**Author’s response**

We sincerely appreciate the reviewer’s thorough evaluation and valuable suggestions, which have been instrumental in improving the clarity and quality of our manuscript. Below, we provide detailed responses to each of the reviewer’s comments. All line numbers mentioned refer to the clean version of the revised manuscript, not the tracked changes version.

**Response to Reviewer – *Geoscientific Model Development***
**Manuscript EGUsphere-2024-4010**
**Authors: Jinhui Zheng, Le Yu *, Zhenrong Du, Liujun Xiao, and Xiaomeng**

Dear Reviewer,

We sincerely thank you for your insightful comments and constructive suggestions, which have helped us improve the quality of our manuscript. Below, we provide detailed responses to each point. The line number is based on the clean version of the revised manuscript, not the track change version. The original *reviewer comments are presented in italic*, while the authors' responses are provided in blue.

*I appreciate the revisions and detailed responses provided by the authors, and I find that the manuscript has improved as a result. Several aspects of the modeling framework are now clearer. However, despite these improvements, I believe that major concerns remain regarding the assumptions and implementation of key components. These concerns build on my earlier comments as well as the clarifications offered in the revision. For clarity and to focus the discussion, I summarize them below as a set of fresh points.*

Response:

We sincerely appreciate your careful review of our manuscript and your valuable feedback. We are pleased to see that you recognize the improvements we have made in the revised manuscript and acknowledge the clarifications and enhancements in several parts of our modeling framework. Regarding the main issues summarized in your current comments, we have carefully analyzed and considered each point and have made corresponding adjustments and supplementary explanations in the revised manuscript. Once again, we thank you for your continued attention to our work and your constructive suggestions, which have played an important role in enhancing the scientific rigor and readability of the manuscript.

● *It's unclear how the correction factor F(EW), generated by the LSTM, is trained end-to-end, given that WOFOST is not a differentiable model. Since WOFOST has many non-differentiable operations (e.g., thresholding, conditional branching), standard gradient-based backpropagation from the loss (e.g., yield error) back to the LSTM parameters isn't directly possible. Could the authors explain how end-to-end optimization of the LSTM is achieved in practice?*

Response:

Thank you for raising this important question. We would like to clarify that the LSTM model in this study is not directly optimized together with WOFOST via end-to-end gradient backpropagation. Instead, a two-stage training approach was adopted. Since the WOFOST model contains many non-differentiable operations, it is not possible to train the LSTM through standard gradient backpropagation across the entire WOFOST. To address this, we used a stepwise training strategy:

- Independent LSTM training
  We first trained the LSTM model using historical extreme weather indices (HDD, LDD, R95P, R10mm, Rx1day, PDSI, VPD) and crop observation data to fit the relationship between the extreme weather adjustment factor F(EW) and crop growth. This process is completely independent of the WOFOST model. The LSTM parameters were updated via a standard gradient descent algorithm (Adam optimizer).

- Integration with WOFOST and validation
  After training, the predicted F(EW)F(EW) was incorporated as an external adjustment factor into WOFOST's DVR calculations for crop growth and yield simulation. Model performance was then validated by comparing simulated results with actual observations.

This approach avoids computing gradients through WOFOST's non-differentiable operations. By using independent LSTM training followed by model integration, F(EW) effectively modulates the crop simulation. This two-stage strategy preserves the mechanistic framework of WOFOST while capturing the nonlinear effects of extreme weather through the LSTM, ultimately improving simulation accuracy. Your comment was crucial for ensuring the clarity of our manuscript, and we have provided a more explicit explanation of the training strategy in Section 2.3.3, lines 215–217 of the revised manuscript to avoid ambiguity.

● *One of the training targets used is yield. However, yield is influenced by many factors beyond phenology. This raises the risk of the model being right for the wrong reason. For example, consider a frost event that has no effect on phenological development but causes yield loss due to tissue damage. In such a case, the model may still adjust F(EW) to compensate for the yield discrepancy — inadvertently altering phenology, even though phenology was not affected by the frost.*

Response:

Thank you for raising this insightful and critical question. Your concern directly addresses the logical connection between the target variables and the core parameter F(EW) in model training. We would like to clarify how this issue is addressed in our study:

- The core role of F(EW) is to adjust phenology, not to directly influence yield
  F(EW) is designed to capture the nonlinear effects of extreme weather on phenology (heading and maturity, as shown in Equation 8), with training data focused on the relationship between extreme weather indices and observed phenological deviations, rather than directly targeting yield. During calibration, phenology simulation errors (differences between observed and WOFOST-simulated heading/maturity dates) are used as the primary loss, while yield errors serve only as an auxiliary constraint. Specifically, the study employs a weighted multi-objective loss function:

$$\mathrm{Loss} = \alpha \times \mathrm{Loss}_{\mathrm{phenology}} + \beta \times \mathrm{Loss}_{\mathrm{yield}}$$

where $Loss_{phenology}$ is the RMSE between simulated and observed phenology, and $Loss_{yield}$ is the RMSE between simulated and observed yield. After extensive testing and cross-validation, the weights were set to $\alpha = 0.8$ and $\beta = 0.2$, giving much higher priority to phenology.

- Supervised phenology training guides LSTM learning
  The LSTM learns the extreme weather adjustment function F(EW) using multiple climate indices (HDD, LDD, R95P, R10mm, Rx1day, PDSI, VPD). F(EW) modifies the daily development rate (DVR, Equation 9) but does not directly alter WOFOST's internal phenology parameters (e.g., base temperature, photoperiod sensitivity). Therefore, even when yield is influenced by non-phenological factors, F(EW) only locally adjusts DVR without violating the physiological logic of phenology. Historical phenology observations (heading and maturity dates) are used as supervision during training (see Section 2.3.4), ensuring that phenology predictions remain accurate regardless of yield error. This prevents over-adjustment of F(EW) in response to yield deviations.

- Validation in extreme weather years confirms consistency
  Sections 3.1 and 3.2 show that WOFOST-EW performs well for both phenology and yield: RMSE of heading decreased from 4.61% to 3.74%, maturity RMSE decreased from 4.74% to 3.98%, and yield $R^2$ increased from 0.67 to 0.76. These results demonstrate that F(EW) improves yield simulation accuracy while maintaining phenology integrity. Yield improvements are therefore a "by-product" of correct phenology adjustment, not a consequence of misdirected intervention by F(EW).

- WOFOST's original modules independently handle non-phenological yield losses
  Non-phenology-related yield processes (e.g., $CO_2$ assimilation, respiration, soil water balance) are handled by WOFOST's original physiological-ecological modules, independent of F(EW). Thus, there is no mechanism by which F(EW) would need to compensate for non-phenological yield loss, avoiding causal misalignment.

In response to the reviewer's comments, we have added clarifications on the loss function design and supervision strategy in Section 2.3.4, lines 231–233, and in Supplementary Text S2. We greatly appreciate your valuable feedback, which has helped enhance the clarity of our manuscript.

● *It is unclear how agromanagement practices are incorporated into the modeling framework. The manuscript mentions that a standard practice involves applying (more than) 300 mm of irrigation water and a base fertilizer with topdressing. However, it is not specified how this is implemented within the WOFOST simulations. Could the authors clarify the timing, quantity, and method of irrigation and fertilizer application used in the simulations? Specifically, how much water and fertilizer are applied, and at which phenological stages or time points?*

Response:

Thank you for raising this important point. We acknowledge that our original manuscript did not provide sufficient detail regarding agricultural management practices—particularly irrigation and fertilization—which may have led to some ambiguity. Here we provide further clarification:

- Model version and applicability
  This study employed the Wofost72_WLP_CWB version within the PCSE framework. This version allows for simulation under water-limited conditions, meaning irrigation strategies can be specified in the model's agromanagement files. However, it does not include nutrient cycling or nitrogen dynamics modules, and thus cannot directly simulate the impact of different fertilization strategies on crop growth (De Wit et al., 2019; De Wit and Boogaard, 2024; Xu et al., 2022). This was not clearly articulated in the original submission. In the revised manuscript, we have added specific details on irrigation configuration and fertilization assumptions in the Methods section (Section 2.2.1, lines 117–121).

- Irrigation practices
  Based on common irrigation practices for winter wheat in the North China Plain and supported by existing survey studies (Li et al., 2012; Sun et al., 2011; Xu et al., 2022), we specified a total irrigation amount of 320 mm across four applications:

  1. ~30 days after sowing (mid-November), 80 mm to ensure safe overwintering;
  2. Late February–early March, 80 mm to promote tillering;
  3. Late March–early April, 80 mm to meet water demand during jointing and booting;
  4. Early May, 80 mm to secure water supply during grain filling.

  These irrigation events and amounts were explicitly recorded in the Agromanagement file of the model, with irrigation efficiency set at 70% to reflect actual field conditions and water use efficiency.

- Simplified fertilization treatment
  Fertilization was treated under the "potential growth" assumption, reflecting both the absence of a nutrient dynamics module in WOFOST and the actual agricultural context of the North China Plain. In this region, winter wheat production is characterized by high fertilizer application rates, which generally eliminate nutrient limitations (Bai et al., 2020; Dai et al., 2021; Liu et al., 2022). Consequently, nutrient supply is not considered a key limiting factor compared to water availability and extreme weather. We therefore did not explicitly define fertilizer amounts or timing but assumed soil nutrients were always sufficient to meet crop demand. This approach, commonly adopted in regional-scale crop modeling (Xue et al., 2024; Zhang et al., 2025), avoids introducing additional uncertainties and keeps the focus on the study's primary objective: the interaction between extreme weather and crop growth.

- Rationale and limitations of the simplified treatment
  The rationality of this approach lies in the fact that the core objective of this study is to improve WOFOST's response to extreme weather conditions, rather than to precisely describe the fertilization process. At the regional scale, it is common to assume that nutrients are non-limiting in order to simplify fertilization effects and highlight the roles of climate and water management (Bai et al., 2020; De Wit et al., 2019). This assumption reduces sources of uncertainty and emphasizes the dominant role of extreme climate factors

in yield formation.

Of course, we also acknowledge the limitation of this approach: it cannot reflect the dynamic regulatory effects of fertilization management and may underestimate potential interactions between nutrient stress and climatic factors. We have discussed this limitation and provided outlooks in the revised manuscript (Discussion Section 4.4, lines 435–442).

In summary, in the revised manuscript we have added specific details of irrigation practices and fertilization assumptions in the Methods section (Section 2.2.1, lines 117–121) and further explanations in the Discussion section (Section 4.4, lines 435–442). We believe these additions will improve the transparency and rigor of the study.

● *It appears that the correction factor F(EW) is computed as a single scalar for the entire season, rather than varying over time. As a result the same correction is applied during both vegetative and reproductive growth, even though weather impacts on phenology may differ substantially between these phases.*

Response:

We sincerely thank the reviewer for this constructive comment. Although the calculation of F(EW) outputs at the growing season level, the temporal features of the input data and the model design implicitly account for stage-specific responses (vegetative growth, reproductive growth), rather than simply applying a single scalar to the whole season. The specific basis is as follows:

- The extreme weather indices used as inputs contain stage-specific dynamic information
  The extreme weather indices for calculating F(EW) are separated by key phenological stages and accumulated dynamically. For example, high temperature degree days (HDD) and low temperature degree days (LDD) are explicitly distinguished between "sowing to heading" and "heading to maturity" thresholds (revised manuscript Section 2.3.1, lines 144–149), and accumulated on a daily basis. This means the indices fed into the LSTM naturally incorporate the intensity differences of extreme weather across different growth stages.

- Limitations of the study design and directions for improvement
  We fully agree with the reviewer that the mechanisms of extreme climate impacts differ significantly across growth stages. We have discussed this limitation in detail in Section 4.4 (lines 442–457) of the Discussion. Specifically, due to data limitations, we only considered heading and maturity stages, while other key phenological stages of winter wheat were not included. This incomplete stage consideration may affect the model's ability to fully capture crop growth dynamics under different conditions. Previous studies have shown that the impact of extreme weather events on crop production varies across growth stages (Feng et al., 2019; Porter and Gawith, 1999). Within the wheat growth cycle, different stages experience different types and intensities of climatic stress, leading to significant differences in yield impact. In particular, severe droughts during the critical growth months of April and May may strongly affect winter wheat yield (Xu et al., 2018; Yang et al., 2020). Moreover, studies on different crops and regions have shown that crop yield is more sensitive to droughts occurring during critical growth stages (Peña-Gallardo et al., 2019; Zipper et al., 2016). These findings highlight important directions for future research and

model improvement. Future studies could improve model accuracy and better capture the effects of extreme weather on wheat growth by explicitly considering stage-specific types and intensities of climatic stress.

- Rationale for the chosen method
  Despite these limitations, our approach was mainly based on two considerations. First, the original WOFOST model lacks direct mechanisms to respond to extreme weather. Therefore, in designing the method, we aimed to incorporate extreme weather impacts in a simplified way, using an overall adjustment factor, to avoid introducing too many free parameters that would be difficult to calibrate. Second, from the perspective of regional-scale winter wheat simulations, introducing overly complex stage-specific adjustments may reduce model stability when applied at large scales, while our research goal was first to verify the feasibility of the "extreme weather–model adjustment" concept.

In summary, while F(EW) appears as a single scalar, its generation process already implicitly contains temporal response logic that reflects stage differences. At the same time, we fully acknowledge the reviewer's comment that future work should further develop stage-specific or dynamic adjustment mechanisms to better capture climatic stress effects at different growth stages. We have added this clarification in the revised manuscript (Discussion, Section 4.4, lines 442–457) and identified it as an important direction for future research.

We thank the reviewer for pointing out this detail, which has helped us present the dynamic response logic of the model more clearly.

● *It seems that the computation of F(EW) requires access to the full-season weather time series, including data beyond the current day. In an in-season context — where the aim is to update the simulation up to today (before harvest) — WOFOST could still be run, but F(EW) would not be available unless future weather were known. Could the authors clarify how this limitation affects the applicability of their approach for real-time or in-season use?*

Response:

We sincerely thank the reviewer for the insightful concern regarding the limitations of the model for real-time applications. This comment is very important, as it indeed touches on the key link between the transition of this approach from academic exploration to practical application. In light of the study design and the practical needs of agricultural production, we would like to further clarify the following points:

- First, regarding the dependence of F(EW) on full-season data
  We acknowledge that in this study, the calculation of F(EW) depends on extreme weather indices across the entire growing season. This design was essentially intended to more accurately capture the cumulative effects and cross-stage interactions of extreme weather on crop phenology. As explained in Section 4.2 of the manuscript, the impacts of extreme weather are often not isolated (for example, drought during the seedling stage can reduce tolerance to later high temperatures, and low temperatures before heading may alter sensitivity to rainfall during the grain-filling stage). Using only partial season data could overlook such cross-stage linkages, leading to bias in the quantification of F(EW), such as misjudging the differential effects of "short-term heat" versus "prolonged heat" on

phenology. Therefore, this design primarily serves the core research objective: improving the accuracy of simulating crop growth mechanisms under extreme weather (particularly for retrospective analyses of historical scenarios and exploring the impact patterns of extreme weather), rather than aiming at real-time forecasting. This also aligns with the characteristics of winter wheat in the North China Plain, which has a fixed single-season growth cycle (October to the following June, as noted in Section 2.1 of the manuscript). For historical data simulations, full-season weather data are available, allowing the advantages of F(EW) to be fully utilized.

- Second, regarding the feasibility of within-season real-time application
  We fully agree with the reviewer's point that "the unknown nature of future weather makes it difficult to calculate F(EW) in real time." However, this limitation can be partially addressed. Considering that agricultural management decisions are often based on stage-specific information, F(EW) can be recalculated in a rolling stage-based manner by combining observed data with short-term forecasts. The key phenological stages of winter wheat in the North China Plain (sowing, regreening, jointing, heading, maturity) occur at well-defined time points (Section 2.2.1 of the manuscript), which makes it possible to divide the growing season into 3–4 stages (e.g., "sowing–regreening," "regreening–jointing," "jointing–heading," "heading–maturity"). At the end of each stage, stage-specific F(EW) can be calculated using complete observed data up to that point (e.g., sowing to regreening) combined with short-term weather forecasts (e.g., 15–30 day forecasts) for the next stage. This stage-specific F(EW) can then be used to adjust phenological simulation of the current stage. For example, at the jointing stage, F(EW) can be generated based on observed data from "sowing–jointing" and forecast data from "jointing–heading," which can then be used to adjust the development rate during the jointing period. As the growing season progresses, forecast data can be iteratively replaced with newly observed data, continuously updating F(EW). While this approach cannot fully eliminate the uncertainty of future weather, it ensures that F(EW) is always based on the "latest available information," with accuracy improving as the season progresses. However, we acknowledge that this study did not further explore the feasibility of such within-season implementation, and this will be an important focus for future work. We clarified this limitation and research direction in Section 4.4 (lines 411–415).

In summary, the dependence of F(EW) on full-season data reflects a mechanism-oriented design that, while presenting challenges for real-time applications, can be mitigated through strategies such as stage-based rolling calculations and historical similarity corrections. We agree with the reviewer's comment and have clarified this limitation and future direction in Section 4.4 (lines 411–415) of the revised manuscript. We thank the reviewer for this constructive suggestion, which has helped us to more clearly define the boundaries of the model's applicability and its potential optimization pathways.

● *Having gained a better understanding of the calibration and evaluation periods, a question remains regarding the representation of extreme weather events during calibration. Does the calibration period include instances of extreme weather? If not, how can the algorithm effectively learn to model their impact if it has not been exposed to such events during training?*

Response:

Thank you for raising this important question, which directly concerns whether the model can effectively learn the mechanisms of extreme weather impacts on crop phenology. Our training/calibration period was 1980–2000 (see Sections 2.3.2 and 2.3.4), during which significant droughts, low-temperature events, and short-duration heavy rainfall occurred (Han and Gong, 2003; Wang et al., 2014; Zheng et al., 2018). More importantly, the extreme weather indices used to train F(EW) are continuous numerical values and were accumulated by key phenological stages, which means that during the calibration/training period the model was exposed to extreme weather scenarios and thus was able to learn the nonlinear relationship between "extreme weather intensity—phenology deviation."

To improve the clarity of the manuscript, we have added relevant explanations in Section 2.3.2, lines 178–179 of the revised version, specifying the coverage of extreme weather during the training period and its effectiveness for learning F(EW). We again thank you for your insightful comment.

● *Although the correction factor F(EW) is computed using a non-linear function (i.e. an LSTM), its integration into the phenology module appears relatively simple — it acts as a multiplicative factor applied to the development rate. Could the authors elaborate on the rationale for choosing this specific coupling mechanism? Are other forms of integration (e.g., additive or stage-specific adjustments) considered or tested?*

Response:

Thank you for raising this important question. The core reason we chose the multiplicative coupling mechanism is that it better aligns with both the physiological logic of phenological development and the compatibility of the model. The detailed explanation is as follows:

- Consistency with the physiological mechanism of "multi-factor coordinated regulation" in crop phenological development
  The phenological development of wheat is the result of the combined effects of temperature (F(T)), photoperiod (F(P)), and vernalization (F(V)), which in WOFOST are already coupled through a multiplicative relationship (Equation 2). This reflects the physiological principle that "if any single factor is limiting, overall development is constrained" (for example, if vernalization is insufficient, development will be delayed even under favorable photoperiods). Studies have shown that crop growth and development are governed by the combined action of multiple environmental factors rather than by a single factor independently. For example, research on wheat has found that both temperature and day length jointly affect the growth cycle, with clear interactions (Porter and Delecolle, 1988). The impact of extreme weather on phenology (such as high temperatures accelerating development or drought altering the pace of development) essentially serves as a "modulation" of this coordinated process. For instance, extreme heat may amplify the weight of the temperature factor, while extreme drought may reduce the sensitivity of photoperiod responses. Therefore, introducing F(EW) in a multiplicative form naturally integrates into this "multi-factor coordination" physiological framework, preserving the original mechanism while capturing the proportional regulatory effect of extreme weather

on overall development. To improve clarity, we have added supplementary explanations in Section 2.3.3, lines 219–222 of the revised manuscript.

- Compatibility with the modular design of the WOFOST model, avoiding logical conflicts
  One of the core strengths of WOFOST is its mechanistically transparent modular structure (Section 2.3.3), where phenological development rate (DVR) calculations strictly follow a "factor product" logic (Equation 2). If an additive form were used instead, it could lead to results with ambiguous physical meaning.

- Stage-specific adjustment
  The "stage-specific adjustment" raised by the reviewer (e.g., applying different correction forms to vegetative and reproductive phases) is not ignored, but rather addressed through the input features of F(EW) and the temporal learning capability of the LSTM:
  • The input to F(EW) consists of stage-specific extreme weather indices (for example, the HDD threshold is 25°C for sowing–heading and 30°C for heading–maturity, Section 2.3.1). During training, the LSTM has already learned that "the correction magnitude of F(EW) under extreme cold in the vegetative stage differs from that in the reproductive stage." Similar stage-based data input strategies have been used in other studies, allowing models to learn differential responses to environmental factors across growth stages and thereby improving accuracy.
  • However, we acknowledge certain limitations in our study, namely that more precise phenological stages were not considered. We discuss this limitation in detail in Section 4.4, lines 442–457. Specifically, due to data constraints, we only considered the heading and maturity stages, while other key phenological stages of winter wheat were not included. This incomplete treatment of growth stages may limit the model's ability to fully capture crop dynamics under varying conditions. Previous studies have shown that the impacts of extreme weather events on crop production differ significantly across growth stages (Feng et al., 2019; Porter and Gawith, 1999). During the wheat growth cycle, different stages experience different types and intensities of climate stress, leading to marked differences in yield impacts. Moreover, severe drought occurring in the critical growth months of April–May can particularly affect winter wheat yields (Xu et al., 2018; Yang et al., 2020). In addition, a number of studies on different crops and regions have shown that yields are more vulnerable to drought occurring during critical growth stages (Peña-Gallardo et al., 2019; Zipper et al., 2016). These findings point to important directions for future research and model improvements. Future work could further refine the model to account for climate stresses of specific types and intensities at different growth stages, thereby improving predictive accuracy and better capturing the impacts of extreme weather events on wheat development.

Once again, we sincerely thank the reviewer for this constructive suggestion, which has helped us more clearly articulate the reasoning behind our model design.

● *The loss function includes both phenological stage errors (predicted vs. observed) and yield errors (predicted vs. observed). The authors should provide a clear rationale for how these components are balanced during training. Are they weighted equally, or is a specific weighting scheme applied? As previously noted, including yield as a target may confound the mechanisms, since yield is* *influenced by many factors beyond phenology. Clarifying this point is important to assess whether the model is learning the intended relationships.*

Response:

Thank you very much for the detailed suggestion. Regarding the design of the loss function, we did consider both phenology error and yield error during model training, but not through a simple "equal weighting." As noted earlier, the weighting coefficients for phenology and yield losses are 0.8 and 0.2, respectively, meaning the phenology error was assigned a much higher weight than the yield error. We have added clarifications in the revised manuscript (Section 2.3.4, lines 231–233) as well as in the Supplementary Material (Text S2) to explain the logic behind the loss function design. The original purpose of F(EW) was to adjust for the effects of extreme weather on phenological development (Section 2.3.3 of the manuscript), rather than to directly simulate yield formation. Therefore, assigning a higher weight to phenological errors ensures that the LSTM prioritizes learning the relationship between extreme weather and phenological responses (e.g., accelerated heading under high temperatures or delayed maturity due to frost) during training, preventing the yield objective from distorting the logic of phenology simulation.

**References**


Bai, S., Kang, Y., and Wan, S.: Drip fertigation regimes for winter wheat in the North China Plain, Agricultural Water Management, 228, 105885, https://doi.org/10.1016/j.agwat.2019.105885, 2020.

Dai, N., SHI, W., and SHI, X.: Optimal nitrogen application rate for winter wheat under multi-objective constraints in the North China Plain, Chinese Journal of Eco-Agriculture, 29, 1512–1523, 2021.

De Wit, A. and Boogaard, H.: A gentle introduction to WOFOST, Wageningen Environmental Research, 2024.

De Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., Supit, I., van der Wijngaart, R., and van Diepen, K.: 25 years of the WOFOST cropping systems model, Agric. Syst., 168, 154–167, https://doi.org/10.1016/j.agsy.2018.06.018, 2019.

Feng, P., Wang, B., Li Liu, D., Waters, C., and Yu, Q.: Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia, Agricultural and Forest Meteorology, 275, 100–113, https://doi.org/10.1016/j.agrformet.2019.05.018, 2019.

Han, H. and Gong, D.: Extreme climate events over northern China during the last 50 years, Journal of Geographical Sciences, 13, 469–479, https://doi.org/10.1007/BF02837886, 2003.

Li, Q., Zhou, X., Chen, Y., and Yu, S.: Water consumption characteristics of winter wheat grown using different planting patterns and deficit irrigation regime, Agric. Water Manage., 105, 8–12, https://doi.org/10.1016/j.agwat.2011.12.015, 2012.

Liu, Y., Han, M., Zhou, X., Li, W., Du, C., Zhang, Y., Zhang, Y., Sun, Z., and Wang, Z.: Optimizing nitrogen fertilizer application under reduced irrigation strategies for winter wheat of the north China plain, Irrigation Science, 40, 255–265, https://doi.org/10.1007/s00271-021-00764-w, 2022.

Peña-Gallardo, M., Vicente-Serrano, S. M., Quiring, S., Svoboda, M., Hannaford, J., Tomas-Burguera, M., Martín-Hernández, N., Domínguez-Castro, F., and El Kenawy, A.: Response of crop yield to different time-scales of drought in the United States: Spatio-temporal patterns and climatic and environmental drivers, Agricultural and forest meteorology, 264, 40–55, https://doi.org/10.1016/j.agrformet.2018.09.019, 2019.

Porter, J. and Delecolle, R.: Interaction of temperature with other environmental factors in controlling the development of plants., Symposia of the Society for Experimental Biology, 133–156, 1988.

Porter, J. R. and Gawith, M.: Temperatures and the growth and development of wheat: a review, European journal of agronomy, 10, 23–36, https://doi.org/10.1016/S1161-0301(98)00047-1, 1999.

Sun, Q., Kröbel, R., Müller, T., Römheld, V., Cui, Z., Zhang, F., and Chen, X.: Optimization of yield and water-use of different cropping systems for sustainable groundwater use in North China Plain, Agricultural water management, 98, 808–814, https://doi.org/10.1016/j.agwat.2010.12.007, 2011.

Wang, J., Wang, E., Yin, H., Feng, L., and Zhang, J.: Declining yield potential and shrinking yield gaps of maize in the North China Plain, Agricultural and forest Meteorology, 195, 89–101, https://doi.org/10.1016/j.agrformet.2014.05.004, 2014.

Xu, X., Gao, P., Zhu, X., Guo, W., Ding, J., and Li, C.: Estimating the responses of winter wheat yields to moisture variations in the past 35 years in Jiangsu Province of China, PLoS One, 13, e0191217, https://doi.org/10.1371/journal.pone.0191217, 2018.

Xu, X., Shen, S., Gao, F., Wang, J., Ma, X., Xiong, S., and Fan, Z.: Considering different water supplies can improve the accuracyof the WOFOST crop model and remote sensing assimilation in predicting wheat yield, Int. Agrophys., 36, 337–349, https://doi.org/10.31545/intagr/154892, 2022.

Xue, J., Sun, S., Luo, L., Gao, Z., Yin, Y., Zhao, J., Li, C., Wang, Y., and Wu, P.: Quantification of wheat water footprint based on data assimilation of remote sensing and WOFOST model, Agricultural and Forest Meteorology, 347, 109914, https://doi.org/10.1016/j.agrformet.2024.109914, 2024.

Yang, J., Wu, J., Liu, L., Zhou, H., Gong, A., Han, X., and Zhao, W.: Responses of winter wheat yield to drought in the North China Plain: spatial–temporal patterns and climatic drivers, Water, 12, 3094, https://doi.org/10.3390/w12113094, 2020.

Zhang, J., Yang, G., Gao, M., Zhang, L., Chen, W., Liu, M., Zhang, Y., Tang, A., and Li, Z.: Wheat yield prediction using an enhanced WOFOST with soil stratified hydrothermal module driven by GLASS and ERA5-land products over Yellow River Basin, Computers and Electronics in Agriculture, 237, 110669, https://doi.org/10.1016/j.compag.2025.110669, 2025.

Zheng, J., Yu, Y., Zhang, X., and Hao, Z.: Variation of extreme drought and flood in North China revealed by document-based seasonal precipitation reconstruction for the past 300 years, Climate of the Past, 14, 1135–1145, https://doi.org/10.5194/cp-14-1135-2018, 2018.

Zipper, S. C., Qiu, J., and Kucharik, C. J.: Drought effects on US maize and soybean production: spatiotemporal patterns and historical changes, Environmental Research Letters, 11, 094021, https://doi.org/10.1088/1748-9326/11/9/094021, 2016.