
# Early Forcasting of Maize Lodging Risk through Multi-period and Multi-source Data Integration

**Lingwei Dong[*1], Yixin Miao[2], Xinbing Wang[3], Pete Berry[4], Damian Hatley[4], Krzysztof Kusnierek[5]**

1. Institute of Tobacco Research, Chinese Academy of Agricultural Sciences, Qingdao 266101, China
2. Precision Agriculture Center, Department of Soil, Water and Climate, University of Minnesota, St. Paul., MN 55108, USA
3. Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
4. ADAS High Mowthorpe, Malton, North Yorkshire YO178BP, UK
5. Center for Precision Agriculture, Norwegian Institute of Bioeconomy Research (NIBIO), Nylinna 226, 2849 Kapp, Norway

**A paper from the Proceedings of the**
**17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture**
**13-15 July 2026**
**Porto Alegre, Brazil**

***Abstract.***

*Lodging is a critical constraint on global maize (Zea mays L.) productivity, primarily through detrimental effects on both grain yield and quality. However, reliable methods to predict maize lodging risk early in the growing season are lacking, which hinders timely implementation of effective agronomic management interventions to increase crop lodging resistance and reduce corresponding yield losses. This work aimed to develop a feasible early season maize lodging risk prediction method by combining proximal sensing, crop growth modeling, and machine learning (ML) using multi-source data fusion. From 2017~2019, maize nitrogen (N) and planting population trials were conducted in China and United Kingdom with contrasting agroecological conditions. Maize growth status information derived from the integration of proximal sensing and crop growth modeling, along with information on management practices, key growth stage dates and environmental factors, were aggerated into multi-period datasets aligned with agronomically critical growth stages and then served as feature variables for developing lodging risk prediction models using ML algorithms. Out of three ML methods tested in this study, random forest regression (RFR) performed the best overall, with stable validation ($R^2$ = 0.71~0.81, RMSE = 0.47~0.67, REr = 6.38~7.98) and satisfactory probability density distribution. Feature importance analysis performed across the tested predictors identified crop growth status and management practices, particularly planting density and N application rate, as critical contributors to lodging risk prediction. Lodging risks can be predicted reliably at the V8 growth stage by combining proximal crop sensing, crop growth modeling and ML based on multi-source data fusion. This study developed an integrated and innovative strategy for early prediction of maize lodging risks,*

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. Dong, L., Miao, Y., Wang, X., Berry, P., Hatley, D. & Kusnierek, K. (2026). Early forcasting of maize lodging risk through multi-period and multi-source data integration. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

*which has the potential to enable precision lodging risk management to enhance lodging resistance and minimize maize yield losses. Future research should focus on expanding data collection across diverse environmental conditions to refine feature selection, simplify prediction models, and improve generalizability, thereby strengthening the robustness and spatial adaptability of this strategy.*

***Keywords.***
*Data fusion, precision management, feature aggregation, early prediction, lodging risk*

## Instruction

Maize (*Zea mays* L.) plays a critical role in global food security, animal feed, and biofuel production (Tanumihardjo et al., 2020). Increasing planting density and N application are widely used to maximize grain yield, but the consequent increase in lodging risk is often overlooked (Xue et al., 2017). Lodging causes permanent stem displacement, reducing canopy photosynthesis and grain filling, leading to severe yield losses (Shah et al., 2017; Gonzalez et al., 2018; Zhao et al., 2020).

Recent lodging risk prediction models are largely empirical, relying on remote sensing indices (Wu et al., 2019; Li et al., 2024; Wang et al., 2024). However, lodging results from complex interactions among crop traits, weather, and management, and purely statistical models lack biological and mechanistic underpinnings. Mechanistic models exist (Berry et al., 2021) but require destructive sampling, limiting practical application. Given the strong correlation between leaf area index (LAI) and lodging risk at early growth stages, integrating crop growth information with environmental factors offers potential (Chauhan et al., 2019; Berry et al., 2021). Crop growth models can simulate dynamic growth processes, and when combined with remote sensing through data assimilation, they improve prediction accuracy (Huang et al., 2019; Jin et al., 2020). ML methods are well-suited for handling multi-source, non-linear data (Chlingaryan et al., 2018; Wang et al., 2021; Ruan et al., 2022). Thus, integrating remote sensing, growth models, environmental variables, and ML offers a promising strategy for advancing lodging risk forecasting.

Because lodging during grain-filling causes major yield losses (Li et al., 2015), this study focuses on predicting lodging risk at earlier growth stages. The objective is to develop a multi-source data fusion approach combining proximal sensing, crop growth modeling, and ML to improve early-season maize lodging risk prediction.

## Materials and methods

### Study area and experimental design

Field experiments were conducted at two contrasting sites. In Northeast China (Lishu County, Jilin Province), black soil with semi-humid continental monsoon climate. During 2018–2019, a randomized split-block design with three replicates was used for maize cultivar Liangyu 66. Main plots had four planting densities (5.5–10.0 plants $m^{-2}$) and sub-plots six N rates (0–300 kg N $ha^{-1}$). Standard agronomic practices were followed. In the UK (ADAS Gleadthorpe, Nottinghamshire), loamy sand with temperate marine climate. During 2017–2018, a two-way factorial design (three replicates) tested planting densities (4–12 plants $m^{-2}$) and N rates (0–200 kg N $ha^{-1}$) on cultivars Dualto (2017) and Ramirez (2018). Both sites received basal P and K, with N split-applied. Detailed information about the experiments is listed in Table 1.

**Table 1. Summary information of field experiments conducted in China and UK.**

| Experiment | Variety | Planting density (Plants $m^{-2}$) | N rate (kg $ha^{-1}$) | P rate (kg $ha^{-1}$) | K rate (kg $ha^{-1}$) | Plot area（$m^2$） | Sowing date | Plant sampling dates (growth stages) | Harvest date |
|---|---|---|---|---|---|---|---|---|---|
| CH18 | Liangyu 66 | 5.5, 7.0. 8.5 | 0, 60, 120, 180, 240, 300 | 90 | 90 | 108 | 28th Apr | 13th Jun (V6), 25th Jun (V8), 10 th Jul (V12), 24th Jul (VT), 16th Aug (R3) | 29th Sep |
| | | 10.0 | 60, 180, 300 | | | | | | |
| CH19 | Liangyu 66 | 5.5, 7.0. 8.5 | 0, 60, 120, 180, 240, 300 | 90 | 90 | 108 | 28th Apr | 14th Jun (V6), 25th Jun (V8), 10th Jul (V12), 24th Jul (VT), 15th Aug (R3) | 27th Sep |
| | | 10.0 | 60, 180, 300 | | | | | | |
| UK17 | Dualto | 4.0, 6.0, 12.0 | 0, 100, 200 | 55.2 | 0 | 36 | 22nd May | 9th Jun (V4), 15th Jun (V6), 22nd Jun (V8), | 11th Oct |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 7th Jul (V12), 15th Jul (VT), 4th Sep (R3) | |
| UK18 | Ramirez | 4.0, 6.0, 12.0 | 0, 100, 200 | 55.2 | 0 | 36 | 18th May | 1st Jun (V4), 15th Jun (V8), 29th Jun (V12), 10th Jul (VT), 28th Aug (R3) | 26th Sep |

## Data acquisition

In China trials, canopy reflectance was acquired at V6 and V8 using a Crop Circle ACS-430 (red, red-edge, NIR). In UK trials, a CROPSCAN Multispectral Radiometer (MSR) radiometer was used at V4, V8, and V12, from which the three corresponding bands were extracted. Plot-level vegetation indices (VIs) were calculated following Dong et al. (2023).

For growth indicators, plant height (PH), aboveground biomass (AGB), and LAI were measured on each sampling date listed in Table 1. In China trials, leaf area was calculated using differentiated coefficients: 0.75 and 0.5 multipliers were applied to fully expanded and remaining leaves, respectively. UK trials used a leaf area meter. Plot-level LAI was derived by multiplying per-plant leaf area by planting density. AGB was determined after oven-drying to constant mass and expressed as t ha$^{-1}$.

At the grain filling stage (R3), lodging-related parameters were measured, including natural frequency (via oscillation test), height of the center of gravity, internode length and diameter, stem breaking strength (three-point bending test), and wall width. Stem failure moment (*Bs*) was calculated as $Bs = (\mathrm{F} \times \mathrm{L}) / 4$. Using these traits, lodging risk (failure wind speed, FWS) for three internode positions (basal internodes 1–3, internode 5, and internode 8) was estimated via a comprehensive maize lodging model (Berry et al., 2021).

Soil samples from different layers (0–20, 20–40, 40–60, 60–80 cm for China; 0–30, 30–60, 60–90 cm for UK) were analyzed for physical and chemical properties, particle composition, and hydrodynamic parameters. Weather data (daily temperature, precipitation, solar radiation) were obtained from local stations. Derived features included growing degree days (GDD), accumulated precipitation (APP), precipitation evenness (SDI), and abundant well-distributed rainfall (AWDR). Abbreviations and descriptions of the weather factors are given in Table 2.

**Table 2 Summary of the weather variables used in this study.**

| Abbreviation | Description | Unit | Calculation | Reference |
|---|---|---|---|---|
| Tmax | Maximum temperature | ℃ | Maximum temperature in given number of days | |
| Tmin | Minimum temperature | ℃ | Minimum temperature in given number of days | |
| Tave | Average temperature | ℃ | Average temperature in given number of days | |
| GDD | Growing degree days | ℃ | (Tmax + Tmin)/2-Tbase, Tbase = 10 ℃. | |
| APP | Accumulated precipitation | mm | Sum of daily precipitation in given number of days | |
| SDI | Shannon diversity index of precipitation evenness | | $-\sum pi\frac{\ln(pi)}{\ln(n)}$, pi = precipitation/APP, n = number of days in given time | Tremblay et al. (2012) |
| AWDR | Abundant and well-distributed rainfall | | SDI × APP | Tremblay et al. (2012) |

## Maize growth status simulation

The decision support system for agrotechnology transfer (DSSAT)-based CERES-Maize model was used to simulate daily crop growth status. Inputs included weather data (daily temperature, precipitation, solar radiation), field management (sowing, fertilization, irrigation, harvesting), soil properties (from laboratory analyses), and cultivar genetic coefficients. Simulated growth indicators (PH, AGB, LAI) were obtained from model runs.

Pearson correlation identified the top ten vegetation indices (VIs) most related to each growth indicator. Using a calibration (n = 240) and validation (n = 120) split, RFR models were developed, incorporating days after planting (DAP) as a temporal predictor alongside the selected VIs to account for phenological dynamics.

The DSSAT-CERES-Maize model was calibrated using particle swarm optimization (PSO). For the China dataset, calibration followed Dong et al. (2024); for the UK dataset, initial parameter ranges were from the same study. PSO assimilated proximal sensing-based crop estimates into the model to minimize discrepancies, improving prediction accuracy and filling data gaps at stages lacking remote sensing measurements (e.g., V12, VT in China; V6, VT in UK).

### Lodging risk prediction modeling

Predictors included incremental crop growth, agronomic management (density, N rate), weather, topsoil properties (pH, N, P, K), and phenological indicators (DAP, DOY). Response variables were failure wind speeds (FWS) at three internode positions. To capture temporal dynamics, data were aggregated into four phenological phases (planting–V6, V6–V8, V8–V12, V12–VT). Two predictive strategies were used: Strategy 1 (at V8, using phases P-6 and 6-8) and Strategy 2 (at V12, additionally including phase 8-12). Separate lodging risk models were built for each internode. The dataset was split into training (n = 129) and testing (n = 51). Three machine learning algorithms (Support Vector Regression (SVR), RFR, Gradient Boosted Decision Tree Regression (GBDT)) were implemented in Python.

Linear regression between observed and predicted lodging risk was assessed using $R^2$, RMSE, and relative error (REr). Kernel density estimation compared data distributions, guiding model selection. Spatial differences (observed vs. predicted) at plot level were visualized per site-year. Feature importance was evaluated using RFR.

## Results

### Growth status information acquisition

The top ten VIs correlated best with AGB ($R^2$ = 0.65–0.70), followed by LAI (0.62–0.67) and PH (0.51–0.56) (Fig 1). Using these VIs together with DAP, planting density, and N rate, RFR models integrating proximal sensing data and crop growth model simulations estimated PH, LAI, and AGB with high accuracy (all $R^2$ > 0.94), as most points clustered near the 1:1 line (Fig 2).

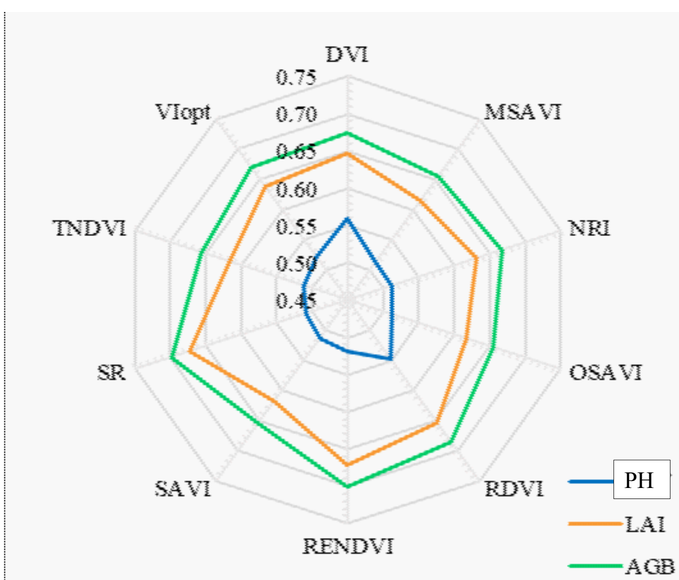


**Fig 1. Coefficient of determination (R2) for correlations between the ten top-performing VIs and maize crop growth indicators based on Pearson correlation analysis**

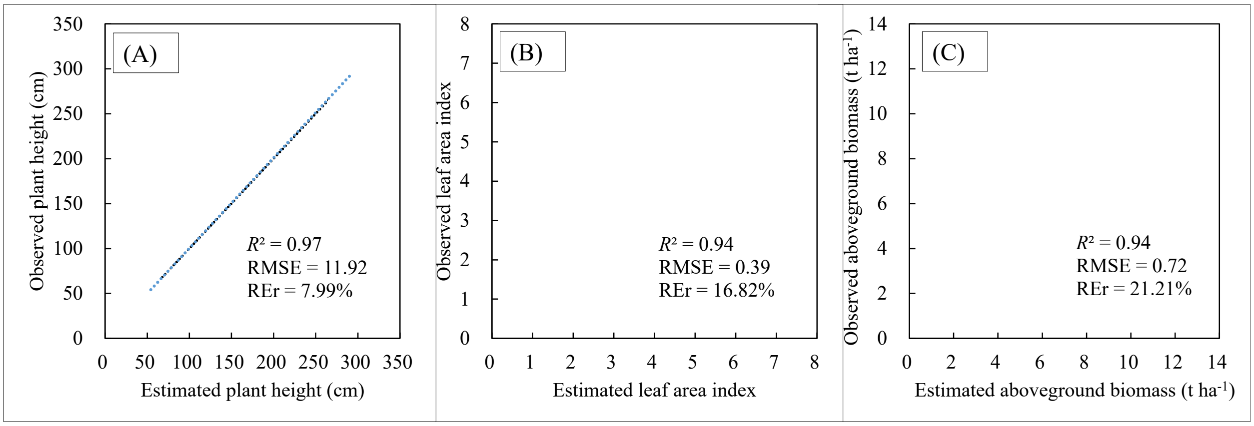


**Fig 2. Scatter plots of observed and estimated maize growth status indicators: (A) plant height, (B) leaf area index and (C) aboveground biomass. The dotted line represents the 1:1 line.**

## Lodging prediction in the early season

Models for predicting lodging risk (failure wind speed, FWS) were developed using three ML algorithms at two predictive timepoints (V8 and V12), with outcomes shown in Fig 3. Both strategies achieved good performance, with validation metrics of $R^2$ = 0.63–0.81, RMSE = 0.47–0.75 $m \cdot s^{-1}$, and REr = 6.38–8.88%. Strategy 2 (V12) slightly outperformed Strategy 1 (V8) in 67% of cases, though the accuracy difference was minor. Among the three algorithms, RFR consistently performed best ($R^2$ = 0.71–0.81, RMSE = 0.47–0.67 $m \cdot s^{-1}$, REr = 6.38–7.98%), followed by GBDT, while SVR showed the lowest accuracy ($R^2$ = 0.63–0.69, RMSE = 0.56–0.75 $m \cdot s^{-1}$, REr = 7.17–8.88%).

To further compare model performance, kernel density curves were plotted (Fig 4). In most cases, predicted FWS distributions from RFR models closely aligned with observations, confirming the stability of RFR for lodging risk prediction. Based on these results, scatter diagrams of observed vs. predicted FWS using the best-performing models (RFR) are shown in Fig 5. Overall, an acceptable linear fit was observed, with scatter points clustered near the 1:1 line for all internode positions. Notably, prediction accuracy was highest for the 5th internode (higher $R^2$, lower RMSE and REr), though models for basal internodes (1–3) and internode 8 also performed well.

Spatial distribution maps of prediction differences for validation plots (Fig 6) were generated to further assess model accuracy. Most predicted values were close to measured values, with small differences visualized by color-coding. Except for a single outlier in the UK17 dataset (Fig 6A), all differences were less than 1.33 $m \cdot s^{-1}$ (FWS1-3), 1.15 $m \cdot s^{-1}$ (FWS5), and 1.83 $m \cdot s^{-1}$ (FWS8). Spatial error mapping further revealed that approximately 69%, 69%, and 71% of predictions deviated by <0.5 $m \cdot s^{-1}$ from measured values for the three internode positions, respectively.

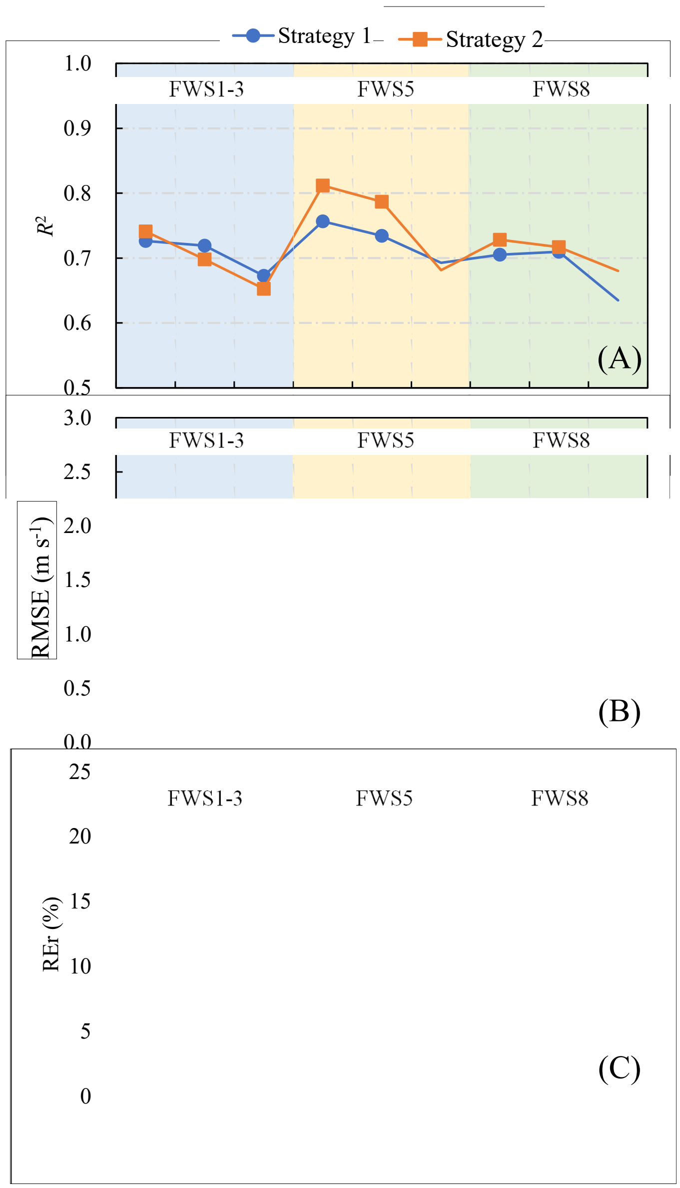


**Fig 3. Validation metrics (A) ($R^2$, (B) RMSE, (C) REr of failure wind speed (FWS) prediction models, at three different internode positions, based on different ML algorithms (RFR, GBDT, SVR) and two sampling strategies (Strategy 1 at V8 and Strategy 2 at V12).**

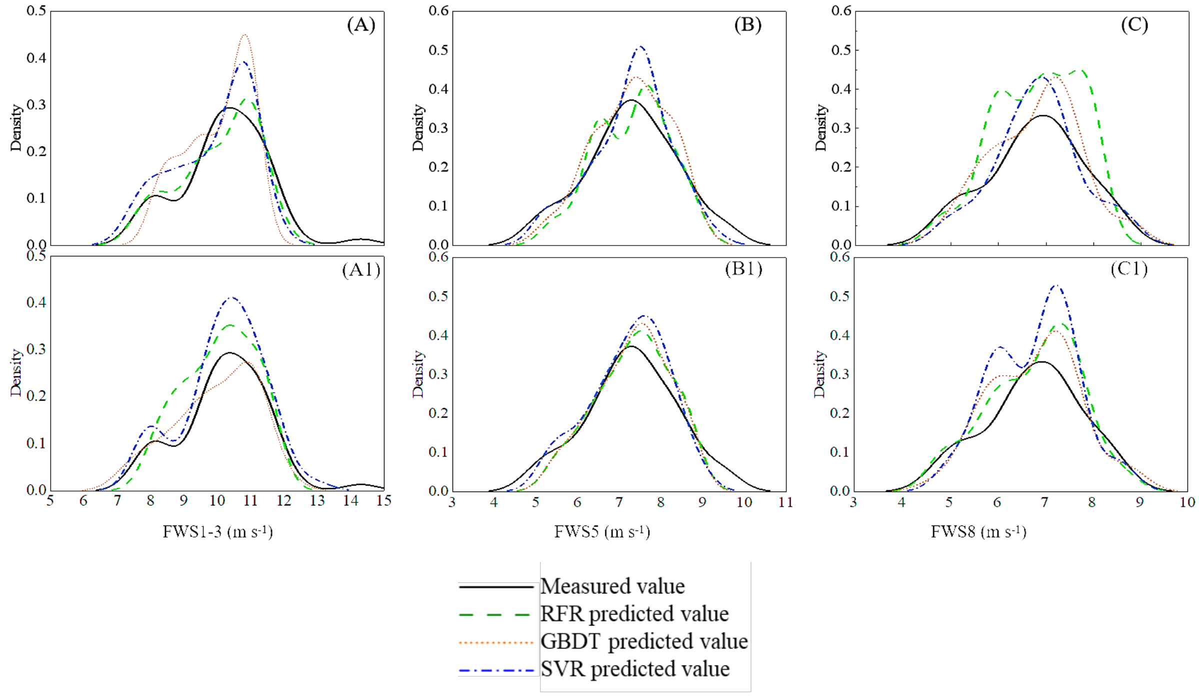


**Fig 4. Kernel density curve plots of failure wind speed (FWS) measured and predicted using three ML methods (RFR, GBDT and SVR), based on two strategies, (A-C) Strategy 1, and (A1-C1) Strategy 2.**

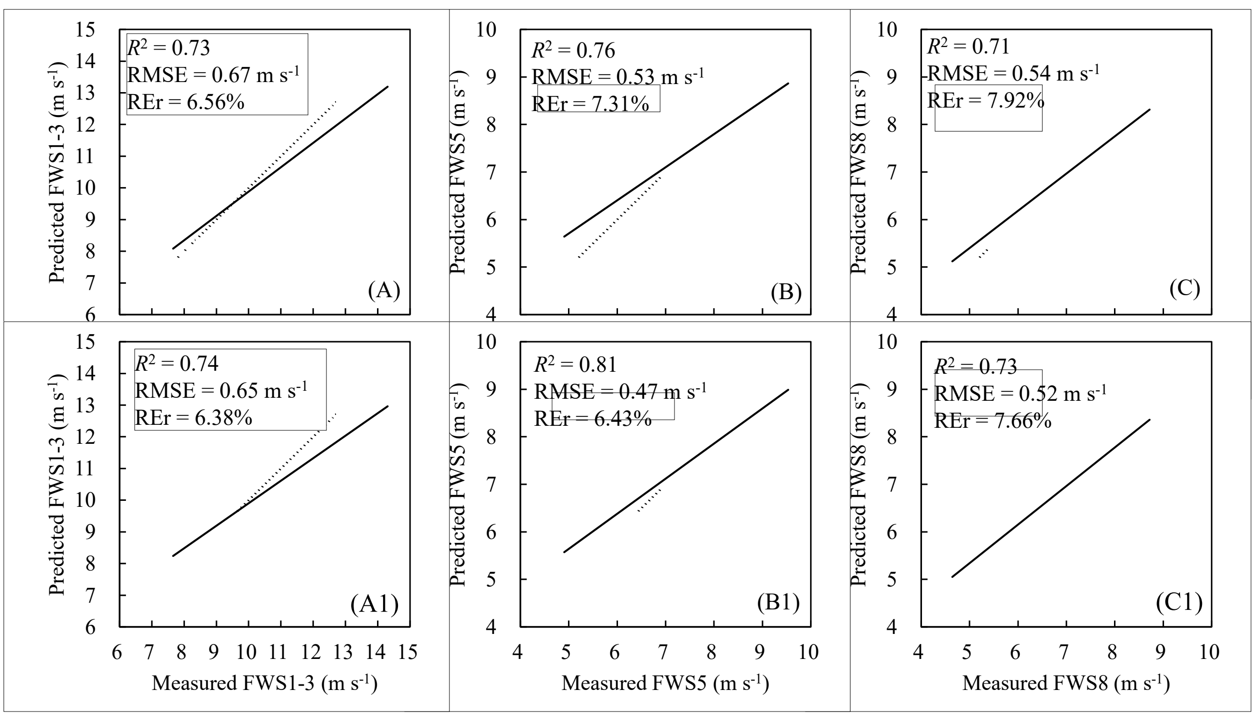


**Fig 5. Scatter plots of measured and predicted FWS based on the best performed models. Strategy 1: A~C; Strategy 2: A1~C1.**

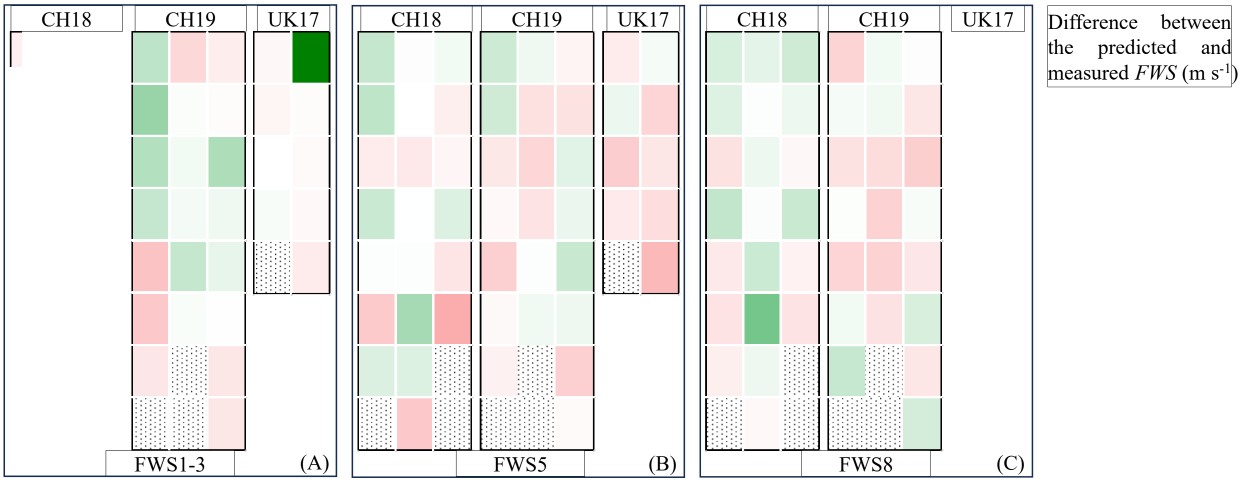


**Fig 6. Spatial distribution of the differences between the measured and predicted values of FWS for the validation plots in the experimental fields based on the precited results using the models. No maize plants were planted in the dotted plots.**

### Feature importance for lodging risk prediction

To further analyze the contribution of key features to lodging risk prediction, the feature importance of the top 20 input variables in RFR was shown in Fig 7. The top 20 features accounted for >75% of total predictive power. Among all the 20 feature variables, growth-related variables collectively contributed 45-60%, indicating crop growth status played a key role in lodging risk prediction, although the importance of each indicator was inconsistent in each RFR model. Management factors, particularly planting density and N application rate were also the major contributors to the RFR prediction, and the planting density had a high importance. Some of the environmental variables also emerged as important but less pronounced at predicting lodging risk, and they varied significantly for predicting the lodging risk at different internode positions..

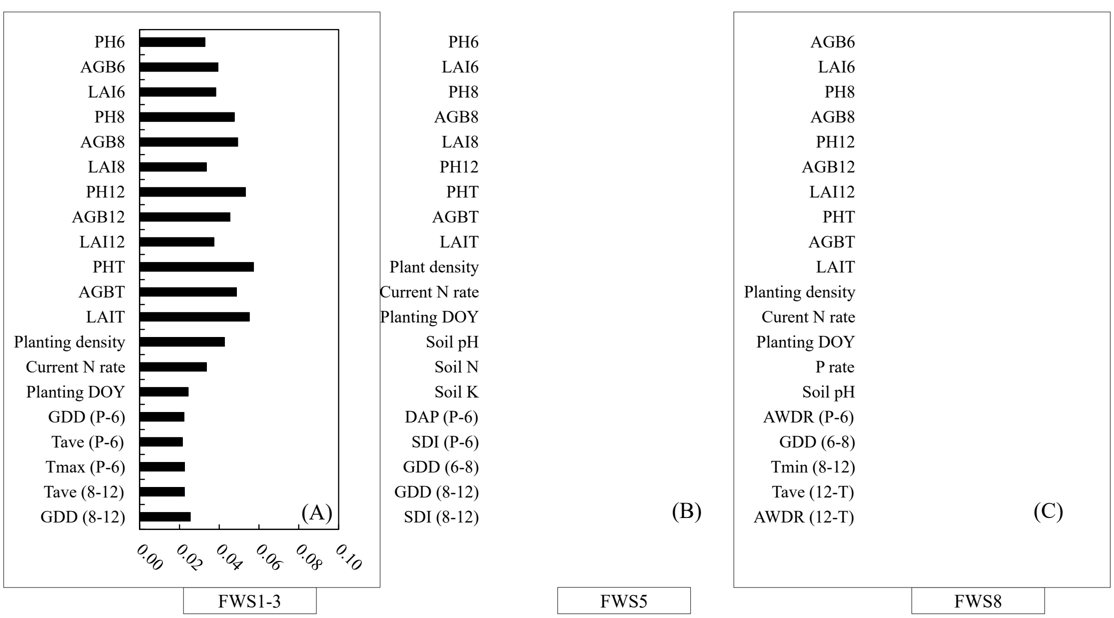


**Fig 7. The importance of the top 20 input variables in RFR for maize lodging risk prediction, at three internode positions: (A) 1st-3rd node, (B) 5th node and (C) 8th node.**

## Conclusions

This study developed a novel framework for early prediction of maize lodging risk using a data fusion strategy based on the integration of proximal sensing, crop growth modeling, and ML. The assimilation of proximal sensing data with the DSSAT-CERES-Maize model provided reliable consecutive maize growth status information for lodging risk prediction. The RFR model outperformed other tested algorithms (GBDT and SVR) and was effective in predicting lodging risk at early growth stages by aggregating multi-source and multi-stage features including crop growth, dates of key growth stages, management and environment factors. Predictions at the agronomically important V8 growth stage proved to be satisfactory for timely agronomic interventions to reduce lodging risk. The results identified the importance of crop growth status and management practices (mainly planting density and N rate) features, while the importance of the environmental features was inconclusive for lodging risk prediction due to the coarse spatial variation. The presented framework can enable precision lodging risk management by early identification of high lodging risk plots to enhance lodging resistance and minimize yield losses. To further improve the robustness and spatial adaptability of the framework, future research is needed to expand datasets across diverse regions and seasons select key features and simplify the lodging risk prediction model.

**Acknowledgments**

The research was financially supported by National Natural Science Foundation of China (Grant No. 32401698), International Foundation of Tobacco Research Institute of Chinese Academy of Agricultural Sciences (Grant No. IFT202504), and The Agricultural Science and Technology Innovation Program (ASTIP-TRIC-QH-1610232023026).

## References

Berry, P. M., Baker, C. J., Hatley, D., Dong, R., Wang, X., & Blackburn, G. A., et al. (2021). Development and application of a model for calculating the risk of stem and root lodging in maize. *Field Crops Research*, 262, 108037.

Chlingaryan, A., Sukkarieh, S., Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers & Electronics in Agriculture,* 151, 61-69.

Dong, R., Miao, Y., Berry, P., Wang, X., Yuan, F., Kusnierek, K., et al. (2023). In-season prediction of maize stem lodging risk using an active canopy sensor. *European Journal of Agronomy*, 151, 126956.

Dong, R., Miao, Y., Wang, X., Kusnierek, K. (2024). An active canopy sensor-based in-season nitrogen recommendation strategy for maize to balance grain yield and lodging risk. *European Journal of Agronomy,* 155, 127120.

Gonzalez, V. H., Tollenaar, M., Bowman, A., Good, B., Lee, E. A. (2018). Maize yield potential and density tolerance. Crop Science, 58, 472-485. Incremental learning for crop growth parameters estimation and nitrogen diagnosis from hyperspectral data. *Computers & Electronics in Agriculture,* 215, 108356.

Li, S., Ma, W., Peng, J., Chen, Z. (2015). Study on yield loss of summer maize due to lodging at the big flare stage and grain fllling stage. Scientia Agricultura Sinica, 19, 3952–3964. (Chinese text).

Ruan, G., Li, X., Yuan, F., Cammarano, D., Ata-UI-Karim, S.T., Liu, X., et al. (2022). Improving wheat yield prediction integrating proximal sensing and weather data with machine learning. *Computers & Electronics in Agriculture,* 195, 106852.

Shah, A. N., Tanveer, M., Rehman, A. U., Anjum, S. A., Iqbal, J., Ahmad, R. (2017). Lodging stress in cereal—effects and management: an overview. *Environmental Science & Pollution*, 24, 5222–5237.

Tremblay, N., Bouroubi, Y. M., B´elec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., et al.

(2012). Corn response to nitrogen is influenced by soil texture and weather. *Agronomy Journal*, 104 (6), 1658–1671.

Wang, X., Miao, Y., Dong, R., Zha, H., Xia, T., Chen, Z., et al. (2021). Machine learning-based in-season nitrogen status diagnosis and side-dress nitrogen recommendation for corn. *European Journal of Agronomy,* 123, 126193.

Xue, J., Xie, R., Zhang, W., Wang, K., Hou, P., Ming, B., et al. (2017). Research progress on reduced lodging of high-yield and -density maize. J*ournal of Integrative Agriculture*, 16 (12), 2717–2725.

Zhao, Y., Huang, Y., Li, S., Chu, X., Ye, Y. (2020). Improving the growth, lodging and yield of different density-resistance maize by optimising planting density and nitrogen fertilization. *Plant Soil and Environment*, 66(9), 453 – 460.