorology, 298-299, 108306, https://doi.org/10.1016/j.agrformet.2020.108306, 2021.

Rangwala, I.: Amplified water vapour feedback at high altitudes during winter, International Journal of Climatology, 33, 897–903, https://doi.org/10.1002/joc.3477, 2012.

Reber, E. E. and Swope, J. R.: On the Correlation of the Total Precipitable Water in a Vertical Column and Absolute Humidity at the Surface, Journal of Applied Meteorology, 11, 1322–1325, https://doi.org/10.1175/1520-0450(1972)011<1322:OTCOTT>2.0.CO;2, 1972.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The global land data assimilation system, Bulletin of the American Meteorological Society, 85, 381–394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

Rosolem, R., Shuttleworth, W. J., Zreda, M., Franz, T. E., Zeng, X., and Kurc, S. A.: The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System, Journal of Hydrometeorology, 14, 1659–1671, https://doi.org/10.1175/jhm-d-12-0120.1, 2013.

Ruckstuhl, C., Philipona, R., Morland, J., and Ohmura, A.: Observed relationship between surface specific humidity, integrated water vapor, and longwave downward radiation at different altitudes, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006jd007850, 2007.

Sangines de Carcer, P., Vitasse, Y., Penuelas, J., Jassey, V. E. J., Buttler, A., and Signarbieux, C.: Vapor-pressure deficit and extreme climatic variables limit tree growth, Glob Chang Biol, 24, 1108–1122, https://doi.org/10.1111/gcb.13973, 2018.

Sarker, I. H.: Machine Learning: Algorithms, Real-World Applications and Research Directions, SN Computer Science, 2, 160, https://doi.org/10.1007/s42979-021-00592-x, 2021.

Sherwood, S. C., Dixit, V., and Salomez, C.: The global warming potential of near-surface emitted water vapour, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aae018, 2018.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Silvestrini, R. A., Soares-Filho, B. S., Nepstad, D., Coe, M., Rodrigues, H., and Assuncao, R.: Simulating fire regimes in the Amazon in response to climate change and deforestation, Ecol Appl, 21, 1573–1590, https://doi.org/10.1890/10-0827.1, 2011.

Snyder, R. L. and Melo-Abreu, J. P.: Frost protection: fundamentals, practice and economics. Volume 1, Fao2005.

Sobolewski, A., Mlynarczyk, M., Konarska, M., and Bugajska, J.: The influence of air humidity on human heat stress in a hot environment, Int J Occup Saf Ergon, 27, 226–236, https://doi.org/10.1080/10803548.2019.1699728, 2021.

Sonak, A. and Patankar, R.: A survey on methods to handle imbalance dataset, International Journal of Computer Science and Mobile Computing, 4, 338–343, 2015.

Sterling, S. M., Ducharne, A., and Polcher, J.: The impact of global land-cover change on the terrestrial water cycle, Nature Climate Change, 3, 385–390, https://doi.org/10.1038/nclimate1690, 2012.

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., and Yan, X.-H.: Super-resolution of subsurface temperature field from remote sensing observations based on machine learning, International Journal of Applied Earth Observation and Geoinformation, 102, https://doi.org/10.1016/j.jag.2021.102440, 2021.

Su, Z., Zhang, H., Wu, S., and Luo, M.: HiMIC-Daily: Code for generating HiMIC-Daily [code], https://doi.org/10.5281/zenodo.20124067, 2026a.

Su, Z., Zhang, H., Wu, S., and Luo, M.: HiMIC-Daily: A high-resolution (daily and 1 km) atmospheric moisture collection over China during 2003–2020 (V1.0) [dataset], https://doi.org/10.11888/Atmos.tpdc.303449, 2026b.

Sun, X., Sun, Q., Yang, M., Zhou, X., Li, X., Yu, A., Geng, F., and Guo, Y.: Effects of temperature and heat waves on emergency department visits and emergency ambulance dispatches in Pudong New Area, China: a time series analysis, Environ Health, 13, 76, https://doi.org/10.1186/1476-069X-13-76, 2014.

Taheri, M., Mohammadian, A., Ganji, F., Bigdeli, M., and Nasseri, M.: Energy-Based Approaches in Estimating Actual Evapotranspiration Focusing on Land Surface Temperature: A Review of Methods, Concepts, and Challenges, Energies, 15, https://doi.org/10.3390/en15041264, 2022.

Tian, H., Zhao, Y., Luo, M., He, Q., Han, Y., and Zeng, Z.: Estimating PM2.5 from multisource data: A comparison of different machine learning models in the Pearl River Delta of China, Urban Climate, 35, https://doi.org/10.1016/j.uclim.2020.100740, 2021.

Tomassini, L.: The Interaction between Moist Convection and the Atmospheric Circulation in the Tropics, Bulletin of the American Meteorological Society, 101, E1378–E1396, https://doi.org/10.1175/bams-d-19-0180.1, 2020.

Wallace, J. M. and Hobbs, P. V.: Atmospheric science: an introductory survey, Elsevier2006.

Wang, J., Hutyra, L., Li, D., and Friedl, M.: Gradients of atmospheric temperature and humidity controlled by local urban land-use intensity in Boston, Journal of

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Applied Meteorology and Climatology, 56, 817–831, https://doi.org/10.1016/j.isci.2023.106837, 2017.

Wang, P., Liu, Z., and Zhang, L.: Sustainability of compact cities: A review of Inter-Building Effect on building energy and solar energy use, Sustainable Cities and Society, 72, 103035, https://doi.org/10.1016/j.scs.2021.103035, 2021.

Wang, P. K.: Moisture plumes above thunderstorm anvils and their contributions to cross-tropopause transport of water vapor in midlatitudes, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002jd002581, 2003.

Willett, K. M., Dunn, R. J. H., Thorne, P. W., Bell, S., de Podesta, M., Parker, D. E., Jones, P. D., and Williams Jr, C. N.: HadISDH land surface multi-variable humidity and temperature record for climate monitoring, Climate of the Past, 10, 1983–2006, https://doi.org/10.5194/cp-10-1983-2014, 2014.

Wolkoff, P. and Kjaergaard, S. K.: The dichotomy of relative humidity on indoor air quality, Environ Int, 33, 850–857, https://doi.org/10.1016/j.envint.2007.04.004, 2007.

Wood, W. H., Marshall, S. J., and Fargey, S. E.: Daily measurements of near-surface humidity from a mesonet in the foothills of the Canadian Rocky Mountains, 2005–2010, Earth System Science Data, 11, 23–34, https://doi.org/10.5194/essd-11-23-2019, 2019.

Wu, W.-B., Ma, J., Banzhaf, E., Meadows, M. E., Yu, Z.-W., Guo, F.-X., Sengupta, D., Cai, X.-X., and Zhao, B.: A first Chinese building height estimate at 10 m resolution (CNBH-10 m) using multi-source earth observations and machine learning, Remote Sensing of Environment, 291, https://doi.org/10.1016/j.rse.2023.113578, 2023.

Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Duan, Q., and Lohmann, D.: Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, Journal of Geophysical Research: Atmospheres, 117, 2011JD016051, https://doi.org/10.1029/2011JD016051, 2012.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D.: A high-accuracy map of global terrain elevations, Geophysical Research Letters, 44, 5844–5853, https://doi.org/10.1002/2017GL072874, 2017.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., Ryu, Y., Chen, G., Dong, W., Hu, Z., Jain, A. K., Jiang, C., Kato, E., Li, S., Lienert, S., Liu, S., Nabel, J. E. M. S., Qin, Z., Quine, T., Sitch, S., Smith, W. K., Wang, F., Wu, C., Xiao, Z., and Yang, S.: Increased atmospheric vapor pressure deficit reduces global vegetation growth, Science Advances, 5, eaax1396, https://doi.org/10.1126/sciadv.aax1396, 2019.

Zhang, H., Luo, M., Zhan, W., Zhao, Y., Yang, Y., Ge, E., Ning, G., and Cong, J.: HiMIC-Monthly: A 1 km high-resolution atmospheric moisture index collection over China, 2003–2020, Scientific Data, 11, 425, https://doi.org/10.1038/s41597-024-03230-2, 2024a.

https://doi.org/10.5194/essd-2026-367
Preprint. Discussion started: 13 July 2026
© Author(s) 2026. CC BY 4.0 License.

Zhang, H., Luo, M., Zhao, Y., Lin, L., Ge, E., Yang, Y., Ning, G., Cong, J., Zeng, Z., Gui, K., Li, J., Chan, T. O., Li, X., Wu, S., Wang, P., and Wang, X.: HiTIC-Monthly: a monthly high spatial resolution (1 km) human thermal index collection over China during 2003–2020, Earth System Science Data, 15, 359–381, https://doi.org/10.5194/essd-15-359-2023, 2023.

Zhang, J., Liu, B., Ren, S., Han, W., Ding, Y., and Peng, S.: A 4 km daily gridded meteorological dataset for China from 2000 to 2020, Scientific Data, 11, 1230, https://doi.org/10.1038/s41597-024-04029-x, 2024b.

Zhang, T., Zhou, Y., Zhu, Z., Li, X., and Asrar, G. R.: A global seamless 1 km resolution daily land surface temperature dataset (2003–2020), Earth System Science Data, 14, 651–664, https://doi.org/10.5194/essd-14-651-2022, 2022.

Zhou, B., Gu, L., Ding, Y., Shao, L., Wu, Z., Yang, X., Li, C., Li, Z., Wang, X., Cao, Y., Zeng, B., Yu, M., Wang, M., Wang, S., Sun, H., Duan, A., An, Y., Wang, X., and Kong, W.: The Great 2008 Chinese Ice Storm: Its Socioeconomic–Ecological Impact and Sustainability Lessons Learned, Bulletin of the American Meteorological Society, 92, 47–60, https://doi.org/10.1175/2010bams2857.1, 2011.

Zhou, H., Ning, S., Yang, R., Nazrollozoda, S., and Hu, Z.: A high-resolution dataset revealing the dynamical variations of the relative humidity in China, Journal of Hydrology, 670, https://doi.org/10.1016/j.jhydrol.2026.135163, 2026.