https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


Open Access Earth System Science Data Discussions

# A Forty-Four-Year Comprehensive Dataset of Maize Phenology in China's Huang-Huai-Hai Plain

Quanjun Zhang[a], Dongli Wu[a*] and Zhaojin Cheng[b*]
[a]*Meteorological Observation Centre, China Meteorological Administration, Beijing, China;*
[b]*Rizhao Meteorological Bureau of Shandong Province, Rizhao, China*
Correspondence: Dongli Wu (wudongli666@126.com) and Zhaojin Cheng (czjau@163.com)

**Abstract**

The dataset presents a FAIR (Findable, Accessible, Interoperable, and Reusable), comprehensive, long-term dataset documenting maize phenology dynamics across China's Huang-Huai-Hai Region (HHHP), a critical area for national grain production. Spanning the period 1981–2024, the dataset integrates observations from 101 agrometeorological stations across eight provinces and municipalities, capturing ten key phenological stages—from sowing to maturity—and deriving four critical growth lengths. A multi-tiered quality control protocol, including automated consistency checks, climate data cross-referencing, and expert arbitration, was applied to ensure data integrity. Analytical outputs include kernel density estimation for characterizing probability distributions and univariate linear regression for quantifying decadal trends. The dataset comprises 1,616 diagnostic plots in JPEG format and two core data tables in XLSX format, with a total uncompressed volume of 1.50 GB.

The dataset is publicly available via the Science Data Bank under the accession code https://doi.org/10.57760/sciencedb.32076 and supports diverse applications in climate impact assessment, crop model improvement, and adaptation strategy development.

**Key words:** maize phenology, Huang-Huai-Hai Region, agrometeorological stations, growth stages, climate adaptation; data quality control

## 1 Introduction

Crop phenology serves as a key physiological indicator linking climate change and agricultural ecosystems, with long-term dynamic records holding irreplaceable scientific value for understanding agricultural response mechanisms in the context of global warming (Tao et al., 2006; Kollah, 2024; Liu & Chen,2025). The Huang-Huai-Hai Region (HHHP), as one of China's most critical strategic zones for grain production, sees its maize phenological patterns influencing not only the rational allocation of regional agro-climatic resources but also directly impacting national food security (Chen et al., 2019; Yan, et al., 2018; Wang, et al, 2024). This region predominantly practices a winter wheat-summer maize rotation system under a temperate monsoon climate, where agricultural production is highly dependent on climatic conditions and particularly vulnerable to extreme weather events (Zhou et al., 2019). Existing research indicates a significant global warming trend, which may increase the intensity and frequency of extreme climate events (ECEs), substantially affecting agricultural production (IPCC, 2022). For instance, extreme heat can impair leaf photosynthesis, disrupt crop pollination, and reduce grain filling rates, while drought stress similarly hinders photosynthesis and adversely affects yield formation (Lobell et al., 2013). In the HHHP, high temperatures during the summer maize growing season pose persistent threats, compounded by prolonged periods without effective precipitation (Xiao et al., 2019).

Although crop models (e.g., DSSAT, WRF-Crop) are essential tools for assessing the potential impacts of climate change on crop productivity, their application often faces challenges such as insufficient localization of phenological module parameters, limited capability in simulating responses to extreme climate events, and a lack of long-term, high-quality station observation data for model calibration and validation (Tao et al., 2018; White et al., 2011; Gawinowski et al., 2025). For example, many process-based crop models used in climate change impact studies fail to account for temporal changes in crop varieties and management practices (Asseng et al., 2013). Additionally, it is difficult to fully incorporate key factors such as pests and diseases, soil quality deficiencies, waterlogging, and heat damage (Challinor et al., 2014; Tokatlidis,2013). Furthermore, while statistical models can quantify crop yield responses to climate change, they lack physiological structure and cannot accurately capture crop growth behaviors (Lobell & Burke, 2010). Historical crop trial data, including precise records of phenology, yield, weather, varieties, and management practices, provide valuable resources for studying how weather influences crop yields, helping to elucidate impact and adaptation mechanisms and improve crop models (Li et al., 2021).

Currently, quantitative analysis of complete maize growth sequences at the regional scale remains challenging, primarily due to a systemic lack of high-quality, long-term, standardized observational data (Piao et al., 2019). Although some studies have utilized historical crop trial data to examine how climate change affects crop growth and productivity under field adaptation conditions, there is still a notable absence of a systematically curated, multi-year, multi-station maize phenology dataset for the HHHP that undergoes rigorous quality control, thereby hindering in-depth mechanistic analysis, reliable model calibration, and the development of targeted adaptation strategies.

To help bridge these research gaps, this study developed a comprehensive maize

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026



phenology dataset for the HHHP spanning 1981-2024, based on the standard framework of the National Agrometeorological Observation Network. The dataset design is grounded in the "climate driver-physiological response-system adaptation" ecological framework, capturing the continuous progression of maize growth and development through complete records of ten key phenological stages from sowing to maturity. Technically, we established a multi-layered quality control protocol incorporating automated rule-based checks, cross-validation with climate data, and expert arbitration, achieving for the first time a systematic integration of paper-based archives and digital records at the regional scale, ensuring data consistency, comparability, and high accuracy over the long term.

This dataset aims to serve as a FAIR-compliant benchmark resource that supports multiple research pathways (Wilkinson et al., 2016). It enables detailed analysis of phenology-climate response mechanisms across different ecological zones by providing continuous phenological sequences, facilitates the refinement of next-generation crop models through high-precision phenological parameters (Bassu et al., 2014), and informs the development of adaptation strategies by clarifying spatiotemporal climate-phenology relationships (Chen et al., 2019).

By systematically integrating 44 years of observational records from 101 stations, this dataset is positioned to serve as key infrastructure for cutting-edge research on agricultural climate change, the formulation of food security strategies, and the practice of smart agriculture. It provides a representative dataset and paradigm from a major agricultural region in China for global crop adaptation studies under climate change.

## 2 Methods

### 2.1 Data Collection Area

The study focuses on the Huang-Huai-Hai Plain (HHHP), a vital agricultural heartland in eastern China spanning 31°N–42°N and 110°E–122°E (Fig. 1). The region has a typical temperate monsoon climate with distinct seasonal variations, featuring hot, rainy summers and cold, dry winters. Mean annual temperatures range from 8°C to 15°C, with annual precipitation between 500 mm and 900 mm, predominantly occurring from June to September. These conditions support highly productive farming systems, particularly the widespread winter wheat-summer maize double-cropping

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026



regime, establishing the HHHP as a cornerstone of China's grain basket. Given its substantial contribution to national maize production and high sensitivity to climate variability, the region serves as an ideal sentinel for detecting crop phenological responses to environmental shifts. Long-term monitoring across this ecologically diverse transect provides critical insights for developing climate-resilient agricultural strategies essential for future food security.

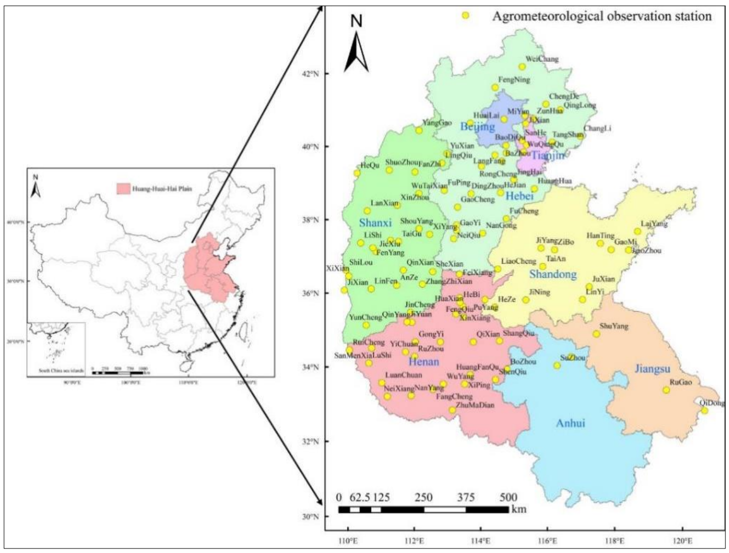



**Fig. 1 Spatial distribution of 101 agrometeorological observation stations in HHHP**

## 2.2 Data Sources

Phenological observations were obtained from 101 national agrometeorological stations operated by the China Meteorological Administration (CMA) spanning the period 1981–2024 across the Huang-Huai-Hai Plain. The station network encompasses Hebei (28 stations), Henan (24 stations), Shandong (12 stations), Shanxi (25 stations), Jiangsu (4 stations), and Anhui (2 stations) provinces, in addition to Beijing (2 stations) and Tianjin (4 stations) municipalities, ensuring comprehensive spatial coverage of this critical agricultural region (Fig. 1).

Station selection was rigorously guided by the principles of spatial representativeness within major maize cultivation zones and long-term record continuity. This strategic selection ensures that the collected data captures phenological behaviors across a wide spectrum of agro-climatic conditions, thereby enhancing the

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


robustness and generalizability of subsequent analyses.

The construction of this dataset involved the integration of heterogeneous data sources, incorporating historical paper-based archives from the 1980s–1990s and modern digital records from the 2000s to the present. This integration required meticulous processes of digitization, cross-verification, and quality control to ensure homogeneity and consistency across the entire temporal range. The resulting dataset provides a unique and valuable resource for analyzing multi-decadal phenological trends, offering high-quality, long-term observational evidence.

The integrated dataset captures the complete growth cycle of maize by documenting 10 key phenological stages: sowing (SO), emergence (EM), three-leaf (TS), seven-leaf (SL), jointing (JO), tasseling (TA), flowering (FL), silking (SI), milking (MI), and maturity (MA). These stages collectively provide a detailed chronicle of maize development from planting to harvest.

### 2.3 Dataset Construction

The methodology for constructing the maize phenology dataset followed a structured, multi-phase workflow, systematically transforming raw observations into a research-ready product, as illustrated in Fig. 2. The process began with rigorous Data Collection and Quality Control, implementing a multi-tiered protocol to ensure data integrity and reliability. This involved the imputation of a minimal proportion of missing values (≈1.1% of records) through a hierarchical strategy that prioritized historical context review, using a 5-year moving window centered on the missing year, and final arbitration by domain experts. Concurrently, anomalies were identified and corrected using a combination of statistical thresholds, such as flagging records beyond ±3 standard deviations from station-specific means, and logical consistency checks to validate the physiological sequence of phenological stages.

Subsequently, a Day of Year (DOY) Conversion was performed, whereby all calendar dates for the 10 phenological stages were converted into numerical values ranging from 1 to 365 (or 366), representing the sequential day from January 1st of each respective year. This standardization is crucial for facilitating subsequent statistical trend analysis and intra-annual comparisons.

Following this, Growth Length Calculation was conducted to derive key developmental phase durations. Four critical growth lengths (GLs) were defined based on fundamental maize developmental milestones: the vegetative phase from sowing to

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


jointing (SO-JO), the pivotal reproductive transition from jointing to silking (JO-SI), the grain-filling period from silking to maturity (SI-MA), and the total lifecycle from sowing to maturity (SO-MA). This multi-stage workflow ensured the generation of a consistent, high-quality dataset suitable for robust spatiotemporal analysis.

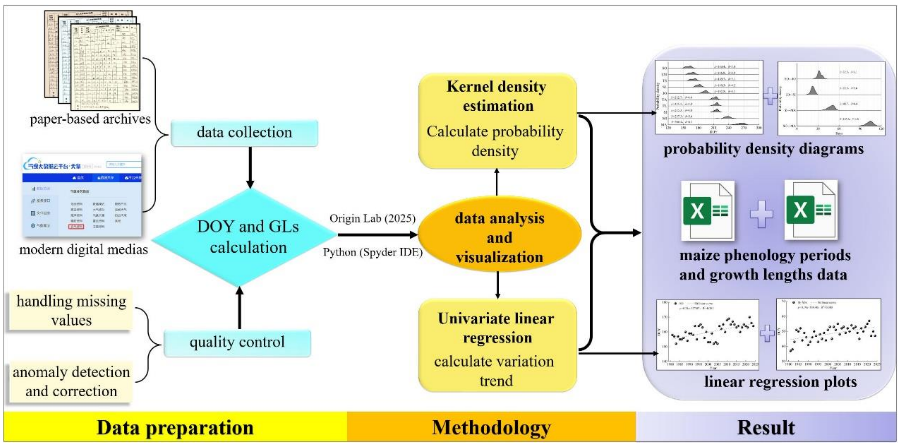



**Fig. 2 Flowchart of the maize phenology dataset construction process for the Huang-Huai-Hai Plain**

### 2.4 Data Analysis and Visualization

A comprehensive analytical approach was employed to characterize the spatiotemporal patterns in maize phenology. Kernel Density Estimation (KDE) was applied as a non-parametric method to model the probability distribution of the Day of Year (DOY) for each phenological event and the calculated Growth Lengths (GLs). Utilizing Scott's rule for optimal bandwidth selection, this technique effectively captured the underlying distribution shapes—including potential multimodality—without imposing restrictive assumptions of normality. For each of the 101 stations, this process generated two key diagnostic plots: one illustrating the probability density of all 10 phenological periods DOYs and another for the 4 GLs, visually summarizing the central tendency and variability of these traits over the study period.

Complementing this distributional analysis, Univariate Linear Regression (ULR) was used to quantify the direction and magnitude of temporal trends in each variable. The slope coefficient derived from ordinary least squares regression served as an indicator of the annual rate of change. To enhance interpretability, the decadal trend was calculated as 10 times the slope. The statistical significance of each trend was rigorously evaluated using a two-tailed t-test at a significance level of $\alpha = 0.05$. This procedure resulted in the creation of 14 distinct linear regression charts per station: 10

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


for the DOY of each phenological stage and 4 for the respective growth phase durations.

The synergistic application of KDE and ULR provides a multi-faceted understanding of phenological dynamics; KDE reveals the static probability structure, while ULR captures the directional temporal shifts. For an in-depth discussion of the specific mathematical formulations and the comparative advantages of these methods in phenological research, readers are referred to Zhang et al. (2025a). The entire analytical workflow was designed to ensure robust, reproducible, and interpretable results from this long-term dataset.

### 2.5 Implementation and Software Workflow

To ensure full reproducibility and computational efficiency, the data processing, statistical analysis, and visualization pipeline was implemented within an integrated software environment. The workflow leveraged the complementary strengths of multiple platforms: initial data integration, sorting, and descriptive statistics were managed within Microsoft Excel. Subsequently, advanced statistical modeling, including kernel density estimation and univariate linear regression, was performed systematically using Origin Lab (2025b). This software was also employed to generate the entire suite of high-resolution (1200 DPI), publication-quality diagnostic plots. Furthermore, custom Python scripts, developed and executed within the Spyder IDE, were deployed to automate batch processing and data validation tasks across the extensive, multi-station dataset. This hybrid approach streamlined the handling of large data volumes while maintaining rigorous, reproducible research standards.

## 3 Data Records

The resulting dataset, "Huang-Huai-Hai Maize Phenology Dataset," is systematically organized to facilitate easy access and reuse. Its hierarchical structure (Fig. 3) begins with a primary root folder, under which data are categorized into provincial-level subfolders (e.g.,*01-Hebei*, *02-Tianjin*). Each provincial directory further contains station-specific subfolders.

At the station level, users will find 16 high-resolution JPEG charts (1,200 DPI) for visual analytics, comprising kernel density estimation plots and linear regression trend diagrams for the 10 phenological periods DOY and the 4 derived GLs. In total, the dataset includes 1,616 such curated diagnostic plots across all stations.

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026

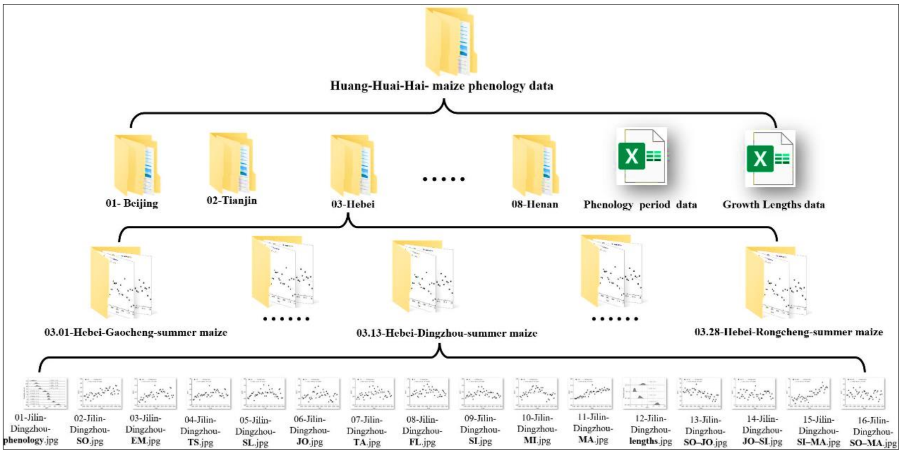




**Fig. 3 The schematic diagram of the folder hierarchy**

The dataset is anchored by two core analytical tables in XLSX format:

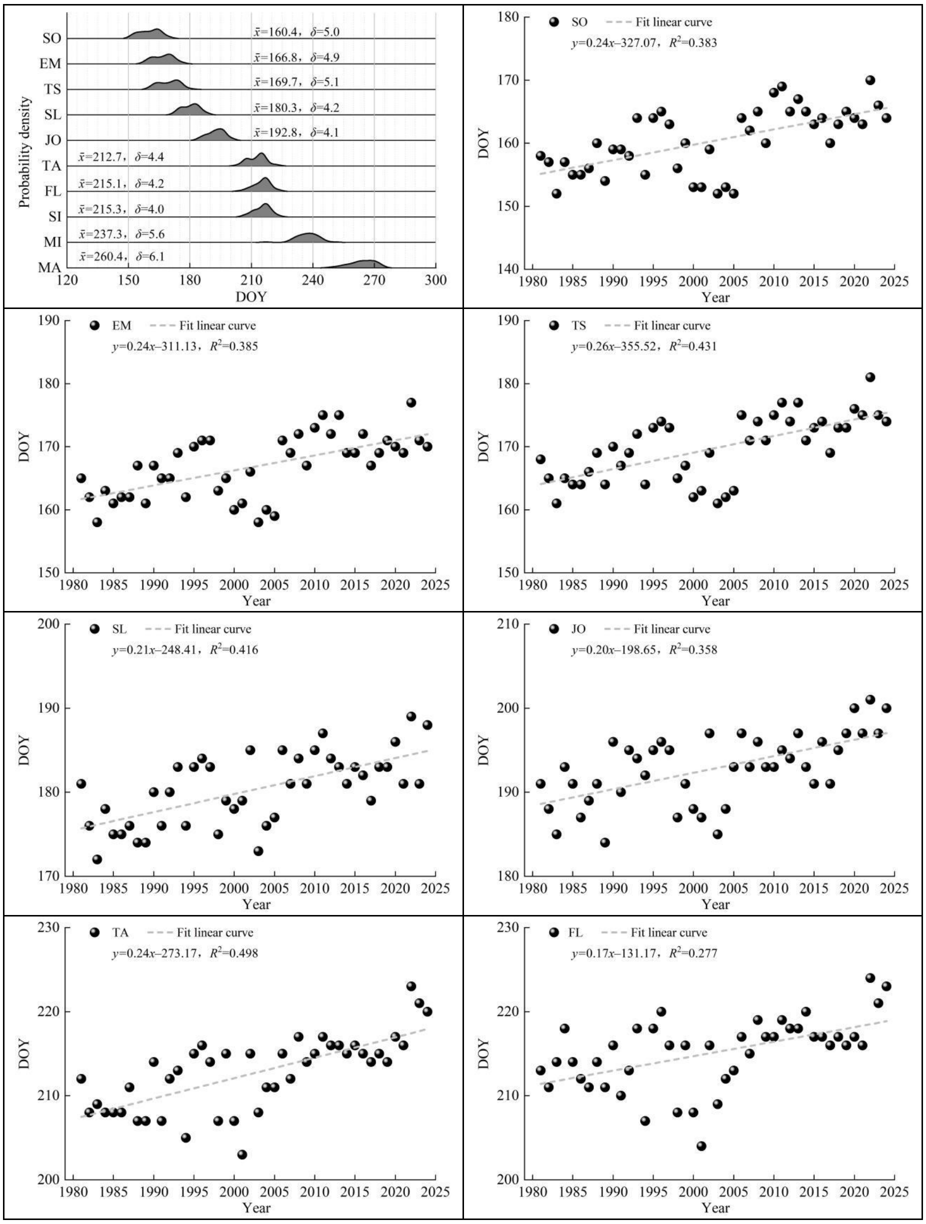
- ***Maize phenology period data.xlsx*** provides comprehensive regression statistics for the DOY of all 10 phenological periods. Key variables include station metadata (province, station ID, station name, crop type, phenology period), descriptive statistics (number of years, maximum, minimum, mean, standard deviation), and linear regression outputs (intercept, annual trend, decadal trend, $R^2$, p-value, F-value, and model mean square).
- ***Maize growth lengths data.xlsx*** contains a regression statistic for 4 GLs. Key variables include station metadata (province, station ID, station name, crop type, growth lengths), descriptive statistics (number of years, maximum, minimum, mean, standard deviation), and linear regression outputs (intercept, annual trend, decadal trend, $R^2$, p-value, F-value, and model mean square).

The total volume of the dataset is 1.50 GB in its uncompressed form, with a compressed archive of 956 MB available for efficient distribution.

Figure 4 presents a curated selection of analytical outputs from the LuanCheng Agrometeorological Station in Hebei Province, serving to illustrate the dataset's structure, analytical depth, and practical utility. This example integrates kernel density estimation (KDE) charts with linear regression diagrams to provide a multi-faceted view of phenological characteristics. The KDE plots visualize the probability distribution of the day of year (DOY) for all ten phenological stages and the four growth lengths (GLs), revealing their temporal concentration and variability. Complementing

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


this, the suite of 14 linear regression diagrams quantifies the temporal trends for each individual phenological DOY and GL, highlighting significant shifts in timing and duration over the study period. Collectively, these visualizations exemplify the dataset's capacity to support robust trend analysis and climatic impact assessment.



https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026

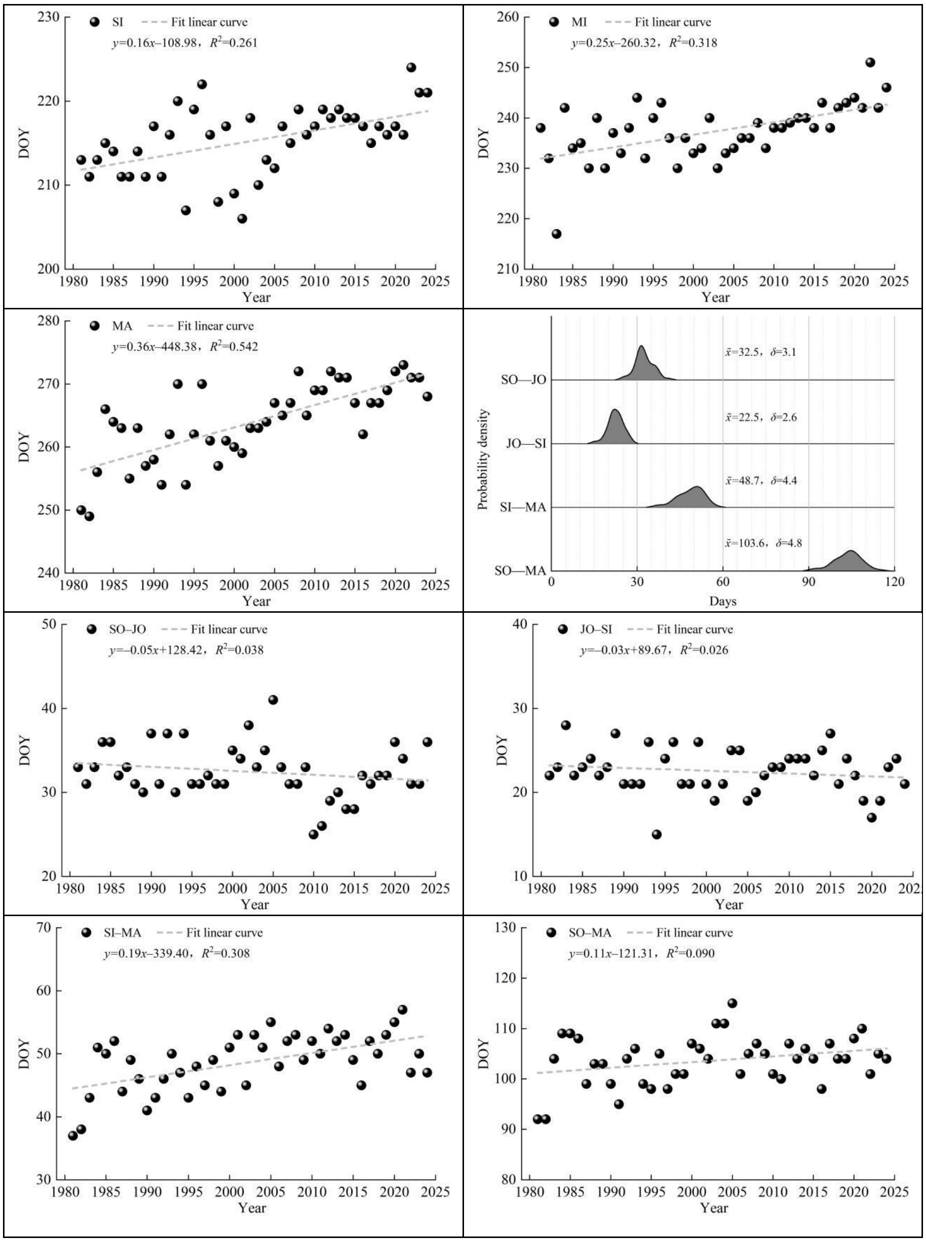




**Fig. 4 Example data: Variation characteristics of maize phenology at LuanCheng Agrometeorological Observation Station, Hebei Province**

Note: SO: sowing, EM: emergence, TS: three-leaf, SL: seven-leaf, JO: jointing, TA: tasseling, FL: flowering, SI: silking, MI: milking, MA: maturity.

# 4 Technical Validation

The credibility and scientific robustness of this maize phenology dataset are established through a comprehensive, multi-faceted validation strategy. This strategy is

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026




built upon a foundation of institutional rigor, standardized observational protocols, and a sophisticated, multi-layered data quality assurance system designed to maximize accuracy and consistency across the 44-year record.

### 4.1 Foundational Framework for Data Integrity

The reliability of the dataset is intrinsically linked to the stringent operational framework of the China Meteorological Administration's (CMA) national agrometeorological network. This framework ensures uniformity and professionalism from data collection to initial processing.

- **Expert Personnel and Standardization:** Data were collected by trained and certified meteorological observers adhering to national standards. Many observers possess over a decade of field experience, ensuring consistent and reliable data recording. Their work strictly adhered to the CMA's "Specification for Agrometeorological Observation," which provides definitive, nationally standardized criteria for identifying and recording each phenological event (detailed in Section 5.3). This minimizes inter-observer variability and ensures methodological continuity across different stations and decades.
- **Systematic Quality Assurance Culture:** Beyond initial collection, a culture of systematic review was embedded within the data generation process. This included regular bimonthly quality inspections, repeated verification cycles, and a structured protocol for addressing discrepancies, forming a proactive first line of defense against data errors.

### 4.2 Multi-Layered Data Quality Control Protocol

To transform raw observations into a research-ready product, a rigorous, multi-tiered Quality Control (QC) protocol was implemented. This protocol leverages both computational automation and domain-expert knowledge to identify and correct anomalies, ensuring a final dataset of high fidelity.

- **Automated Rule-Based Filtering:** The initial QC layer involved a suite of automated checks applied to the digitized records. These checks were designed to flag physiologically implausible data and included:
    - **Temporal Logic Checks:** Verification of the chronological sequence of phenological stages (e.g., ensuring that emergence always follows sowing, and silking follows tasseling).

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026



- **Plausibility Range Checks:** Identification of records where the Day of Year (DOY) for a given stage fell outside biologically reasonable bounds for the region and cultivar.
- **Statistical Outlier Detection:** Flagging of values that deviated by more than ±3 standard deviations from the station-specific long-term mean for each phenological event, helping to identify potential transcription or recording errors.

- **Climate Data Cross-Referencing:** To add an independent layer of validation, the dates of key temperature-sensitive phenological stages (e.g., jointing, flowering) were cross-referenced with daily temperature records from the co-located meteorological stations. Records indicating that a phenological event occurred during a period of climatically implausible conditions (e.g., jointing reported during a sustained cold spell) were flagged for further expert review.
- **Contextual Expert Arbitration:** The final and most critical layer of validation involved human expertise. All records flagged by the automated and cross-referencing checks were submitted to a formal arbitration process by a panel of senior agrometeorologists. This panel, with access to station history logs and regional climatic context, made consensus-based decisions on whether to correct, impute, or exclude questionable data points. A minimal proportion of missing records (<1.1%) was imputed using a hierarchical strategy, which prioritized station-specific historical averages, followed by regional analogues, and finally, expert estimation.

### 4.3 Standardized Phenological Criteria

The consistent application of unambiguous, objective criteria for defining each phenological stage is paramount for the long-term comparability of the data. The determination of each stage was not based on subjective estimation but on clearly prescribed morphological developments of the maize plant, as codified in national standards (National Meteorological Administration, 1993, 2025). These criteria, summarized in Table 1, ensure that an observation of "silking" or "maturity" represents the same plant developmental state regardless of the year, location, or individual observer, thereby solidifying the dataset's utility for trend analysis.

Table 1 Basis for determining maize phenology

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026




| Phenology | Observation basis |
|---|---|
| Sowing (SO) | Actual sowing date. |
| Emergence (EM) | First leaf emerges from coleoptile, approximately 3.0 cm long. |
| Three-leaf (TS) | Third leaf emerges from second leaf sheath, approximately 2.0 cm long |
| Seven-leaf (SL) | Seventh leaf emerges from sixth leaf sheath, approximately 2.0 cm long. |
| Jointing (JO) | Basal internodes transition from flattened to round shape; near-ground stem nodes become round and firm to touch, approximately 3.0 cm long. Tassel differentiation begins at this stage. |
| Tasseling (TA) | Tassel tip emerges from leaf sheath. |
| Flowering (FL) | Anthers from upper-middle section of tassel become visible and release pollen. |
| Silking (SI) | Silks emerge from husks of female ear. |
| Milking (MI) | Silks darken to reddish-brown, outer husks fade slightly but remain green. Kernels reach normal size with thick white milky fluid in lower-middle section of ear. |
| Maturity (MA) | Over 80% of plants exhibit yellowed outer husks, dried silks, hardened kernels displaying variety-specific color. Kernels resist fingernail penetration. |

### 4.4 Assessment of Final Data Quality

The culmination of this exhaustive validation process is a dataset characterized by exceptional integrity. The overall data completeness rate exceeds 98.9% across all stations and years. The estimated error rate for the final phenological dates is conservatively assessed to be below 1%, a level of accuracy that meets or exceeds the standards for similar long-term ecological and agricultural datasets. While minor uncertainties, particularly associated with the digitization of early paper-based records (1980s-1990s), are an inherent limitation of any historical archive, the comprehensive QC measures applied have effectively minimized their impact. Consequently, this dataset is deemed highly suitable for robust statistical analysis, climate impact studies, and the parameterization and validation of process-based crop models.

## 5 Dataset Value

This rigorously quality-controlled dataset delivers an unparalleled resource for the agricultural and climate research communities, distinguished by its exceptional quality, strategic importance, and strict adherence to the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles.

**Unprecedented Data Quality and Scientific Value:** As the first long-term, quality-controlled phenological dataset dedicated to the Huang-Huai-Hai Region, it bridges a critical data gap. Its scientific value is rooted in the integration of multi-decadal, spatially explicit observations (1981–2024) from a dense network of 101 stations, underpinned by a rigorous, multi-tiered quality control protocol that ensures high data integrity (completeness >98.9%, error rate <1%). This high-quality data enables robust quantification of spatiotemporal shifts in maize phenology, which is fundamental for:

- **Detecting Climate Impacts:** Precisely attributing changes in growth timing and duration to climate drivers like warming temperatures and shifting precipitation patterns.
- **Advancing Crop Modeling:** Enhancing the parameterization, calibration, and validation of process-based crop models, leading to more reliable yield projections and climate impact assessments.
- **Informing Breeding and Management:** Identifying climate-resilient phenological traits and optimizing regional adaptation strategies. For instance, the data can directly inform decisions on adjusting sowing dates and selecting cultivars with appropriate growth cycles to avoid critical heat and drought stress periods.

**Strategic and Policy Relevance:** The dataset's focus on China's primary grain basket gives it direct strategic significance for national food security. By elucidating the phenological basis of yield formation under climate change, it provides an evidence-based foundation for policymakers and agricultural stakeholders to develop targeted interventions, safeguarding the stability and sustainability of one of the world's most critical food production systems.

**FAIR Principles Compliance:** The dataset is architected for maximum utility and reuse, fully adhering to the FAIR principles. It is assigned a persistent Digital Object Identifier (DOI) for findability, freely accessible under a CC BY 4.0 license, structured in standardized formats (XLSX, JPEG) with detailed descriptions for interoperability with common analytical tools, and supported by comprehensive documentation to ensure its reusability for independent validation and novel research.

## 6 Data Availability Statement

The "Huang-Huai-Hai Maize Phenology Dataset" supporting this study is publicly available without restrictions from the Science Data Bank (Zhang et al., 2025b). The persistent dataset can be accessed via the following permanent links:

DOI : https://doi.org/10.57760/sciencedb.32076;

STR: https://cstr.cn/31253.11.sciencedb.32076;

The data are released under a Creative Commons Attribution 4.0 International (CC

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


Open Access Earth System Science Data Discussions

BY 4.0) license.

## 7 Conclusion

This study presents a comprehensive, long-term maize phenology dataset for China's Huang-Huai-Hai Plain spanning the period 1981–2024, addressing a critical gap in agricultural climate research. By systematically integrating observations from 101 agrometeorological stations and implementing a rigorous multi-tiered quality control protocol, we have created a FAIR-compliant resource that captures the complete developmental cycle of maize through ten key phenological stages and four derived growth lengths.

The dataset's principal strength lies in its unprecedented combination of temporal depth (44 years), spatial coverage (eight provinces/municipalities), and data quality (completeness >98.9%, error rate <1%), making it uniquely positioned to support diverse research applications. The integrated analytical outputs, including kernel density estimation and univariate linear regression analyses, provide immediate insights into phenological distributions and decadal trends while enabling deeper investigation of climate-crop interactions.

This dataset serves as a valuable foundation for multiple research pathways: it enables detailed mechanistic analysis of phenology-climate relationships across different ecological zones; facilitates improved parameterization and validation of process-based crop models; and supports the development of targeted adaptation strategies to enhance climate resilience in one of China's most critical agricultural regions. By providing a representative dataset from a major global food production area, this resource contributes to international efforts to understand and adapt to climate change impacts on agricultural systems.

Future work will focus on expanding temporal coverage, integrating complementary environmental variables, and developing specialized tools to enhance the dataset's utility for both research and operational applications in precision agriculture and food security planning.

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


Open Access Earth System Science Data Discussions

## Author contribution

Quan-Jun Zhang: Conceptualization, Methodology, Validation, Data Curation, Formal Analysis, Visualization, Writing - Original Draft.
Dong-Li Wu: Resources, Supervision, Validation, Writing - Review & Editing, Project Administration.
Zhaojin Cheng: , Data Curation, Validation, Writing - Review & Editing.
All authors reviewed and approved the final manuscript.

## Competing interests

The authors declare that they have no conflict of interest.

## Usage Notes

To maximize the utility of this dataset, users are recommended to employ statistical computing software (e.g., R, Python, or Origin Pro) for comprehensive data extraction and analysis. The provided high-resolution JPEG plots serve as a direct tool for visual assessment of phenological distributions and temporal trends at each station. For quantitative modeling and in-depth statistical analysis, the underlying data in the XLSX tables should be utilized. To investigate the climatic drivers behind observed phenological shifts, researchers are strongly encouraged to integrate these data with corresponding daily meteorological records (e.g., temperature and precipitation). Proper attribution is essential for sustaining data-sharing initiatives; any publication that uses this dataset should cite the accompanying data paper as referenced in the Data Availability Statement.

## Acknowledgements

We extend our sincere gratitude to the dedicated agrometeorological observers across the China Meteorological Administration (CMA) network, whose long-term commitment to high-quality data collection made this work possible. We also acknowledge the support from various administrative levels within the CMA. Special thanks are due to Dr. Zhu Yongchao, Dr. Liu Cong, Dr. Yang Dasheng, Mr. Kong Xiangsheng, Ms. Zhang Jing, and Mr. Zhang Shihao for their invaluable contributions to data curation and processing. The authors also sincerely thank all colleagues and data contributors who supported this research.

## Financial Supports

This work was supported by the following funding sources: National Key

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


Research and Development Program of China (Grant No. 2024YFD2301301); Innovative Development Special Project of China Meteorological Administration (Grant No. CXFZ2023J069); Agricultural Observation Technology Innovation Team Project of the Meteorological Observation Centre, CMA; State Key Laboratory of Environment Characteristics and Effects for Near-space; Engineering Technology Research Center for Meteorological Observation, CMA.

## References

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., et al. (2013). Uncertainty in simulating wheat yields under climate change. Nature Climate Change, 3(9), 827–832.

Bassu, S., Brisson, N., Durand, J.L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., et al. (2014). How do various maize crop models vary in their responses to climate change factors? Global Change Biology, 20(7), 2301–2320.

Challinor, A.J., Watson, J., Lobell, D.B., Howden, S.M., Smith, D.R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. Nature Climate Change, 4(4), 287–291.

Chen G., Cao H., Chen D., Zhang L., Zhao W., Zhang Y., Ma W., Jiang R., Zhang H., Zhang F. (2019) Developing sustainable summer maize production for smallholder farmers in the North China Plain: An agronomic diagnosis method. Journal of Integrative Agriculture, 18(8):1667-1679

Gawinowski, M., Aubry, M., Buis, S., Cécile Garcia, Deswarte, J. C., & Bancal, M. O., et al. (2025). Refining type and timing of measured crop variables for the calibration of a new winter wheat cultivar in the STICS crop model. bioRxiv. https://doi.org/10.1101/2025.02.10.637374

IPCC (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg2/

Kollah, B., Parmar, R., Devi, M. H., & Mohanty, S. R. (2024). Prediction of Crop Response to Atmospheric Greenhouse Gas Concentration and Climate Parameters. Springer, Cham.

Liu, L., & Chen, X. (2025). Estimation of the impact of climate warming on spring wheat (*triticum aestivum* l.) phenology from observations and modelling in the arid region of northwest China. Journal of Agronomy and Crop Science, 211(1).

Lobell, D.B., & Burke, M.B. (2010). On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11), 1443–1452.

Lobell, D.B., Hammer, G.L., McLean, G., Messina, C., Roberts, M.J., & Schlenker, W. (2013). The critical role of extreme heat for maize production in the United States. Nature Climate Change, 3(5), 497–501.

National Meteorological Administration. Specifications for Agrometeorological Observations[M].

https://doi.org/10.5194/essd-2025-726
Preprint. Discussion started: 13 February 2026


Beijing: China Meteorological Press, 1993.

National Meteorological Administration. Specifications for Agrometeorological Observations[M]. Beijing: China Meteorological Press, 2025.

Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. Global Change Biology, 25(6), 1922–1940.

Tao, F., Xiao, D., Zhang, S., Zhang, Z., & Rötter, R.P. (2018). Adaptation of maize production to climate change in North China Plain: Quantify the relative contributions of adaptation options. Journal of Cleaner Production, 197, 1049–1057.

Tao, F., Yokozawa, M., Xu, Y., Hayashi, Y., & Zhang, Z. (2006). Climate changes and trends in phenology and yields of field crops in China, 1981-2000. Agricultural and Forest Meteorology,138 (1-4), 82-92.

Tokatlidis, I. (2013) Adapting maize crop to climate change[J]. Agronomy for Sustainable Development, 2013, 33(1): 63-79.DOI:10.1007/s13593-012-0108-7.

Wang, X.; Li, X.; Lou, Y.; You, S.; Zhao, H. (2024) Refined Evaluation of Climate Suitability of Maize at Various Growth Stages in Major Maize-Producing Areas in the North of China. Agronomy, 14, 344. https://doi.org/10.3390/agronomy14020344

White, J.W., Hoogenboom, G., Kimball, B.A., & Wall, G.W. (2011). Methodologies for simulating impacts of climate change on crop production. Field Crops Research, 124(3), 357–368.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018.

Xiao, D., Zhao, Y., Bai, H., Hu, Y., Cao, J. (2019) Impacts of climate warming and crop management on maize phenology in northern China. Journal of Arid Land, 11(6):892-903

Yan, H., Liu, F., Niu, E., Gu, X., Yang Y. (2018) Changes of multiple cropping in Huang-Huai-Hai agricultural region, China[J] Journal of Geographical Sciences, 28(11):1685-1699

Zhang, Q.J. (2025b) A Forty-Four-Year Comprehensive Dataset of Maize Phenology in China's Huang-Huai-Hai Plain [DS/OL]. V1. Science Data Bank, [2025-11-19]. https://doi.org/10.57760/sciencedb.32076. DOI:10.57760/sciencedb.32076.

Zhang, Q.J., Wu, D.L., Gao J. (2025a) Variation of winter wheat phenology dataset in Huang Huai Hai Plain of China from 1981 to 2021 [J]. Scientific Data 12, 1203. https://doi.org/10.1038/s41597-025-05368-z

Zhou, B., Ma, W., Sun, X., Gao Z., Ding Z., Li., Zhao M. (2019) Effects of different sowing and harvest dates of winter wheat-summer maize under double cropping system on the annual climate resource distribution and utilization. Scientia Agricultura Sinica,52(9):1501-1517 (in Chinese with English abstract)