Huijing Wang[1], Junjun Cao[1,4, *], Jinxiao Li[2, *], Qun Tian[3], Dev Niyogi[5,6]

[1] Key Laboratory for Geographical Process Analysis & Simulation of Hubei province/School of Urban and Environmental Sciences, Central China Normal University, Wuhan 430079, China.

[2] State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

[3] Guangzhou Institute of Tropical and Marine Meteorology/Guangdong Provincial Key Laboratory of Regional Numerical Weather Prediction, CMA, Guangzhou 510641, China

[4] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

[5] Department of Geological Sciences, Jackson School of Geosciences, and Department of Civil, Architecture, and Environmental Engineering, University of Texas at Austin, Austin, TX 78712, USA.

[6] Department of Agronomy, and Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA.

Corresponding author: Junjun Cao(`caojun@ccnu.edu.cn`), Jinxiao Li(`lijinxiao@lasg.iap.ac`.cn).

Key Points:

- A new model that couples a dynamical subseasonal-to-seasonal atmospheric prediction system and machine learning algorithms is proposed.
- The proposed model significantly improves the accuracy of winter wheat yield forecasting compared to traditional statistical-based modeling.
- The proposed prediction model achieves a skilled prediction one season before harvest.

Abstract

Subseasonal-to-seasonal (S2S) prediction of winter wheat yields is crucial for farmers and decision-makers to reduce yield losses and ensure food security. Recently, numerous researchers have utilized machine learning (ML) methods to predict crop yield using observational climate variables and satellite data. Meanwhile, some studies also illustrate the potential of state-of-the-art dynamical atmospheric prediction in crop yield forecasting. However, the potential of coupling both methods has not been fully exploited. Herein, we aim to establish a skilled ML–dynamical hybrid model for crop yield forecasting (MHCF v1.0), which hybridizes ML and a global dynamical atmospheric prediction system, and apply it to northern China at the S2S time scale. MHCF v1.0 demonstrates that crop yield forecasting with S2S dynamical predictions generally outperforms that with observational climate data. The coupling of ML and

S2S dynamical atmospheric prediction provides a useful tool for yield forecasting, which could guide agricultural practices, policy-making and agricultural insurance.

Plain Language Summary

More cushioning time can be provided for the insurance industry and policy-makers in China by timely and accurate prediction of winter wheat yields. Subseasonal-to-seasonal (S2S) atmospheric prediction and machine learning (ML) algorithms are emerging and advanced technologies that have both been proven to boost crop yield forecasting, but coupling them together into one, hybrid approach has not been fully exploited, and their joint contribution to crop yield prediction needs to be quantified. In this study, we evaluate the contribution of using S2S atmospheric prediction and ML algorithms to improve winter wheat yield prediction in northern China. To this end, we establish an ML–dynamical hybrid ensemble prediction model for winter wheat by incorporating skilled S2S atmospheric prediction outputs and advanced ML algorithms. We find that the ML algorithms demonstrate superior prediction performance compared to traditional statistical models, and S2S atmospheric prediction also achieves better prediction accuracy compared to models based on historical observed climate data or remote sensing data. This study proves that combining S2S atmospheric prediction and ML is a novel and promising technique for regional yield prediction and is expected to be extended to forecasting crop yields in other regions and even at the global scale.

1 Introduction

Growing populations and an increasing frequency of extreme weather events are together bringing great challenges to global food security (Cole et al., 2018; Prosekov & Ivanova, 2018; Ray et al., 2012; Tilman et al., 2011). Besides, timely crop yield forecasting on subseasonal-to-seasonal (S2S) time scales (2 weeks to 3 months) is of great interest to agricultural production, decision-making, market futures, and the insurance industry (Chipanshi et al., 2015; Iizumi et al., 2018; Jiang et al., 2020). Wheat is the world's most widely distributed cereal with the largest planting area (FAO et al., 2021)—especially for China, the world's largest producer and consumer of wheat, where the planting area and yield of winter wheat account for about 94% and 95% of the total planting area and wheat yield, respectively (Huang et al., 2017). Hence, this study examples the winter wheat planting regions of northern China.

Numerous studies have been carried out on crop yield prediction, based on either crop growth modeling or statistical regression. Crop growth modeling aims to reproduce the key processes of plant growth and development in detail from daily meteorological data, cultivar features, soil properties, and agro-management information (Pagani et al., 2017).As such, crop growth models are crucial for providing farmers or specialists with real-time information about their crops, giving risk-assessment information and monitoring decision-making relevant to agricultural management by quantifying the impact of weather, soil, and management interaction on crops (Benami et al., 2021). However, the high computational costs and data requirements involved hinder scaling the approach to multiple crops and regions (Kostková et al., 2021; Li et al., 2020). Traditional linear regression models are based on the empirical relationships between historical yields and other factors, such as climate variables, agrometeorological factors, and/or remote sensing data. However, as this method is unable to consider dynamical meteorological factors with the changing of growth stages, it has limited ability to disentangle the complex nonlinear relationships between independent variables and yields (Feng et al., 2020; Bolton & Friedl, 2013; Pan et al., 2012; Wang et al., 2014; Zhang et al., 2014). By contrast, machine learning (ML) algorithms—more advanced regression methods and more popular in agricultural production—use data or experience to improve the performance of specific algorithms (Goldberg et al., 1988), particularly by explaining a higher-order and nonlinear relationship. Increasingly, ML has been used for agricultural applications, such as crop type classification and crop yield prediction (Klompenburg et al., 2020), and is becoming an indispensable and mainstream tool in precision agriculture.

Climate variables, such as temperature and precipitation, have been confirmed to have significant impacts on crop production, and explain approximately one third of global crop yield variability (Ray et al., 2015; Cai et al., 2019; Guo et al., 2021; Liu et al., 2020). Remote sensing data enable the rapid monitoring and forecasting of agricultural information such as crop growth and grain yield (Liaqat et al., 2017; Rembold et al., 2015; Wu et al., 2014) to be achieved on a large scale. Therefore, the majority of previous studies have used observational climate data and remote sensing data. Recently, some studies have illustrated that dynamical atmospheric prediction is more advantageous than observational climate data in predicting crop yields, which raises the idea that seasonal agricultural production forecasting could benefit directly from dynamical atmospheric prediction (Brown et al., 2018 ; Iizumi et al., 2018; Peng et al., 2018).

S2S dynamical atmospheric prediction, which aims to bridge the gap between medium-range weather forecasts and seasonal prediction, is an emerging and fast-developing field. The science community has rallied under the World Weather Research Programme–World Climate Research Programme S2S Prediction Project, dedicated to improving the forecasting skill and understanding of the sources of S2S predictability. Substantial progress has been made recently on predicting the onset, evolution and decay of some large-scale extreme events, such as heat waves and tropical cyclones (Vitart & Robertson, 2018). Considering the potential significant advantages of S2S atmospheric prediction paired with ML, to the best of our knowledge, no study has combined these two methods in a hybrid approach to predict crop yields. In the present study, we seek to address this knowledge gap.

Specifically, we use S2S dynamical prediction system outputs and a variety of algorithms to build an ML–dynamical hybrid model for crop yield forecasting (MHCF v1.0). The motivations behind this study are to (1) compare the performance in crop yield forecasting based on observed meteorological data/remote sensing data and S2S atmospheric prediction system outputs; (2) investigate the potential of various algorithms for S2S crop yield forecasting; and (3) evaluate how early MHCF v1.0 can forecast winter wheat yield with reasonable accuracy. Section 2 and Section 3 describe the data and methods used in this study, respectively. Section 4 presents results from a multi-model comparison and leave-one-year-out cross-validation, and reports findings on the spatial distribution of yield forecasting and the optimum lead time. Sections 5 and 6 provide some further discussion and conclusions, respectively.

2 Data

2.1 Cropland and winter wheat yield data

We collected county-level winter wheat yield data (t/ha) in the winter wheat producing regions of northern China from 2004 to 2014, which were gathered by the Agricultural Statistical Yearbook of the Ministry of Agriculture of China. The winter wheat planting areas are used to mask the winter wheat yields with a resolution of 1 km in China (Luo et al., 2020). To match the spatial scale of other variables, the county-level winter wheat yields were assigned to a 0.5° grid through weighted averaging.

2.2 Satellite data

Enhanced Vegetation Index (EVI) has been proven to be superior in crop yield prediction than Normalized Difference Vegetation Index, which is more sensitive to higher canopy Leaf Area Index (Bolton & Friedl, 2013; Franch et al., 2015; Tilman et al., 2011; Zhou et al., 2019; Cao et al., 2021; Ma et al., 2021). Thus, we chose EVI as the satellite indicator, which was derived from the MOD13C1 (Collection 6) product with a 16-day repeat and 0.05° spatial resolution (Text S1).

2.3 Observational climate data

Observational climate variables were obtained from the Climatic Research Unit (CRU), including monthly maximum temperature, minimum temperature, mean temperature, precipitation, vapor pressure deficit (VPD), and growing degree day (GDD). One another variable was Standardized Precipitation–Evapotranspiration Index (Beguería et al., 2014). VPD and GDD were calculated from the CRU variables (Text S1) (Cai et al., 2019).

2.4 FGOALS-f2 S2S climate prediction data

The S2S atmospheric prediction outputs were obtained from the FGOALS-f2 dynamical forecasting system, which has been applied at China National Climate Center for real-time S2S prediction (Li et al., 2021(a), 2021(b); Vitart et al., 2017). Studies have shown that FGOALS-f2 is skilled in predicting extreme events, such as summer drought and tropical cyclones genesis, which inevitably affect crop growth and factual yield (Feng et al., 2020; Ren et al., 2019). In this study, we used the monthly atmospheric prediction outputs of FGOALS-f2 with a 0.5° spatial resolution for the next whole month to forecast the winter wheat yield, including 925-hPa air temperature in K, 925-hPa eastward wind in m/s, 925-hPa northward wind in m/s, 925-hPa specific humidity in kg/kg, ground temperature in K, surface (2-m) air temperature in K, total precipitation rate in mm/h, and surface net shortwave radiation in W/m^2. A summary of the datasets and detailed information on these variables are given in Table S1 and S2.

3 Methods

3.1. Model development

Four ML methods (MLR, SVR, RF, XGBoost) were adopted to establish prediction models between input variables and winter wheat yield. More information about the selected algorithms can be found in Text S2.

3.2 Model evaluation

Data preprocessing was carried out by dividing the dataset into training data and testing data, and then preprocessing the training data and testing data separately. We randomly divided the whole dataset into 70% training data and 30% testing data. To have a mean of 0 and a standard deviation of 1, the training data and testing data were normalized respectively by the Z-score (Cai et al., 2019). Additionally, the ten-fold cross-validation technique was adopted to evaluate the performance of the developed models. That is, the entire dataset was randomly divided into 10 subsets, each subset being a testing set, and the rest were used as a training set. Next, the best hyper-parameters for each model were determined by five-fold cross-validation using the GridSearchCV package (Cawley & Talbot, 2010; Molinaro et al., 2005).

Finally, we conducted a "leave-one-year-out" prediction to assess the practicality of the models, i.e., using all years' data during 2004 to 2014 except the target year to train the model and then make a prediction for the target year (Peng et al., 2018). This approach is an extensively used cross-validation method because of its simplicity, universality, and superiority in avoiding the issue of over-fitting (Cao et al., 2021; Han et al., 2020; Peng et al., 2018).

To evaluate the model performance, the coefficient of determination (R^2), root-mean-square error (RMSE), and percent error (PE) were selected as the evaluation metrics in this paper (Text S1).

3.3 Experimental design

Three experiments were designed to address the aims of our study as outlined in the introduction. Firstly, we trained crop yield models with observational climate data and S2S atmospheric prediction outputs separately in order to identify which data source is superior for crop yield forecasting. Secondly, we compared several yield prediction models with different algorithms to find the model with the highest accuracy. And thirdly, we performed in-season prediction to assess the lead time that can reasonably predict the winter wheat yield. The in-season prediction started at the beginning of the growing season and ended one month before harvest.

4 Results

4.1 Model evaluation and multi-model comparison

In order to compare the performance of observational climate data with S2S dynamical atmospheric prediction outputs in yield forecasting, three groups of data were integrated into the selected models—namely, S2S dynamical atmospheric prediction outputs (S2S), observational climate data, and the combination of the two. We found that the performance of the S2S dynamical atmospheric prediction outputs was better than the observational climate data in all selected ML models, but not for MLR (Figure 1). S2S atmospheric prediction outputs alone outperformed the observational climate data, as well as the combination of both, among the ML-based models, with an R^2 of 0.85, 0.84 and 076 in XGBoost, RF and SVR (RMSEs of 0.78 t/ha, 0.82 t/ha and 0.99 t/ha), respectively. Meanwhile, the prediction R^2 values of the observational climate variable–based models were 0.81, 0.8 and 0.65 in XGBoost, RF and SVR (RMSEs were 0.89 t/ha, 0.92 t/ha and 1.21 t/ha), respectively. Additionally, we noticed that S2S+observation performed almost equivalently to S2S alone for XGBoost and RF, and worse for SVR, indicating that S2S atmospheric prediction outputs plus observational climate variables as model inputs do not add extra contributions. This phenomenon may be largely because the S2S atmospheric prediction outputs are generated based on a coupled prediction system of the atmosphere, ocean, land and sea ice, which is different from the observational climate data that are collected based on the conditions of the historical atmosphere. Therefore, the overlapping of variables from different systems will result in them restricting each other and will ultimately worsen the model performance for the ML methods. For MLR, the performance of S2S atmospheric prediction outputs was the worst among the three groups of input variables, followed by observational climate, and then S2S+observation, which had the most input data, achieved the highest R^2, which was completely different to the results obtained by the ML-based models. One possible explanation for this is that the MLR method essentially captures the correlation between input variables and yield (Bouras et al., 2021), and the correlation between historical climate observations and yield is much greater than that between S2S atmospheric prediction and yield (Figure S1). In addition, the ML models (i.e. XGBoost, RF and SVR) outperformed the linear method (i.e. MLR) due to the fact that most relationships between yield and different variables are nonlinear, and ensemble learning methods are better able to capture these relationships than the linear method (Cai et al., 2019). We also evaluated the contribution of EVI to the candidate models (Figure S2), the results of which show that EVI contributed little to the predictive capability because climate variables act on the whole growing season scale while EVI is only present in specific months, such as the jointing stage and heading stage (Cao et al., 2021). The performance of EVI for yield prediction was insignificant for predictions based on the whole growing season (Figure S3).

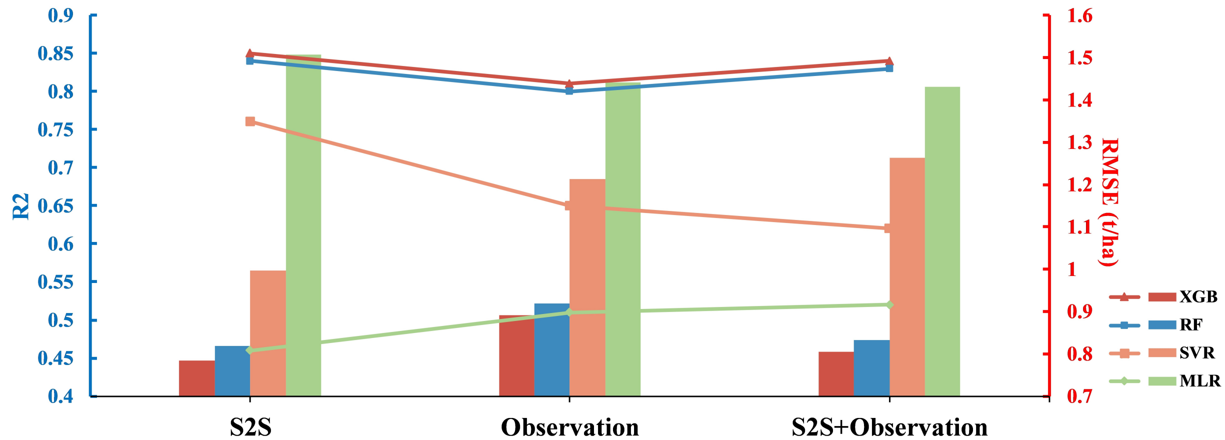

Figure 1. Model performance of the three groups of variables. The lines represent the prediction R^2 of the four models (XGBoost, RF, SVR and MLR), while the bars denote the RMSE (unit: t/ha) of the predicted yield.

4.2 Leave-one-year-out cross-validation

We conducted leave-one-year-out cross-validation, only using XGBoost and RF, to further validate the model performance with S2S atmospheric prediction outputs and observational climate data, respectively, as they showed the best prediction skill among the four candidate models. All R^2 (Table S3 and Table S4) and RMSE (Table S5 and Table S6) values for each testing year were averaged by 10 repeated predictions. In general, both XGBoost and RF demonstrated good forecasting capability, with an average R^2 greater than 0.8 and RMSE less than or equal to 0.91 (t/ha). The S2S atmospheric prediction outputs exhibited a consistently higher R^2 and lower RMSE (mean R^2 and RMSE ranging from 0.836 to 0.853 and from 0.784 to 0.823 t/ha, respectively) than the observational climate data (mean R^2 and RMSE ranging from 0.802 to 0.815 and from 0.881 to 0.91 t/ha, respectively) in predicting winter wheat yield. Overall, the S2S atmospheric prediction outputs were superior to the observational climate data in yield forecasting, which could explain about 84% of yield variations for winter wheat using ML methods (average R^2 of 0.836–0.853).

4.3 Spatial patterns of yield estimation at grid level

The spatial distributions of predicted yields are also provided (Figure 2 and Figures S4–S6). As in section 4.2, we only evaluate the predictions made by XGBoost and RF, using the S2S atmospheric prediction and observational climate data as training inputs, respectively. Figure 2 and Figure S4 show the spatial patterns of predicted yield by XGBoost, and Figure S5 and Figure S6 show the results predicted by RF. In general, the spatial patterns of the predicted yield were consistent with the recorded yield. The high-yield grids were mainly concentrated in the east of the planting area, while the grids with low yield were mainly distributed in the west. As shown in Figure 2, some high-yield grids in the east were slightly underestimated or overestimated, while some high-yield grids in the west were also underestimated, especially for the years 2013 and 2010. The prediction yields obtained by RF in Figure S5 and Figure S6 are roughly similar to those of XGBoost.

To further compare the performances of different sources of input data, we present the mean prediction PE of winter wheat yield predicted by XGBoost. It was found that grids with low yield tended to have larger PE (25%) (Figures 3a and 3b). The PE differences between S2S and observation were obtained by the result trained with observational climate data to minus that trained with S2S atmospheric prediction outputs (Figure 3c). The model performance from integrating different sources of meteorological data also corresponded with crop yield. In other words, the high-yield grids using the S2S-trained model achieved a satisfactory R^2 compared to the low-yield grids; whereas, in contrast, the combination of low-yield grids and model trained with observational climate data achieved better predictive precision.

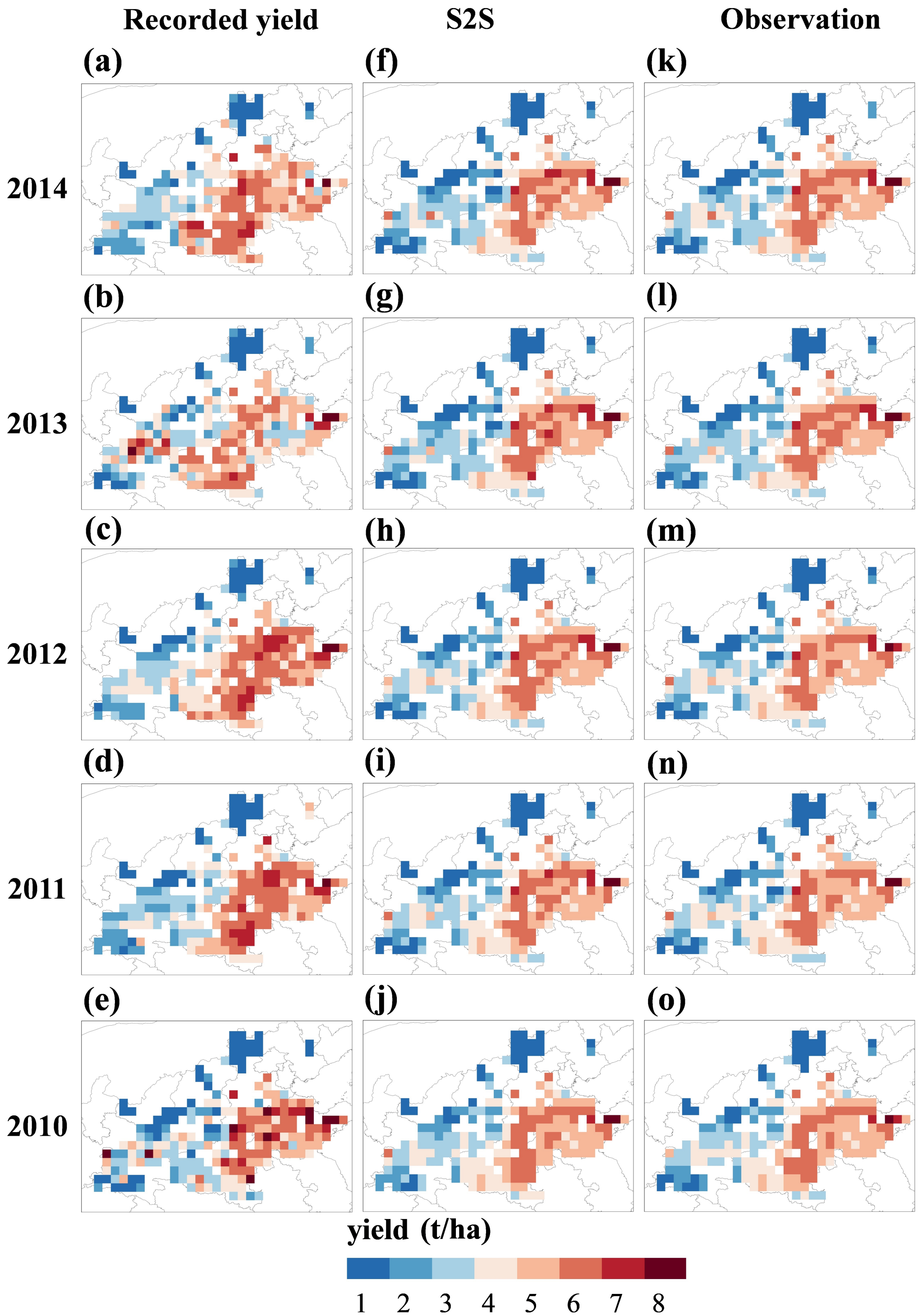
Recorded yield
S2S
Observation
(a)
(f)
(k)
2014
(b)
(g)
(l)
2013
(c)
(h)
(m)
2012
(d)
(i)
(n)
2011
(e)
(j)
(o)
2010
yield (t/ha)
1
2
3
4
5
6
7
8

Figure 2. Spatial patterns of the yield predicted with XGBoost (2010–2014). The first column is the recorded yield, the second column is the simulated yield using the XGBoost model trained with S2S atmospheric prediction, and the third column is the simulated yield using the XGBoost model trained with observational climate data.

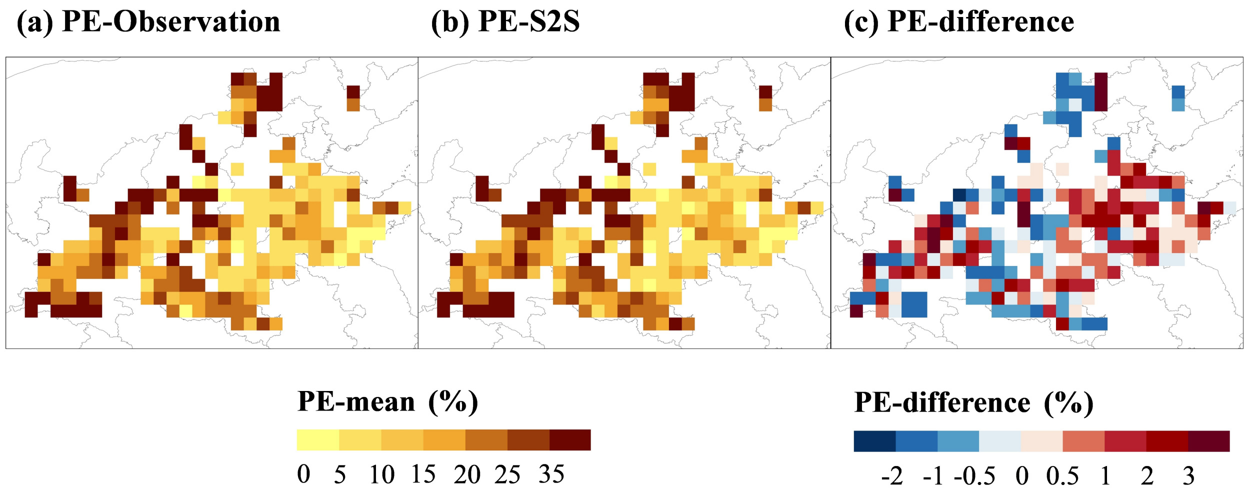

Figure 3. Spatial patterns of the mean PE for XGBoost: (a) PE calculated by XGBoost trained with observational climate; (b) PE calculated by XGBoost trained with S2S atmospheric prediction; (c) PE difference between (a) and (b) (former minus the latter). Positive values indicate that the model trained with observational climate data has greater error, while negative values indicate that the model trained with S2S has greater error.

4.4 Optimum lead time of yield forecasting

In order to investigate the limit of making a trustworthy early yield prediction, we conducted an in-season prediction experiment on winter wheat yield using XGBoost and RF. The in-season predictions were made in monthly intervals, which means we added the climate information for a more recent month at each time step. In general, better performance of the model was achieved with more input data, i.e., as the prediction period approached the end of the growing season, the R^2 increased and the RMSE decreased (Li et al., 2021). As shown in Figure 4, XGBoost and RF performed comparably in terms of predictive capability for yield forecasting, sharing the same trend. Although RF achieved a stable R^2 and RMSE earlier than XGBoost, XGBoost had a higher R^2 and lower RMSE than RF. Overall, the performance of both XGBoost and RF increased with the accumulation of data. In terms of XGBoost, the R^2 ranged from 0.77 to 0.85 and 0.75 to 0.81 (RMSE from 0.97 to 0.78 t/ha and 1.01 to 0.89 t/ha) when S2S-trained and climate-observational-data-trained, respectively (Figure 4); while for RF, the R^2 ranged from 0.78 to 0.84 and 0.75 to 0.8 (RMSE from 0.94 to 0.82 t/ha and 1.03 to 0.9 t/ha) for the two models, respectively (Figure

4). The results confirmed that the S2S atmospheric predictions were superior to the observational climate variables as model inputs, and XGBoost was slightly superior to RF in predicting winter wheat yield.

Regarding the optimum forecasting time, the models trained by S2S atmospheric prediction (XGBoost and RF) reached a stable R^2 of 0.85 and 0.84 in February and January, respectively, with a lead time of about four months before harvesting of winter wheat (Figure 4a). By contrast, the models using observational climate data resulted in the highest R^2 of 0.81 and 0.8 in March and January for XGBoost and RF, respectively, with a lead time of three months (Figure 4b). Our work achieves a more satisfactory lead time prior to harvest compared with previous studies (Cai et al., 2019; Guo et al., 2021; Li et al., 2021). The findings demonstrate that S2S atmospheric predictions have the potential to achieve an earlier optimum yield forecast than observational climate data as model inputs.

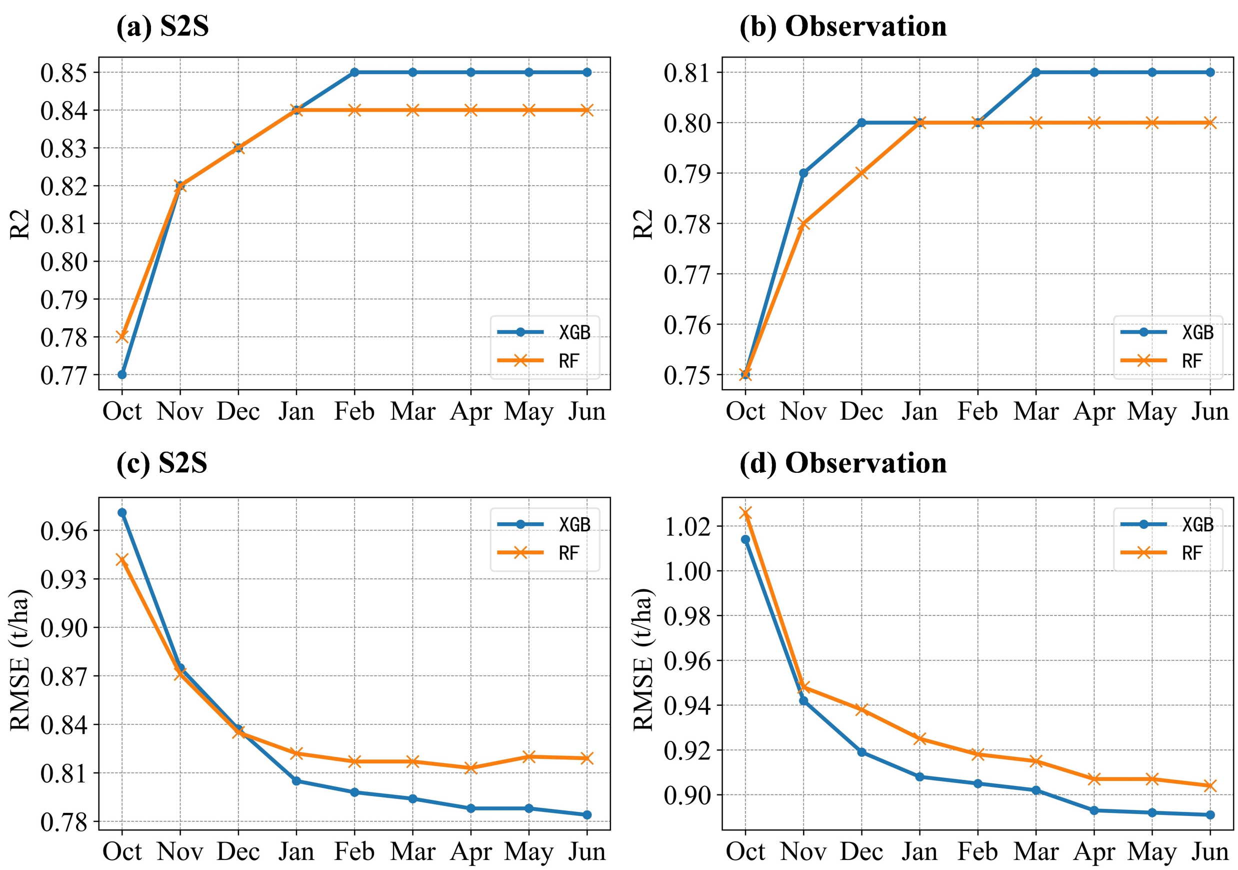

Figure 4. Temporal progression of model performance (R^2 and RMSE) based on the four models: (a, c) models trained with S2S atmospheric predictions; (b, d) models trained with observational climate variables.

5 Discussion

5.1 Model performance

In this work, we have further developed the four most commonly used models (XGBoost, RF, SVR, MLR) for winter wheat yield forecasting. Factors affecting crop yield are complex and diverse, so the combination of multi-source variables can be considered in further studies (Cao et al., 2021; Liu et al., 2022). Linear regression generally explains the linear relationship between yield and input variables. In this study, the MLR method performed the worst for simulating yield among the developed models, which simply confirmed this point. ML algorithms have shown strong predictive capability, especially for XGBoost and RF. Thus, ML methods can clearly achieve superior performance in capturing the complex relationships between S2S atmospheric predictions and yield, and they provide a window of opportunity to predict crop yield at regional or even global scales. Here, we built yield prediction models for winter wheat within the framework of ML, due to its simplicity and efficiency. However, deep learning, with its higher accuracy at the cost of computational intensity and model complexity, may have great potential for improving global grain yield forecasting, and it is essential to find a balance between complexity and efficiency (Reichstein et al., 2019). The hybridization of process-based crop growth models and/or ML with deep learning as well multi-source data (Cao et al., 2021) has the potential to offer an improved and optimized technique for yield forecasting (Feng et al., 2020; Shahhosseini et al., 2021).

5.2 Potential of S2S atmospheric prediction systems

Our findings demonstrate that the S2S atmospheric predictions outperformed those based on observational climate data. This may be largely due to S2S dynamical atmospheric predictions being able to provide more information over the simple use of observational climate data to simulate possible future scenarios, confronted with the climate change together with climate teleconnections between a region of interest and other parts of the globe (Brown et al., 2018; Ogutu et al., 2018). Compared with previous studies, the in-season prediction showed that the winter wheat yield can be forecasted with a 3–4-month lead time, and that has increased as the S2S atmospheric prediction outputs can normally be available. These findings show that the inclusion of S2S atmospheric predictions improve the yield prediction performance. More S2S dynamic prediction models from various institutions (e.g., FGOALS-f2, ECMWF, NECP-CFSv2 etc.) could be used for comparison to improve yield forecasting (Feng et al., 2020; Ren et al., 2019). Thus, the potential of S2S atmospheric prediction should be extensively exploited. In further research, deep learning methods should be considered to improve the accuracy of S2S atmospheric prediction (Ham et al., 2019; Yu & Ma, 2021). Besides, the horizontal resolution of S2S atmospheric prediction systems should be increased to a convection-permitting resolution ($<$ 10 km) in order to provide crop yield predictions at finer spatial scales.

6 Conclusions

The aim of this study was to establish a forecasting model (MHCF v1.0) for the major production areas of winter wheat in northern China using ML driven by an S2S atmospheric prediction system. To this end, we designed several experiments to test the model performance and compare the predictive performance of S2S atmospheric predictions, observational climate data, and the combination of meteorological and remote sensing data. Ultimately, a skilled prediction one season before harvest was achieved. These experiments illustrate that S2S atmospheric predictions are superior in their forecasting performance in terms of crop yield, as compared with observational climate data. Moreover, the study demonstrates that ML methods, especially ensemble learning models, perform significantly better than linear regression–based methods. This research proves that MHCF v1.0 is a novel and promising technique for regional yield prediction, and it is our hope that it can be extended to crop yield forecasts in other regions and even at the global scale.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China program (41701111). The authors acknowledge the ArcGIS software provided by Esri and the Python software provided by the Python Software Foundation.

Open Research

The datasets that support our work are as follows. The FGOALS-f2 S2S dynamical atmospheric prediction outputs were provided by the Institute of Atmospheric Physics, Chinese Academy of Sciences (https://apps.ecmwf.int/datasets/data/s2s-realtime-daily-averaged-anso/levtype=sfc/type=cf/). The winter wheat yield data were obtained from the Ministry of Agriculture of China (https://doi.org/10.6084/m9.figshare.18093674). The winter wheat planting areas were used to mask the winter wheat yields (https://data.mendeley.com/datasets/jbs44b2hrk/2). EVI data were derived from MOD13C1 (Collection 6) records (https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD13C1). Historical observational climate variables were obtained from the CRU (https://crudata.uea.ac.uk/cru/data/hrg/).

References

Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, *34*(10), 3001–3023. https://doi.org/10.1002/joc.3887Benami, E., Jin, Z., Carter, M. R., Ghosh, A., Hijmans, R. J., Hobbs, A., et al. (2021). Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth and Environment*, *2*(2), 140–159. https://doi.org/10.1038/s43017-020-00122-yBolton, D. K., & Friedl, M. A. (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology*, *173*, 74–84. https://doi.org/10.1016/j.agrformet.2013.01.007Bouras, E. H., Jarlan, L., Er-Raki, S., Balaghi, R., Amazirh, A., Richard, B., & Khabba, S. (2021). Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in morocco. *Remote Sensing*, *13*(16). https://doi.org/10.3390/rs13163101Brown, J. N., Hochman, Z., Holzworth, D., & Horan, H. (2018). Seasonal climate forecasts provide more definitive and accurate crop yield predictions. *Agricultural and Forest Meteorology*, *260–261*(February), 247–254. https://doi.org/10.1016/j.agrformet.2018.06.001Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*(February), 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010Cao, Juan, Zhang, Z., Tao, F., Zhang, L., Luo, Y., Zhang, J., et al. (2021). Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches. *Agricultural and Forest Meteorology*, *297*(November 2020), 108275. https://doi.org/10.1016/j.agrformet.2020.108275Cao, Juan, Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., & Tao, F. (2021). Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine. *European Journal of Agronomy*, *123*(March 2020), 126204. https://doi.org/10.1016/j.eja.2020.126204Cao, Junjun, An, Q., Zhang, X., Xu, S., Si, T., & Niyogi, D. (2021). Is satellite Sun-Induced Chlorophyll Fluorescence more indicative than vegetation indices under drought condition? *Science of the Total Environment*, *792*. https://doi.org/10.1016/j.scitotenv.2021.148396Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, *11*, 2079–2107.Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., et al. (2015). Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. *Agricultural and Forest Meteorology*, *206*, 137–150. https://doi.org/10.1016/j.agrformet.2015.03.007Cole, M. B., Augustin, M. A., Robertson, M. J., & Manners, J. M. (2018). The science of food security. *Npj Science of Food*, *2*(1), 1–8. https://doi.org/10.1038/s41538-018-0021-9FAO, IFAD, UNICEF, WFP, & WHO. (2021). *The State of Food Security and Nutrition in the World 2021: Transforming food systems for food security, improved nutrition and affordable healthy diets for all. Fao.*Feng, P., Wang, B., Liu, D. L., Waters, C., Xiao, D., Shi, L., & Yu, Q. (2020). Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. *Agricultural and Forest Meteorology*, *285–286*(January), 107922. https://doi.org/10.1016/j.agrformet.2020.107922Feng, X., Klingaman, N., Zhang, S., & Guo, L. (2020). Building sustainable science partnerships between early-career researchers to better understand and predict east asia water cycle extremes. *Bulletin of the American Meteorological Society*, *101*(6), E785–E789. https://doi.org/10.1175/BAMS-D-20-0039.1Franch, B., Vermote, E. F., Becker-Reshef, I., Claverie, M., Huang, J., Zhang, J., et al. (2015). Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment*, *161*, 131–148. https://doi.org/10.1016/j.rse.2015.02.014Goldberg, David E.; Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine Learning*, *3*(2), 95–99.Guo, Y., Fu, Y., Hao, F., Zhang, X., Wu, W., Jin, X., et al. (2021). Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, *120*(June 2020), 106935. https://doi.org/10.1016/j.ecolind.2020.106935Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, *573*(7775), 568–572. https://doi.org/10.1038/s41586-019-1559-7Han, J., Zhang, Z., Cao, J., Luo, Y., Zhang, L., Li, Z., & Zhang, J. (2020). Prediction of winter wheat yield based on multi-source data and machine learning in China. *Remote Sensing*, *12*(2). https://doi.org/10.3390/rs12020236Huang, J. kun, WEI, W., CUI, Q., & XIE, W. (2017). The prospects for China's food security and imports: Will China starve the world via imports? *Journal of Integrative Agriculture*, *16*(12), 2933–2944. https://doi.org/10.1016/S2095-3119(17)61756-8Iizumi, T., Shin, Y., Kim, W., Kim, M., & Choi, J. (2018). Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Climate Services*, *11*, 13–23. https://doi.org/10.1016/j.cliser.2018.06.003Jiang, Z., Liu, C., Ganapathysubramanian, B., Hayes, D. J., & Sarkar, S. (2020). Predicting county-scale maize yields with publicly available data. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-71898-8van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, *177*(July), 105709. https://doi.org/10.1016/j.compag.2020.105709Kostková, M., Hlavinka, P., Pohanková, E., Kersebaum, K. C., Nendel, C., Gobin, A., et al. (2021). Performance of 13 crop simulation models and their ensemble for simulating four field crops in Central Europe. *Journal of Agricultural Science*, *159*(1–2), 69–89. https://doi.org/10.1017/S0021859621000216Li, J., Bao, Q., Liu, Y., Wu, G., Wang, L., He, B., et al. (2021(a)). Dynamical seasonal prediction of tropical cyclone activity using the fgoals-f2 ensemble prediction system. *Weather and Forecasting*, *36*(5), 1759. https://doi.org/10.1175/WAF-D-20-0189.1Li, J., Bao, Q., Liu, Y., Wang, L., Yang, J., Wu, G., et al. (2021(b)). Effect of horizontal resolution on the simulation of tropical cyclones in the Chinese Academy of Sciences FGOALS-f3 climate system model. *Geoscientific Model Development*, *14*(10), 6113–6133. https://doi.org/10.5194/gmd-14-6113-2021Li, L., Wang, B., Feng, P., Wang, H., He, Q., Wang, Y., et al. (2021). Crop yield forecasting and associated optimum lead time analysis based on multi-source environmental data across China. *Agricultural and Forest Meteorology*, *308–309*(June). https://doi.org/10.1016/j.agrformet.2021.108558Li Sanai, David Fleisher, Dennis Timlin, Vangimella R.Reddy, Z. W., & Mcclung, A. (2020). Evaluation of Different Crop Models for Simulating Rice Development and Yield in the U.S. Mississippi Delta. *Agronomy*, 1–21.Liaqat, M. U., Cheema, M. J. M., Huang, W., Mahmood, T., Zaman, M., & Khan, M. M. (2017). Evaluation of MODIS and Landsat multiband vegetation indices used for wheat yield estimation in irrigated Indus Basin. *Computers and Electronics in Agriculture*, *138*, 39–47. https://doi.org/10.1016/j.compag.2017.04.006Liu, Yuan, Li, N., Zhang, Z., Huang, C., Chen, X., & Wang, F. (2020). The central trend in crop yields under climate change in China: A systematic review. *Science of the Total Environment*, *704*. https://doi.org/10.1016/j.scitotenv.2019.135355Liu, Yuanyuan, Wang, S., Wang, X., Chen, B., Chen, J., Wang, J., et al. (2022). Exploring the superiority of solar-induced chlorophyll fluorescence data in predicting wheat yield using machine learning and deep learning methods. *Computers and Electronics in Agriculture*, *192*(November 2021), 106612. https://doi.org/10.1016/j.compag.2021.106612Luo, Y., Zhang, Z., Li, Z., Chen, Y., Chen, Y., Zhang, L., et al. (2020). Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources. *Environmental Research Letters*, *15*(7). https://doi.org/10.1088/1748-9326/ab80f0Ma, Y., Zhang, Z., Kang, Y., & Özdoğan, M. (2021). Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach. *Remote Sensing of Environment*, *259*. https://doi.org/10.1016/j.rse.2021.112408Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: A comparison of resampling methods. *Bioinformatics*, *21*(15), 3301–3307. https://doi.org/10.1093/bioinformatics/bti499Ogutu, G. E. O., Franssen, W. H. P., Supit, I., Omondi, P., & Hutjes, R. W. A. (2018). Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts. *Agricultural and Forest Meteorology*, *250–251*(January), 243–261. https://doi.org/10.1016/j.agrformet.2017.12.256Pagani, V., Stella, T., Guarneri, T., Finotto, G., van den Berg, M., Marin, F. R., et al. (2017). Forecasting sugarcane yields using agro-climatic indicators and Canegro model: A case study in the main production region in Brazil. *Agricultural Systems*, *154*(March), 45–52. https://doi.org/10.1016/j.agsy.2017.03.002Pan, Y., Li, L., Zhang, J., Liang, S., Zhu, X., & Sulla-Menashe, D. (2012). Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sensing of Environment*, *119*, 232–242. https://doi.org/10.1016/j.rse.2011.10.011Peng, B., Guan, K., Pan, M., & Li, Y. (2018). Benefits of Seasonal Climate Prediction and Satellite Data for Forecasting U.S. Maize Yield. *Geophysical Research Letters*, *45*(18), 9662–9671. https://doi.org/10.1029/2018GL079291Prosekov, A. Y., & Ivanova, S. A. (2018). Food security: The challenge of the present. *Geoforum*, *91*(February), 73–77. https://doi.org/10.1016/j.geoforum.2018.02.030Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications*, *3*, 1–7. https://doi.org/10.1038/ncomms2296Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1Rembold, F., Meroni, M., Urbano, F., Royer, A., Atzberger, C., Lemoine, G., et al. (2015). Remote sensing time series analysis for crop monitoring with the SPIRITS software: New functionalities and use examples. *Frontiers in Environmental Science*, *3*(JUL), 1–11. https://doi.org/10.3389/fenvs.2015.00046Ren, H. L., Wu, Y., Bao, Q., Ma, J., Liu, C., Wan, J., et al. (2019). The China Multi-Model Ensemble Prediction System and Its Application to Flood-Season Prediction in 2018. *Journal of Meteorological Research*, *33*(3), 540–552. https://doi.org/10.1007/s13351-019-8154-6Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Scientific Reports*, *11*(1), 1–16. https://doi.org/10.1038/s41598-020-80820-1Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2017). The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, *98*(1), 163–173. https://doi.org/10.1175/BAMS-D-16-0017.1Vitart, Frédéric, & Robertson, A. W. (2018). The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *Npj Climate and Atmospheric Science*, *1*(1), 1–7. https://doi.org/10.1038/s41612-018-0013-0Wang, M., Tao, F. L., & Shi, W. J. (2014). Corn yield forecasting in northeast china using remotely sensed spectral indices and crop phenology metrics. *Journal of Integrative Agriculture*, *13*(7), 1538–1545. https://doi.org/10.1016/S2095-3119(14)60817-0Wu, B., Meng, J., Li, Q., Yan, N., Du, X., & Zhang, M. (2014). Remote sensing-based global crop monitoring: Experiences with China's CropWatch system. *International Journal of Digital Earth*, *7*(2), 113–137. https://doi.org/10.1080/17538947.2013.821185Yu, S., & Ma, J. (2021). Deep Learning for Geophysics: Current and Future Trends. *Reviews of Geophysics*, *59*(3), 1–36. https://doi.org/10.1029/2021RG000742Zhang, J., Feng, L., & Yao, F. (2014). Improved maize cultivated area estimation over a large scale combining MODIS-EVI time series data and crop phenological information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *94*, 102–113. https://doi.org/10.1016/j.isprsjprs.2014.04.023Zhou, Y., Yang, B., Chen, H., Zhang, Y., Huang, A., & La, M. (2019). Effects of the Madden–Julian Oscillation on 2-m air temperature prediction over China during boreal winter in the S2S database. *Climate Dynamics*, *52*(11), 6671–6689. https://doi.org/10.1007/s00382-018-4538-z