Predicting crop phenology: a simple logistic regression model approach.

Guillermo Leale[1], Bruno Cocitto[2], Ana Laura Cardoso[2], Pedro Lafluf[2], Ligia Tantucci[2], and Fernanda Mendez[1].

[1] Universidad Austral, Facultad de Ciencias Empresariales.
gleale@austral.edu.ar
[2] SIMA soft LLC.

Abstract. Crop yield prediction plays a central role in the agricultural planning and decision-making processes. In this paper, we analyze the phenology as a crucial aspect of this topic. We propose a simple model to predict phenology groups on maize and wheat crops at the field-level in Argentina. Our model uses logistic regression and includes photoperiod as an explanatory variable, which is very simple to calculate taking into account latitude and date as input. A large number of data records are used to obtain accurate results. Our model has been tested with over 77% accuracy for both crops. It was also benchmarked with Random Forest, which gives comparable results. However, our study shows that a very simple approach could be used with logistic regression, with very little loss of performance. Our model obtains phenology groups and also performs well with certain critical phenology stages for both crops. Our study aims to provide a simple and effective method for predicting phenology, which can be an aid to crop prediction and for farmers to make accurate decisions. Our work emphasizes the simplicity of the model, the use of a large number of data records, and the inclusion of the photoperiod as an input variable.

Keywords: Phenology prediction · Logistic Regression · Photoperiod.

1 Introduction

Crop yield prediction is a crucial aspect of agricultural planning and decision-making processes. One important factor in crop yield prediction is crop phenology, which refers to the timing of the different stages of crop growth. Crop phenology is influenced by various factors such as temperature, photoperiod, and weather variability [9]. In recent years, remote sensing data has been used to monitor crop phenology in near-real-time [7]. Several studies have proposed models to predict crop phenology using different approaches such as logistic regression [6], deep learning [11], and dynamic threshold methods [10]. These models have been tested with good accuracy and have the potential to aid in crop prediction and for farmers to make accurate decisions [6]. Additionally, high-resolution crop phenological datasets have been produced for different regions, such as China [12]. Understanding the spatiotemporal variability of crop phenology is important for monitoring and modeling land surface phenology dynamics and crop management and production [12].

In this context, the present paper proposes a simple logistic regression model approach to predict crop phenology groups on maize and wheat crops at the field-level in Argentina [6]. The model uses photoperiod as an explanatory variable, which is very simple to calculate taking into account latitude and date as input. It also uses a strategically defined zonification as a categorical variable. The model has been tested with very good accuracy and aims to provide a simple and effective method for predicting phenology, which can be an aid to crop prediction and for farmers to make accurate decisions [6].

Existing literature, such as the study by [2], demonstrates the potential for predicting wheat phenology using photoperiod and temperature as input variables. The study highlighted the significant influence of these factors on wheat development and applied logistic regression for prediction. However, this research focused exclusively on wheat, leaving a gap in predicting maize phenology using similar methods. Our study addresses this gap by proposing a model to predict phenology groups for maize and wheat crops in Argentina, using logistic regression with photoperiod and zone as input variables. In contrast to the model in [2], which employed growing degree days (GDD) and photoperiod for spring wheat phenology in Pakistan, our model emphasizes photoperiod and zone for predicting phenology groups in maize and wheat crops in Argentina. Additionally, our approach benefits from a large dataset, resulting in over 77% accuracy at the field level for both crops. Our results show that the proposed method outperforms Random Forest in maize. Although benchmarking with Random Forest yields better results in wheat, our study underscores the simplicity and effectiveness of logistic regression with minimal loss of accuracy or performance for this case. Overall, the proposed model provides a practical and efficient method for predicting maize and wheat crop phenology groups in Argentina, supporting farmers in their decision-making processes.

This work is organized as follows: In Section 2, we discuss the methods used for predicting phenology groups in maize and wheat crops, including the definition and importance of phenology groups, a primer to the logistic regression approach, and an explanation of the benchmark with the Random Forest model. Section 3 describes the experimental design, including the input dataset and the setup for the experiments conducted on maize and wheat crops. In Section 4, we present and discuss the results of our experiments, including the performance of the logistic regression model and the benchmarking with the Random Forest model. Finally, in Section 5, we draw conclusions about the effectiveness of our proposed logistic regression model for predicting phenology groups and its potential implications for agricultural decision-making. Additionally, we outline future work to be conducted on the model, expanding the scope of our research.

2 Methods

In this section, we describe the methods used for predicting phenology groups in maize and wheat crops in Argentina. We first introduce how phenology groups are defined and their importance in yield prediction. Then, we outline our logistic regression approach for predicting phenology, followed by the benchmark model using Random Forest, in order to show the simplicity and efficiency of our proposed method.

2.1 Phenology Groups

Phenology groups refer to the categorization of crop development stages based on observable growth characteristics. These groups play a crucial role in understanding crop yield prediction as they are strongly correlated with crop performance and productivity [8, 5]. In maize and wheat crops, phenology stages are defined by key events such as germination, vegetative growth, reproductive growth, and physiological maturity [13].

In maize, these stages range from Ve (Emergence) to R6 (Physiological Maturity). Ve - V3 denotes the Emergence to 3-Leaf Collar stage, during which the plant begins its life, developing its first three leaves. V4 - V6, the 4-Leaf to 6-Leaf Collar stage, witnesses the further vegetative development for the plant. The Blister (R2) to Dent (R5) stages are when the kernels mature, changing from blister-like to dented appearances due to starch accumulation. The 7-Leaf to 10-Leaf Collar stage (V7 - V10) marks continued vegetative growth and rapid stem elongation. The Tasseling to Silking stage (Vt - R1) signifies tassel emergence and pollination initiation. Finally, R6 indicates the Physiological Maturity stage, where maximum kernel dry weight is attained, and the plant is ready for harvest [14].

In wheat, phenology stages are classified differently. The Germination to Emergence stage is when the seedling emerges from the soil. Terminal spikelet is the stage where the final spikelet on the head (or spike) of the plant is formed. The First Node stage signifies the vegetative to reproductive transition as the stem grows vertically, and the first node becomes detectable. Subsequently, the Heading stage arrives, where the wheat head becomes visible as it emerges from the sheath, signifying the imminent flowering stage. The Anthesis stage follows, during which self-pollination occurs, setting the foundation for grain development. Then comes the Grain Filling stage, where the grains fill with starch and proteins, ultimately determining the final grain weight and impacting yield. Lastly, the Maturity stage signifies the point at which the grain achieves its maximum dry weight, indicating the plant is ready for harvest [1].

Accurate prediction of these stages can inform agricultural planning and decision-making processes, allowing farmers to optimize crop management practices and improve yield outcomes.

2.2 Logistic Regression Approach

The logistic regression model is a powerful statistical method for modeling the probability of the occurrence of an event by fitting data to a logistic function. It is particularly useful for predicting binary outcomes, making it suitable for predicting phenology group membership. In our study, we apply logistic regression to estimate the probability of a crop belonging to a specific phenology group based on photoperiod and zone as explanatory variables. The logistic function is represented as:

$$P(y=1) = \frac{1}{1+e^{-(b_0+b_1x_1+b_2x_2+...+b_nx_n)}}$$

where $P(y=1)$ is the probability of the event occurring (i.e., the crop belonging to a specific phenology group), b_0 is the intercept, $b_1, b_2, ..., b_n$ are the coefficients for the explanatory variables (photoperiod, zones), and $x_1, x_2, ..., x_n$ are the values of the explanatory variables.

The zone is a categorical variable representing a geographical division in Argentina. A binary transformation approach was used to include this variable in the logistic regression model.

Using logistic regression for predicting phenology, which incorporates photoperiod and zone as explanatory variables, we can provide a simple, yet effective, method for estimating the probability of crop development stage based on easily obtained input variables such as latitude, date, and geographical location.

2.3 Random Forest as a Benchmark Model

In this study, we also employ a more complex model, Random Forest [4], as a benchmark for comparison with our logistic regression model. Random Forest is an ensemble learning method that operates by constructing multiple decision trees during training and outputting the class that represents the mode of the classes for classification or the mean prediction for regression [15]. It offers robustness against overfitting and can handle large datasets with high dimensionality. However, Random Forest has a known higher degree of complexity and computational demand compared to logistic regression. By using it as a benchmark, we aim to demonstrate the simplicity and efficiency of the proposed logistic regression model. The primary goal of this comparison is to emphasize the performance of the logistic regression approach without a substantial loss in accuracy relative to the more complex benchmark model.

3 Experiments

In this section, we describe the experimental design used to test the performance of the logistic regression model in predicting phenology groups.

Table 1: Columns from the original wheat dataset, filled with dummy data

planting_date	**plot_lat**	**plot_lon**	**plot_zone**	**scouting_date**	**phenology_group**
2014-07-29	LAT1	LON1	VII	2014-10-12	First Node
2014-07-29	LAT2	LON2	VII	2014-10-22	Heading
2014-06-09	LAT3	LON3	VI	2014-08-15	Terminal Spikelet
2014-06-09	LAT4	LON4	VI	2014-09-03	First Node
2014-06-09	LAT5	LON5	VI	2014-09-17	First Node

3.1 Input Dataset

The dataset used in this study includes records of maize and wheat crops in Argentina, comprising a total of 39,035 and 77,131 observations, respectively. The source of the dataset is from the production database from an Argentine-based company. The raw data is private and is not publicly available due to confidentiality agreements signed by the company with both investors and customers. However, we can provide the structure of the original dataset. Table 1 shows the column names of the original dataset, filled with dummy data for wheat. The columns, in order, are the planting date, the latitude, the longitude, the zone (in Roman numbers, following the PAS nomenclature), the scouting date and the observed phenology group for the observation.

The dataset was subjected to several preprocessing steps, including cleaning, removal of missing values, and transformation of variables as needed. The preprocessing also involved filtering out records where the monitoring date was before the sowing date or more than 180 days after the sowing date and records where the phenology group did not correspond to the time elapsed between sowing and monitoring. The sowing date was obtained as additional information from another part of the database, and is not shown explicitly here. Additionally, records with different phenology groups for the same field on the same date were also removed. To demonstrate some of the preprocessing work, Figure 1 depicts the distribution of non-processed observations for maize phenology groups based on the number of days since planting. Several outliers and inconsistent data points were present in the phenology groups, and we removed them as part of the preprocessing steps.

One explanatory variable used in the final dataset is photoperiod, which was calculated from latitude and date. To obtain the photoperiod, we used the `daylength.py`[3] function in Python, which computes the length of the day given the day of the year and latitude. The final dataset includes 13,953 records for maize and 12,245 records for wheat. Table 2 shows dummy data for the resulting columns in wheat after the preprocessing steps, namely the plot zone, the calculated photoperiod length in hours, and the phenology group. Table 3 summarizes the observations per phenology group in the final dataset for maize and wheat. The data is not evenly distributed along all the phenology stages for both

[3] A Python function to compute the length of the day given day of the year and latitude. `https://gist.github.com/anttilipp/ed3ab35258c7636d87de6499475301ce`

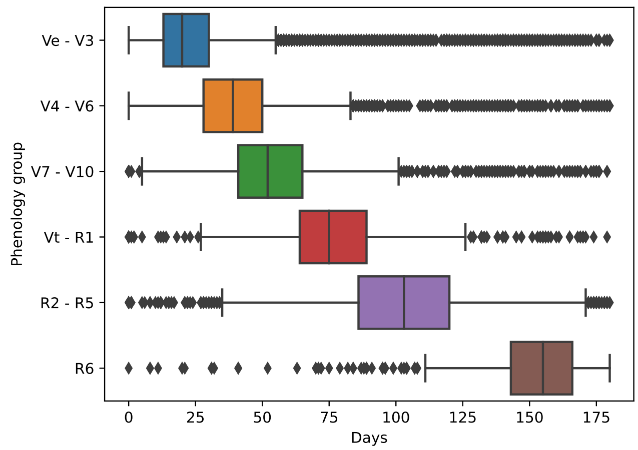


Fig. 1: Distribution of unfiltered observations for maize phenology groups by days since planting

Table 2: Columns from the processed wheat dataset, filled with dummy data

plot_zone	**cphotoperiod**	**phenology_group**
VII	866.752	First Node
VII	996.190	Heading
VI	1470.559	Grain Filling
VI	744.040	First Node
VI	1115.402	Anthesis

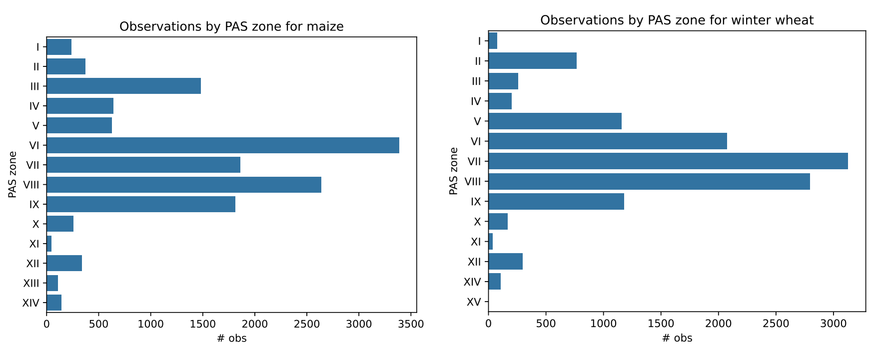
crops. This is due to the different perceived importance of each group. Farmers consider that some phenology groups are more relevant than others in order to track the growing status of each plant, and the consequence of these choices are clearly evidenced on the different number of observations per stage in the source dataset. It is worth noting that for both crops, all available observations in the database were used for the original dataset.

As stated above in Section 1, the other input variable is the zone. In particular, in this paper the zone refers to a zonification named *zona PAS* (PAS zone), created by the Buenos Aires Grain Exchange. The PAS zone comprises a geographic region of Argentina which is used to study the planted area and production of the main extensive crops. Argentina is divided into 15 zones, each with its own nomenclature, which are used to analyze crop variables. The guidelines for the PAS zone zonification are detailed in [3]. To include the zone in the logistic regression model, we used a one-hot encoding approach, which creates binary variables for each zone category. This allows the model to capture the influence of different geographical zones on phenology group prediction. Figure 2 shows the distributions of the observations in each PAS zone for maize (left) and wheat (right). It can be seen that the distribution is uneven for both crops,

Table 3: Observations per phenology group in the final dataset.

Crop	Phenology Group	Observations
Maize	Ve - V3	3,540
	V4 - V6	2,928
	R2 - R5	2,538
	V7 - V10	2,487
	Vt - R1	1,607
	R6	853
	Total	**13,953**
Wheat	Terminal Spikelet	3,671
	Germination to Emergence	2,917
	Grain Filling	2,224
	First node	1,643
	Anthesis	752
	Heading	551
	Maturity	487
	Total	**12,245**

this is due to the number of available observations in the dataset for each PAS zone.



(a) Maize (b) Wheat

Fig. 2: Distribution of observations by PAS zone on maize and wheat

This input dataset comprised a substantial number of observations for maize and wheat crops in Argentina, which were carefully preprocessed to ensure data quality. It is to be noted that for both crops, the same variables (columns) were used for the original and the processed final dataset, respectively.

3.2 Logistic Regression Experiment

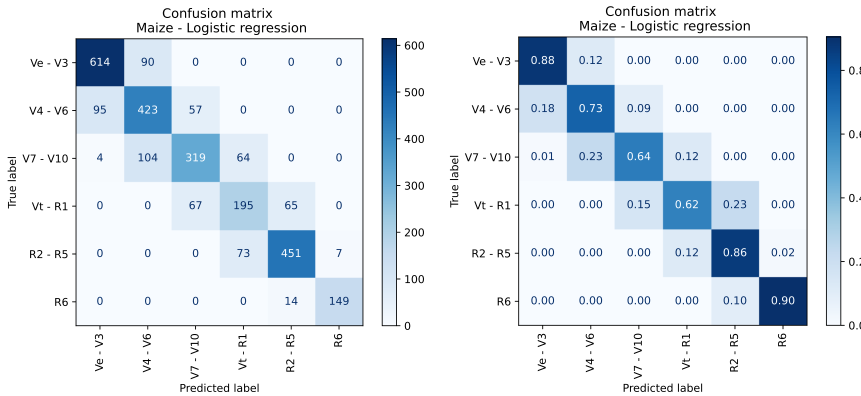
The logistic regression model was applied to the input dataset using photoperiod and zone as input variables and phenology group membership as the output variable. The experiment was designed to test the performance of the logistic regression model in predicting phenology groups. A training-test stratified split of 80%-20% was performed on the final dataset. A classical 10-fold cross-validation approach was applied for the training set. Both accuracy and F1-score were obtained as performance metrics for the model.

4 Results

In this section, we present and discuss the results of our experiments on the logistic regression model and the benchmark with the Random Forest model.

4.1 Results

For our experiments, all results were obtained with respect to test data, using the trained model. From the original dataset sizes of 13,953 observations for maize and 12,245 observations for wheat, the 20% test size was splitted using stratification with respect to the phenology group as the target class, obtaining test sets with sizes of 2791 observations for maize and 2449 observations for wheat. In all cases, the total number of observations in each phenology stage was computed. As well as this, the normalized number of observations in the range [0,1] for each stage was added, in order to have a fair comparison among all stages. The results from the logistic regression experiments indicate very good accuracy in predicting phenology groups for both maize and wheat crops. Regarding maize, the confusion matrix is depicted in Figure 3 with the total observations on the left, and the normalized observations on the right. The predicted labels are shown on the columns and the true labels on the rows. Each cell contains a numerical value which represents the number of observations for Figure 3a and the normalized number of observations for Figure 3b. The color intensity of the cell provides a visual aid for this value, with darker shades indicating higher values. For the following analysis, the normalized values corresponding to the actual values are written in parentheses next to the actual values. It can be seen that the proposed logistic regression model achieved reasonably good results. The model demonstrated high accuracy in the Ve - V3 and V4 - V6 stages, with 614 (0.88) and 423 (0.73) correct classifications, respectively, indicating its effectiveness in these stages. The model also performed relatively well in the V7 - V10 and Vt - R1 stages, with 319 (0.64) out of 491 and 195 (0.62) out of 327 observations correctly classified, respectively. However, the model performance was lower in the R2 - R5 and R6 stages, with only 451 (0.86) out of 531 and 149 (0.90) out of 163 observations correctly classified, respectively, suggesting that there might be room for improvement in these stages. The accuracy of the maize classification model is approximately 73.1% and the F1 score is approximately



(a) Maize (b) Maize, normalized

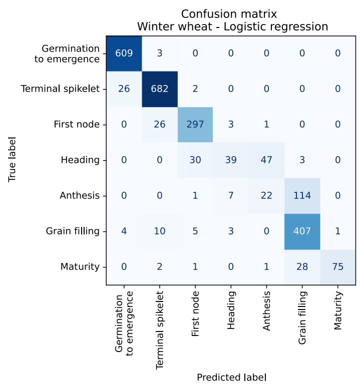
Fig. 3: Confusion matrix for Logistic Regression in maize

81.8%. Therefore, the proposed model can be considered useful for predicting the phenological stages of maize crops effectively.

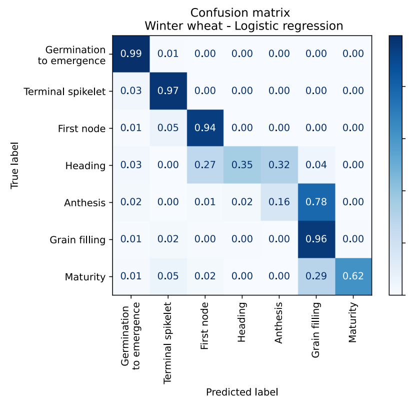
Regarding wheat, the logistic regression model achieved overall good results, as demonstrated by the confusion matrix in Figure 4. The highest number of correct classifications were in the Germination to Emergence and Terminal Spikelet stages, with 609 (0.99) and 682 (0.97) correct classifications, respectively, indicating that the model is effective in these stages. The model also performed well in the First Node stage, with 297 (0.94) out of 327 observations correctly classified. Additionally, the model demonstrated acceptable performance in the Heading and Anthesis stages, with 39 (0.35) out of 119 observations correctly classified in the Heading stage and 114 (0.78) out of 144 observations correctly classified in the Floración stage. The model demonstrated relatively high performance in the Grain Filling and Maturity stages, with 407 (0.96) out of 430 and 75 (0.62) out of 107 observations correctly classified, respectively. The results indicate that the model has satisfactory overall performance and is appropriate for predicting phenological stages of wheat crops with a high degree of accuracy. Specifically, the wheat logistic regression model achieves an accuracy rate of approximately 94.2% and an F1 score of approximately 93.4%. The effectiveness of the model is supported by the confusion matrix, demonstrating that it is a valuable tool for predicting phenological stages in wheat crops.

4.2 Benchmarking

To benchmark the logistic regression model, we applied the Random Forest model to the same input dataset using the same experimental setup. The performance metrics of the Random Forest model were also calculated for each experiment.



(a) Wheat



(b) Wheat, normalized

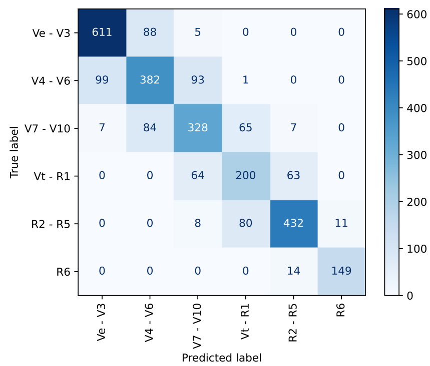
Fig. 4: Confusion matrix for Logistic Regression in wheat

The confusion matrices shown in Figure 5 provide insight into the accuracy of Random Forest in predicting the phenological stages of maize and wheat crops. For maize (Figure 5b), the highest number of correctly predicted observations is in the Ve-V3 stage with 611 (0.87), while the lowest number of correctly predicted observations is in the R6 stage with only 149 (0.92). In stage Vt - R1 there is a lower performance with 127 misclassified observations out of 327 total observations for the group. Overall, the accuracy rate of the model for maize is approximately 72.7%.

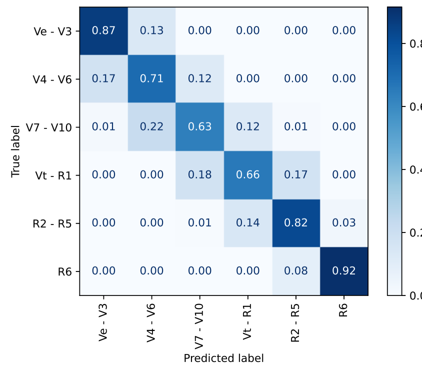
For wheat (Figure 5d), the highest number of correctly predicted observations is in the Terminal Spikelet stage with 680 (0.96), while the lowest number of correctly predicted observations is in the Maturity stage with only 94 (0.85). The model yields 45 misclassified observations in the Grain Filling stage across all the other phenological groups, out of the 430 total observations for the group. Overall, the accuracy rate of the model for wheat is approximately 96.9%.

Comparing the results, the Random Forest model performs better in predicting the phenological stages of wheat crops than maize crops. This is evident in the higher accuracy rate for wheat compared to maize. The model for wheat was able to correctly predict more observations across all stages than the model for maize. This behavior is consistent with that of the Logistic Regression Approach.

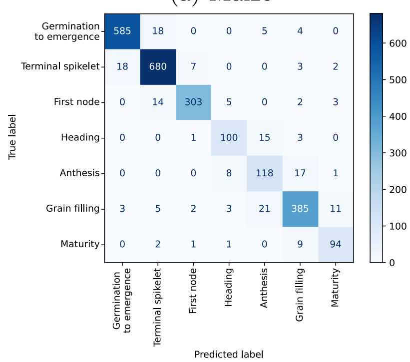
To compare the Logistic Regression approach against the Random Forest approach, an ANOVA test was performed. To ensure statistical significance, the experiment was repeated 100 times using different random splits of the dataset into training and testing subsets, and applying each split equally to both models. Results from the repeated experiments with different dataset splits were summarized using performance metrics such as accuracy and F1-score, taking into account the variability in the performance metrics across the repeated experi-



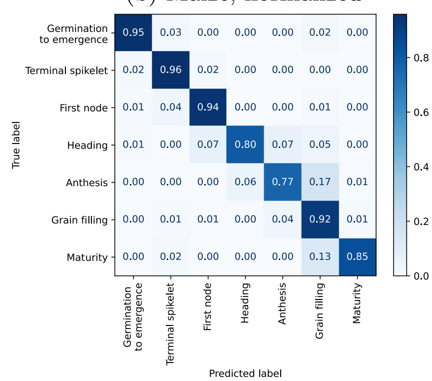
(a) Maize



(b) Maize, normalized



(c) Wheat



(d) Wheat, normalized

Fig. 5: Results for Random Forest on maize and wheat

ments. These metrics provide a comprehensive understanding of the performance in predicting the phenology groups for both models.

The results of the ANOVA test for maize indicate that there is a highly significant difference between the mean values of the two groups, with p-values of 2.87×10^{-69} for accuracy and 5.54×10^{-62} for the F1 score. This suggests that our Logistic Regression model outperforms Random Forest in terms of predicting the phenological stages of maize crops. Specifically, our model yields both a higher mean accuracy and F1 score and therefore is comparatively better than Random Forest. It is worth noting that both approaches achieved relatively high levels of accuracy, with mean values of 0.775 and 0.764 for the Logistic Regression accuracy and the Random Forest accuracy, respectively. Regarding wheat, from the ANOVA test we obtained very small p-values for both accuracy (5.65×10^{-188}) and F1 score (2.33×10^{-183}), indicating a significant difference in mean values between the Random Forest and logistic regression models for predicting phenological stages in wheat crops. While Random Forest outperformed logistic regression, with mean accuracy values of 0.92 and 0.85, respectively, as well as mean F1 scores of 0.92 and 0.83, respectively, the difference in accuracy and F1 score is not substantial. The results show that while the Random Forest model exhibits comparable performance to the logistic regression approach, the latter achieves comparable accuracy with a significantly simpler approach and less computational demand. This demonstrates the effectiveness of our proposed logistic regression model for predicting phenology groups in maize and wheat crops.

As well as the above, a computational demand assessment was performed for each method. A 11th Gen Intel® Core™ i5-1135G7 at 2.42 GHz with 12GB RAM with Windows® 10 Pro was used for all the experiments. The average time for each training run ranged between 14 and 16 seconds, with no noticeable difference between both methods.

5 Conclusions

In this paper, we presented the results of our experiments on predicting the phenological stages of maize and wheat crops using a logistic regression model and benchmarked it against the Random Forest model. Our logistic regression model achieved very good accuracy rates for both maize and wheat crops, demonstrating high effectiveness in several stages. Although the Random Forest model yielded slightly better results for wheat, the difference in accuracy and F1 score was not substantial. We performed an ANOVA test to compare the performance of both models, which showed that our Logistic Regression model outperformed the Random Forest model in predicting the phenological stages of maize crops. The Logistic Regression approach achieved comparable accuracy rates with a significantly simpler and less computationally demanding approach. Therefore, our proposed Logistic Regression model is an effective and cost-efficient tool for predicting the phenological stages of maize and wheat crops.

By using a large dataset and including photoperiod and zone as explanatory variables, our model demonstrates good accuracy in predicting critical phenology stages for both crops. Furthermore, the simplicity of our model, when benchmarked against the more complex Random Forest model, highlights the potential for this approach to be a valuable tool for farmers and agricultural decision-makers. Future research could explore additional input variables, other main crops such as soybean, and model refinements to further improve the accuracy and generalizability of the logistic regression model in predicting crop phenology.

6 Acknowledgements

As a double-blind peer review process is required, we are unable to acknowledge any specific individuals or organizations in this section. We would like to express our gratitude to all those who have contributed to this research project, including colleagues who have provided valuable feedback and support throughout the

study. We look forward to the completion of the peer review process. Should this work be selected for the conference, we will update this section to acknowledge those who have contributed to this work.

References

1. Acevedo, E., Silva, P., Silva, H.: Wheat growth and physiology. Bread wheat, improvement and production **30**, 39–70 (2002)
2. Aslam, M., Aslam, M.A., Ahmed, M., Stöckle, C.O., Higgins, S.S., Hassan, F.u., Hayat, R.: Can growing degree days and photoperiod predict spring wheat phenology? Frontiers in Environmental Science (2017). https://doi.org/10.3389/fenvs.2017.00057
3. Bolsa de Cereales de Buenos Aires: Guía metodológica de estimaciones agrícolas. introducción, metodología y manual de procedimientos para el relevamiento y análisis de datos (junio 2020), `https://www.bolsadecereales.com/imagenes/GuiaPASyECC.pdf`
4. Breiman, L.: Consistency for a simple model of random forests. University of California at Berkeley. Technical Report **670** (2004)
5. Foerster, S., Kaden, K., Foerster, M., Itzerott, S.: Crop type mapping using spectral–temporal profiles and phenological information. Computers and Electronics in Agriculture **89**, 30–40 (2012)
6. Foerster, S., Kaden, K., Foerster, M., Itzerott, S.: Crop type mapping using spectral–temporal profiles and phenological information. Computers and Electronics in Agriculture (2012). https://doi.org/10.1016/j.compag.2012.07.015
7. Gao, F., Zhang, X.: Mapping crop phenology in near real-time using satellite remote sensing: Challenges and opportunities. Journal of Remote Sensing (2021). https://doi.org/10.34133/2021/8379391
8. Gao, F., Zhang, X.: Mapping crop phenology in near real-time using satellite remote sensing: Challenges and opportunities. Journal of Remote Sensing **2021** (2021)
9. Hill, C.B., Li, C.: Genetic architecture of flowering phenology in cereals and opportunities for crop improvement. Frontiers in Plant Science (2016). https://doi.org/10.3389/fpls.2016.01906
10. Huang, X., Liu, J., Zhu, W., Atzberger, C., Liu, Q.: The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method. Remote Sensing (2019). https://doi.org/10.3390/rs11232725
11. Lin, T., Hu, H., Zhong, R., Xu, J., Xu, J., Huang, J., Wang, S., Ying, Y.: A deep learning approach to conflating heterogeneous geospatial data for corn yield estimation: A case study of the us corn belt at the county level. Global Change Biology (2019). https://doi.org/10.1111/gcb.14885
12. Luo, Y., Zhang, Z., Chen, Y., Li, Z., Tao, F.: Chinacropphen1km: A high-resolution crop phenological dataset for three staple crops in china during 2000–2015 based on leaf area index (lai) products. Earth System Science Data (2020). https://doi.org/10.5194/essd-12-197-2020
13. McIntoch, R.: Wheat: Prospects for global improvement. Dordrecht, Boston, London **71**, 86 (1997)
14. Nleya, T., Chungu, C., Kleinjan, J.: Corn growth and development. Grow Corn Best Manag. Pract (2016)

15. Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P.: Random forest: a classification and regression tool for compound classification and qsar modeling. Journal of chemical information and computer sciences **43**(6), 1947–1958 (2003)