Effects of wind speed and wind direction on crop yield forecasting using dynamic time warping and an ensembled learning model

Bright Bediako-Kyeremeh Corresp., Equal first author, [1], **Ting Huai Ma** Equal first author, [2], **Huan Rong** [3], **Benjamin Kwapong Osibo** [2], **Lord Amoah** [2], **Isaac Kofi Nti** [4], **Lorenzo Mamelona Ambinintsoa** [1]

[1] School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

[2] School of Computer Science, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

[3] School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China, Nanjing, Jiangsu, China

[4] Department of Information Technology, University of Cincinnati, Cincinnati, Ohio, USA

Corresponding Author: Bright Bediako-Kyeremeh
Email address: brytkyeremeh@gmail.com

The cultivation of cashew crops carries numerous economic advantages, and countries worldwide that produce this crop face a high demand. The effects of wind speed and wind direction on crop yield prediction using proficient deep learning algorithms are less emphasized or researched. We suggest employing a combination of advanced deep learning techniques, specifically focusing on Long Short-Term Memory (LSTM) and random forest models. We intend to enhance this ensemble model using Dynamic Time Warping (DTW) to assess the spatiotemporal data similarities within Jaman North, Jaman South, and Wenchi. In the Bono region of Ghana, these three areas are crucial for cashew production. The model achieved an R^2 score of 0.847. Among the three municipalities, Jaman South had the highest evaluation scores for the model, with an RMSE of 0.883, an R^2 of 0.835, and an MBE of 0.212 when comparing actual and predicted values for Wenchi. In terms of the annual average wind direction, Jaman North recorded (270.5 SW°), Jaman South recorded (274.8 SW°), and Wenchi recorded (272.6 SW°). The DTW similarity distance for the annual average wind speed across these regions fell within specific ranges: Jaman North (±25.72), Jaman South (±25.89), and Wenchi (±26.04). Following the DTW similarity evaluation, Jaman North demonstrated superior performance in wind speed, while Wenchi excelled in wind direction. This underscores the potential efficiency of DTW when incorporated into the analysis of environmental factors affecting crop yields, given its invariant nature. The results obtained can guide further exploration of DTW variations in combination with other machine learning models to predict higher cashew yields. Additionally, these findings emphasize the significance of wind speed and direction in vertical farming, contributing to informed decisions for sustainable agricultural growth and

development.

Effects of wind speed and wind direction on crop yield forecasting using dynamic time warping and an ensembled learning model

Bright Bediako-Kyeremeh[1*], Ting Huai Ma[2*], Huan Rong[3], Benjamin Kwapong Osibo[2], Lord Amoah[2], Isaac Kofi Nti[4], Lorenzo Mamelona Ambinintsoa[1]

[1] School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China
[2] School of Computer Science, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China
[3]School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China
[4] Department of Information Technology, University of Cincinnati, Ohio, USA

Corresponding Author:
Bright Bediako-Kyeremeh[1]
219 Ning Liu Road, Nanjing, Jiangsu, 210044, China
Email address: brytkyeremeh@gmail.com.com

Abstract

The cultivation of cashew crops carries numerous economic advantages, and countries worldwide that produce this crop face a high demand. The effects of wind speed and wind direction on crop yield prediction using proficient deep learning algorithms are less emphasized or researched. We suggest employing a combination of advanced deep learning techniques, specifically focusing on Long Short-Term Memory (LSTM) and random forest models. We intend to enhance this ensemble model using Dynamic Time Warping (DTW) to assess the spatiotemporal data similarities within Jaman North, Jaman South, and Wenchi. In the Bono region of Ghana, these three areas are crucial for cashew production. The model achieved an R^2 score of 0.847. Among the three municipalities, Jaman South had the highest evaluation scores for the model, with an RMSE of 0.883, an R^2 of 0.835, and an MBE of 0.212 when comparing actual and predicted values for Wenchi. In terms of the annual average wind direction, Jaman North recorded (270.5 SW°), Jaman South recorded (274.8 SW°), and Wenchi recorded (272.6 SW°). The DTW similarity distance for the annual average wind speed across these regions fell within specific ranges: Jaman North (±25.72), Jaman South (±25.89), and Wenchi (±26.04). Following the DTW similarity evaluation, Jaman North demonstrated superior performance in

wind speed, while Wenchi excelled in wind direction. This underscores the potential efficiency of DTW when incorporated into the analysis of environmental factors affecting crop yields, given its invariant nature.
The results obtained can guide further exploration of DTW variations in combination with other machine learning models to predict higher cashew yields. Additionally, these findings emphasize the significance of wind speed and direction in vertical farming, contributing to informed decisions for sustainable agricultural growth and development.

Introduction

Recently, there has been a significant focus on crop yield, which is influenced by various factors such as crop genotype, environment, and management practices (Khaki et al., 2020). Machine learning and deep learning models have been used in different forms to predict crop yield, providing valuable insights throughout the supply chain from pre-production to post-production. In the global economy, one crucial objective of accurately predicting crop yield is to ensure an adequate food supply for nations, including livestock feed and energy resources. This necessitates the development of a crop prediction model that can deliver high-precision results to facilitate effective decision-making. Can it be demonstrated that Dynamic Time Warping (DTW) can be utilized to assess the similarity of targeted features, such as wind speed and wind direction, in a spatial dataset of cashew crops and produce better predictions than what has been claimed in the literature about Time-Weighted Dynamic Time Warping (TWDTW) being superior to DTW? Can DTW be integrated into a learning mechanism to achieve a more accurate model? To address these questions, we have adopted a learning mechanism framework that combines Long Short-Term Memory (LSTM), Dynamic Time Warping (DTW), and Random Forest Regressor (RF). DTW will enhance spatial analysis within the framework by incorporating specific environmental features obtained from the Predictable of Worldwide Energy Resources, enhanced Data Access Viewer (POWER | DAVe, n.d.) to predict yield in a selected cashew-growing geographical area. Machine learning is a set of statistical methods designed to solve specific tasks such as classification or regression by automatically detecting patterns and anomalies in data and making decisions or acquiring skills similar to humans, improving their learning independently over time (Nti et al., 2022; Sagan et al., 2021). Deep learning models, including CNN and LSTM, have been employed by various authors (Cao et al., 2020; Srivastava et al., 2022; X. Wang et al., 2020) to predict crop yield in wheat and other crops. Hybrid deep-learning models have also been studied (Khaki & Wang., 2019) to predict crop yield based on environmental and genotype features. Additionally, machine learning models have been utilized by (Kumar et al., 2015; Ganapathi et al., 2020; Kalimuthu et al., 2020) and to predict crop yield. Over the years, deep learning techniques have been extensively applied to predict crop yield with high accuracy in various crops by authors such as (Khaki & Wang., 2019; Khaki et al., 2020; X. Wang et al., 2020; Sagan et al., 2021; Tian et al., 2021) investigated crop yield enhancement in winter wheat using LSTM and remote sensing data. LSTM, a special type of Recurrent Neural Network (RNN), is capable of capturing long-term dependencies (Bhimavarapu et al., 2023). It can bridge long time intervals between inputs and analyze temporal patterns at different frequencies, which is advantageous for analyzing crop-growing cycles of varying lengths (Omdena, 2022). Bhimavarapu et al. (2023) also highlighted that LSTM considers historical values, adjusts itself based on complete patterns, and makes future forecasts. Furthermore, machine learning regression models have proven to be effective for crop yield prediction, as demonstrated by authors such as (L. Wang et al., 2016; Rale et al., 2019; Keerthana et al., 2021; Panigrahi et al., 2023). In fact, (L. Wang et al., 2016) showed that the random forest model produced more accurate estimates in their research. DTW, as defined by (Xiao et al., 2023), is an effective method for limited-samples-based crop classification that compares the similarity between two time-series curves, exhibiting reduced sensitivity to training samples. We aim to integrate DTW into an ensemble of LSTM and RF models to achieve a higher accuracy model. Peng et al. (2023) investigated and demonstrated that the revised TWDTW effectively utilizes crop phenological information and improves the accuracy of extracting summer crop planting areas on a large scale. This indicates that the initial TWDTW model did not achieve the desired accuracy, necessitating its reinforcement in their research. Therefore, we believe that adopting a learning mechanism integrated with DTW to assess similarities and differences in instrumental environmental features is essential for predicting crop yield.

Materials & Methods

Data

We utilized a dataset compiled from various sources by the Ghana Meteorological Agency (GMet, 2021). This dataset included environmental variables like solar radiation, relative humidity, and rainfall, collected throughout the entire year from 1999 to 2018, encompassing a span of 20 years. This dataset specifically covered the three municipalities where cashew is grown.
We sought cashew yield production data from the Ministry of Food and Agriculture (MoFA, 2021) for the municipalities being investigated. The data encompassed the study period of 1999-2018 and were focused on cashew-growing regions, namely Jaman North, Jaman South, and Wenchi.
We acquired remote sensing information for the three designated study regions from (POWER | DAVe, 2023). The provided weather parameters included soil moisture, wind speed at 2m, and wind direction at 10m, spanning the study period of 1999-2018. These supplementary parameters are vital for ensuring sustainable crop yields, particularly in practices like vertical farming (van Delden et al., 2021). Figure 1 illustrates the geographical positions of our study areas.

Model framework

LSTM has shown promising results in crop cultivation, as demonstrated by the study conducted by (J. Wang et al., 2022). Additionally, similar positive outcomes have been observed with machine learning regressors, as highlighted by (Rale et al., 2019). A prospective avenue involves integrating mapping techniques to enhance crop prediction by considering land use and land cover, a concept explored in the works of (Chaves et al., 2021; Feng et al., 2021).

Our proposed model is a fusion of LSTM and RF regression, incorporating Dynamic Time Warping (DTW). This combined approach is well-suited for time series regression tasks. DTW serves as a valuable tool to measure the similarity between two-time series sequences, particularly when the parameters may have varying lengths or exhibit evidence of time-based warps. The key purpose of integrating DTW is to align the sequences through warping and temporal stretching, thus identifying the optimal alignment that minimizes discrepancies among corresponding nodes. In Figure (2), we illustrate the framework of DTW integrated into the ensemble deep learning model.

Model Construction

The cashew production dataset was loaded into pandas and ensure data quality by removing missing values using the dropna function. Outliers were filtered using the interquartile percentage technique. To scale the features, we applied the MinMax scaler. For the target variable transformation scaling, we used a one-hot-encoding technique on the categorical feature and production. To analyze the temporal dependencies in the time series data and identify essential patterns, we employed an LSTM model. The data was split into training (70%) and testing (30%) sets. The LSTM model had 50 neurons in the LSTM units and 1 dense unit. The activation function used was sigmoid. We compiled the model using an Adam optimizer with a mean square error loss. The LSTM model was trained for 20 epochs with a batch size of 32. Next, we applied the DTW algorithm to compare the predicted sequence with the ground truth sequence. This allowed us to measure the similarities of environmental features through feature engineering. We calculated the DTW distances between each time series in the training dataset and the test dataset. We then printed the DTW distance and the optimal alignment path, which indicates the indices of the points in time_series1 (wind speed) and time_series2 (wind direction). The matrix representation of Dynamic Time Warping (DTW) can be computed using the dot product (DP) between matrices.

$$DTW_q\left(x, y^1\right) = \min_{\pi \in A(x,x^1)} \langle A_\pi, D_q\left(x, x^1\right)\rangle^{\frac{1}{q}}$$

Where $D_q\left(x, y^1\right)$ stores distance $d\left(x_i, x_j^1\right)$ at the power q

Algorithm for DTW

Input: X(t), $0 \leq t \leq nT + L$ *is the historical wind speed/direction time series*

T: represents the length of a complete cashew seasonal period

N: represents the number of seasonal periods

L: represents the length of the time series of the last incomplete season

Output: X_{nT+L+1}, X_{nT+L+2}, … $X_{nT+L+\text{predicted length}}$

for i = 0: predict length -1 do

$A = \{X_{nT} + {}_i, X_{nT} + 1 + {}_i, X_{nT} + 2 + {}_i, \ldots X_{nT+L+i}\}$

For j = 0:nT – L do

$B_j = \{X_j, X_{j+1}, X_{j+2}, \ldots X_{j+L}\}$
$C_j = D_{DTW}(A, B_j)$ //{The methods for calculating DTW distance}
End for
K = minindex(C)
$X_{nT+L+i+1} = X_{k+L+i+1}$
End for
Return $X_{nT+L+1}, X_{nT+L+2}, \ldots X_{nT+L+\text{predict length}}$

The framework incorporated a random forest regression model to enhance the prediction of the target variable. This was achieved by utilizing DTW similarity scores for wind direction and wind speed as additional features, along with other relevant input features, to improve the accuracy of yield prediction. The random forest model was constructed using a library like sci-kit-learn. The input features were a combination of LSTM and DTW, while the target variable was the output. The model was configured with 100 estimators and a random state of 42. To assess the performance of the random forest regression model, metrics such as Mean Bias Error (MBE), Root Mean Square Error (RMSE), and coefficient of determination (R2) were used to evaluate the goodness of fit. The expressions for these regression metrics are as follows:

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)$$

Where O_i is the observation value and P_i is the predicted value

$$RMSE = \sqrt{\sum_{i-1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}}$$

Where $\hat{y}_i$ the predicted value y_i is the observed value, n number of a given dataset.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}$$

Where $\hat{y}_i$ is the predicted value and $\overline{y}_i$ is the mean value

Results

LSTM is highly effective in handling time series data, thanks to its ability to control data flow through its gates. DTW is used to assess similarities in additional features, while the random forest regressor helps evaluate model errors and biases. The accuracy of the model was measured using an R^2 score of (0.847). The various metrics for three municipalities were Jaman North (MBE= 0.231, RMSE= 0.802, R^2= 0.742), Jaman South (MBE= 0.22, RMSE= 0.883, R^2= 0.835), and Wenchi (MBE= 0.212, RMSE= 0.746, R^2= 0.702). Table 1 displays the metric scores of the overall model's performance and individual municipal performance. Figure 3 displays a line chart comparing the model's performance with and without the inclusion of the target variables, namely wind speed and wind direction. This provides a complete overview of the model's performance outcomes with and without these variables.

These scores indicate the exceptional performance of our model, which aligns with the findings of (J. Wang et al., 2022). The dynamic time-warping technique allowed us to evaluate similarities in spatiotemporal data such as wind speed and wind direction. Unlike (Chaves et al., 2021), who only considered the harvest period, we analyzed the entire year, including the flowering and fruit development stages of the cashew crop, as well as the impact of wind speed and direction during the harvest period on yield. Figure 4 (a), (b), and (c) showcase a line regression chart for RMSE, R^2, and MBE. These charts utilize evaluation metrics to visually illustrate the correlation between the model (wind speed and wind direction) and observed (production) variables in the dataset. This is crucial as production is primarily influenced by environmental factors, particularly wind speed and wind direction.

The chart, which has a slope trend line of 1:1, visually demonstrates the performance of the three municipalities based on the metric index. The blue "dot" symbols represent a variable (wind speed, wind direction, and production).

The annual average wind direction for Jaman North was (270.5 SW), Jaman South was (274.8 SW), and Wenchi was (272.6 SW). The DTW similarity distance for the annual average wind speed ranged from (±25.72) for Jaman North, (±25.89) for Jaman South, and (±26.04) for Wenchi. Graphical representations in Figure 5 (a) and (b) showcase the performance of the DTW model in relation to wind speed (measured in km/h) and wind direction (measured in x^o) across the three municipalities and the twenty-year duration of the study.

Discussion

We evaluated our model using metrics such as MBE, RMSE, and R^2 for the three (3) cashew crop-growing areas over the study period of 1999 to 2018. To compare the similarity or calculate the distance between two arrays or time series with different lengths, we used Dynamic Time Warping (DTW) since it is invariant to time shifts between series. Our research aimed to calculate the average wind speed in km/h and wind direction in the North, South, East, and West on the cardinal directional compass for the three study areas of the cashew crop and assess their effects on crop yield.

Previous research, such as that by (X. Wang et al., 2020), focused on using the LSTM model with Modis LAI products and the time-weighted dynamic time warping (TWDTW) variant of DTW to predict the yield of winter wheat in Henan Province, China, but did not consider the effects of wind speed and wind direction on yield. Similarly, (Chaves et al., 2021; J. Wang et al., 2022) discussed the role of TWDTW in determining the area of the crop but overlooked the impact of wind speed and wind direction. Our results align with their findings, showing high yields in Jaman South, Jaman North, and Wenchi, and highlighting the influence of wind speed and direction in the under-studied municipalities.

The matrix table indicates that Jaman South performed the best among the cashew-growing areas, excelling in wind speed determination, while Wenchi ranked second and performed better in wind direction. The geographical location of Wenchi on the western side of Jaman North and South may explain its higher wind direction. Jaman South had a high wind speed, which contributed to its high-yield production. The DTW model effectively analyzed the similarity of spatial data, demonstrating its effectiveness for spatiotemporal analysis.

While previous studies have used TWDTW as a standalone model for classification, our research suggests that using DTW with the right ensemble deep learning models can yield better results when identifying key environmental parameters for crop yield. The DTW model accurately captures the appropriate wind speed for Wenchi, aligning with its production levels, while Jaman South experiences a notable influx of air in its direction. The significance of wind speed and wind direction in cultivating crops, especially cashew crops, supports the implementation of vertical farming for sustainable agriculture, as emphasized by (Kalantari et al., 2018; Beacham et al., 2019; van Delden et al., 2021).

Conclusions

The necessity of increasing crop yield to meet growing supply demands has been emphasized, and cashews are no exception due to their significant benefits. This highlights the need for further research to enhance cashew yield using DTW and its variants. Our study aimed to explore how DTW could be utilized to identify similarities in targeted features (wind speed and wind direction) within the cashew spatiotemporal dataset, surpassing previous claims that TWDTW is superior to DTW in terms of prediction accuracy. Additionally, we aimed to determine the most effective integration of DTW into a learning mechanism to achieve a higher accuracy model. Our proposed model yielded successful results, exhibiting fewer errors with regression metrics such as MBE, RMSE, and R2.

Our proposed method leverages LSTM for sequence modeling and DTW for similarity assessment. Additionally, the method incorporates random forest regression for ensemble-based predictions.

References

Beacham, A. M., Vickers, L. H., & Monaghan, J. M. (2019). Vertical farming: a summary of approaches to growing skywards. *J. Hortic. Sci. Biotechnol.*, *94*(3), 277–283. https://doi.org/10.1080/14620316.2019.1574214

Bhimavarapu, U., Battineni, G., & Chintalapudi, N. (2023b). Improved Optimization Algorithm in LSTM to Predict Crop Yield. *Computers*, *12*(1), 10. https://doi.org/10.3390/computers12010010

Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Han, J., & Li, Z. (2020). Identifying the Contributions of Multi-Source Data for Winter Wheat Yield Prediction in China. *Remote Sensing*, *12*(5), 750. https://doi.org/10.3390/rs12050750

Ganapathi, N. P., Sudarshan, K., & Bhatta, A. (2020). Agriculture Crop Prediction Using Machine Learning Algorithms. *International Journal of Research in Engineering, Science and Management*.

Kalantari, F., Tahir, O. M., Joni, R. A., & Fatemi, E. (2018). Opportunities and challenges in sustainability of vertical farming: a review. *J. Landsc. Ecol.*, *11*(1), 35–60. https://doi.org/10.1515/jlecol-2017-0016

Kalimuthu, M., Vaishnavi, P., & Kishore, M. (2020). Crop prediction using machine learning. *In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 926–932.

Keerthana, M., Meghana, K. J. M., Pravallika, S., & Kavitha, M. (2021). An Ensemble Algorithm for Crop Yield Prediction. *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 963–970. https://doi.org/10.1109/ICICV50876.2021.9388479

Khaki, S., & Wang, L. (2019). Crop Yield Prediction Using Deep Neural Networks. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00621

Khaki, S., Wang, L., & Archontoulis, S. V. (n.d.). A CNN-RNN Framework for Crop Yield Prediction. *Frontiers in Plant Science*.

Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN Framework for Crop Yield Prediction. *Frontiers in Plant Science*.

Kumar, R., Singh, M. P., Kumar, P., & Singh, J. P. (2015). Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique. *International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)*.

Nti, I. K., Akyeramfo-Sam, S., Bediako-Kyeremeh, B., & Agyemang, S. (2022). Prediction of social media effects on students' academic performance using Machine Learning Algorithms (MLAs). *Journal of Computers in Education*, *9*(2), 195–223. https://doi.org/10.1007/s40692-021-00201-z

Omdena. (2022, September 30). *Crop Yield Prediction Using Deep Neural Networks*. Omdena.

Panigrahi, B., Kathala, K. C. R., & Sujatha, M. (2023). A Machine Learning-Based Comparative Approach to Predict the Crop Yield Using Supervised Learning with Regression Models. *Procedia Computer Science*, *218*, 2684–2693. https://doi.org/10.1016/j.procs.2023.01.241

Peng, Q., Shen, R., Dong, J., Han, W., Huang, J., Ye, T., Zhao, W., & Yuan, W. (2023). A new method for classifying maize by combining the phenological information of multiple satellite-based spectral bands. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.1089007

POWER | DAVe. (n.d.). *Prediction of Worldwide Energy Resource (POWER) | Data Access Viewer*. NASA.

Rale, N., Solanki, R., Bein, D., Andro-Vasko, J., & Bein, W. (2019). Prediction of Crop Cultivation. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 0227–0232. https://doi.org/10.1109/CCWC.2019.8666445

Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D. R., Sidike, P., & Fritschi, F. B. (2021). Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *174*, 265–281. https://doi.org/10.1016/j.isprsjprs.2021.02.008
Srivastava, A. K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F., Gaiser, T., & Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. *Scientific Reports*.
Tian, H., Wang, P., Tansey, K., Zhang, J., Zhang, S., & Li, H. (2021). An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China. *Agricultural and Forest Meteorology*, *310*, 108629. https://doi.org/10.1016/j.agrformet.2021.108629
van Delden, S. H., SharathKumar, M., Butturini, M., Graamans, L. J. A., Heuvelink, E., Kacira, M., Kaiser, E., Klamer, R. S., Klerkx, L., Kootstra, G., Loeber, A., Schouten, R. E., Stanghellini, C., van Ieperen, W., Verdonk, J. C., Vialet-Chabrand, S., Woltering, E. J., van de Zedde, R., Zhang, Y., & Marcelis, L. F. M. (2021). Current status and future challenges in implementing and upscaling vertical farming systems. *Nature Food 2021 2:12*, *2*(12), 944–956. https://doi.org/10.1038/s43016-021-00402-w
Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal*, *4*(3), 212–219. https://doi.org/10.1016/j.cj.2016.01.008
Wang, X., Huang, J., Feng, Q., & Yin, D. (2020). Winter Wheat Yield Prediction at County Level and Uncertainty Analysis in Main Wheat-Producing Regions of China with Deep Learning Approaches. *Remote Sensing*, *12*(11), 1744. https://doi.org/10.3390/rs12111744
Xiao, X., Jiang, L., Liu, Y., & Ren, G. (2023). Limited-Samples-Based Crop Classification Using a Time-Weighted Dynamic Time Warping Method, Sentinel-1 Imagery, and Google Earth Engine. *Remote Sensing*, *15*(4), 1112. https://doi.org/10.3390/rs15041112

Figure 1

Study Area of Three Municipalities (Jaman North, Jaman South, and Wenchi). Insert map of West Africa and Ghana. Source: ArcGIS

Geographical map showing the location of the three municipalities known for large cashew production in Ghana

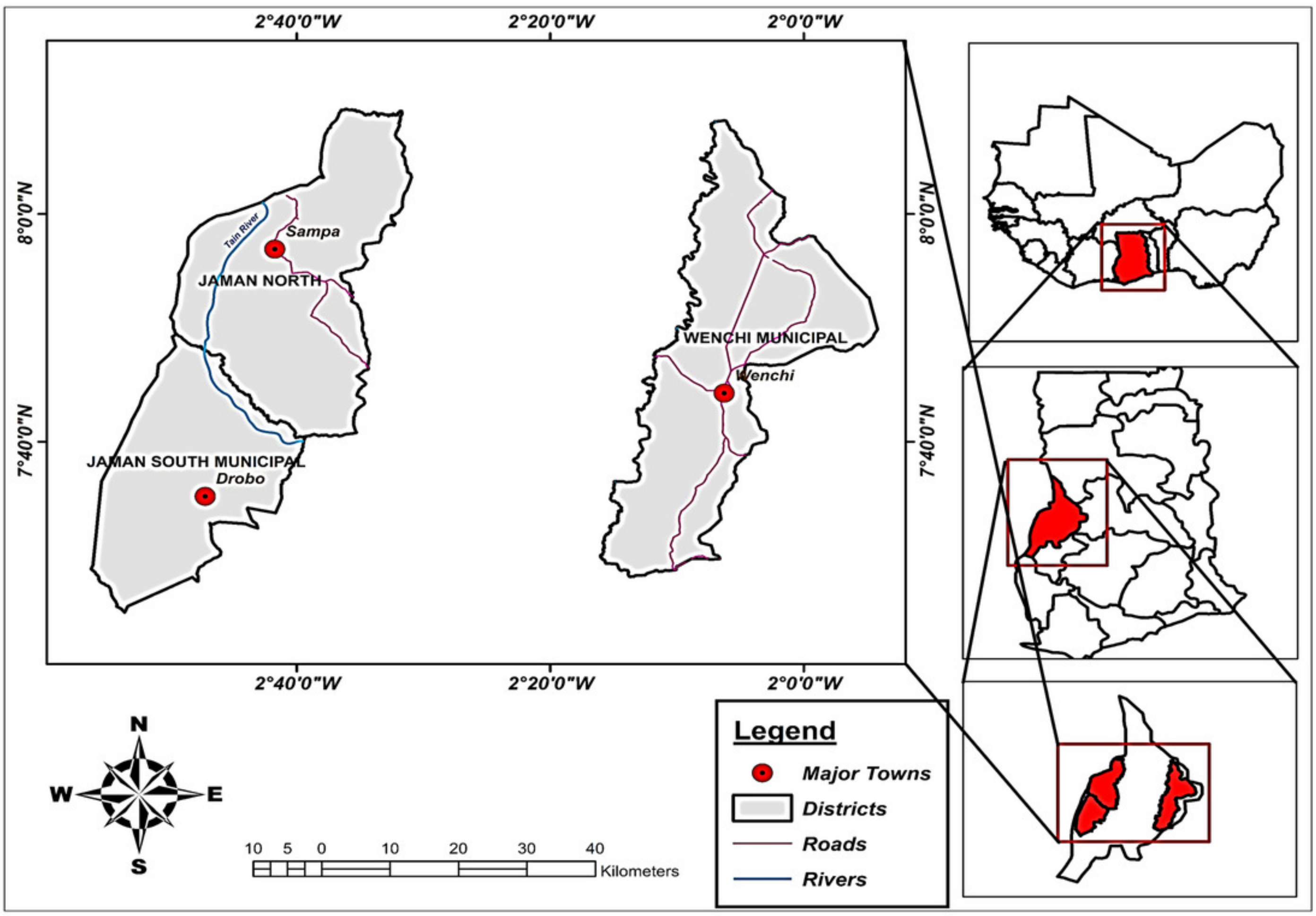


Figure 2

The architecture of LSTM/DTW/RF

The model architecture informs readers how the inception point of the model where data is accepted through the LSTM model, then the DTW model where the targeted variable/parameter similarity evaluation is computed then thr last layer where random forest predict result through best decision

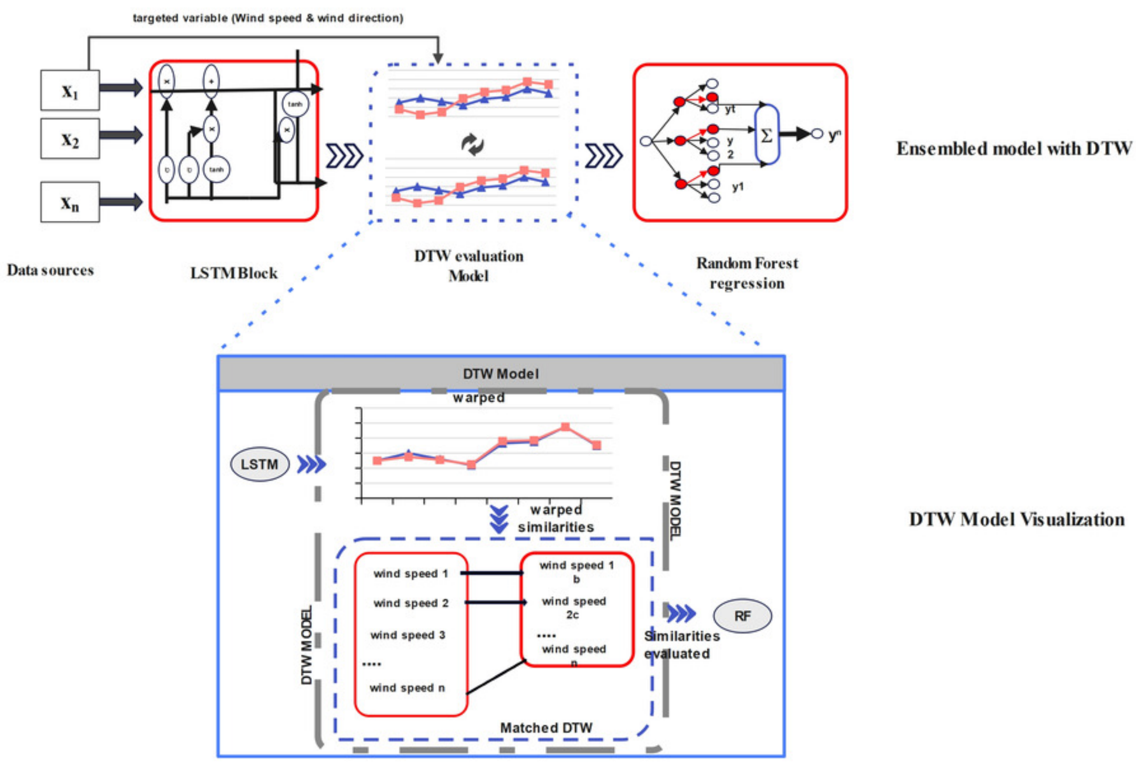


Figure 3

Overall model performance visualization with an overlay bar-with-scatter Plot

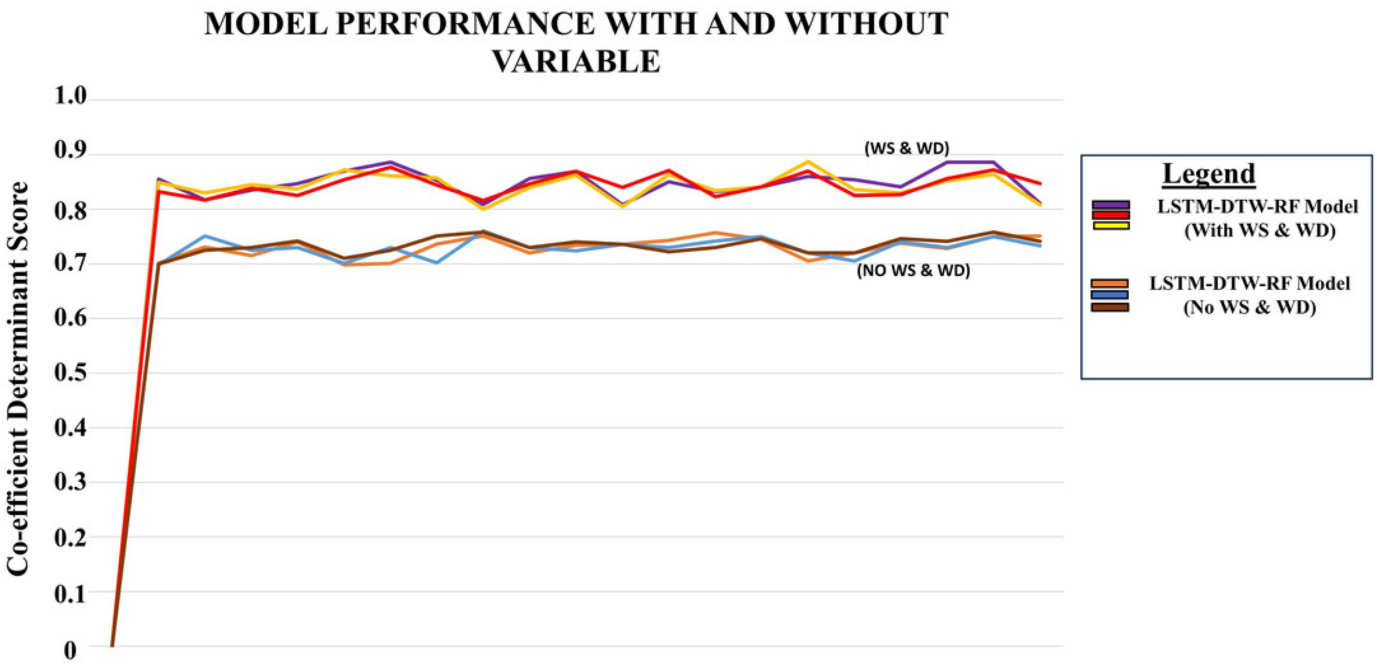


Figure 4

Regression Charts for the study area performance evaluation metric (RMSE, MBE and R^2)

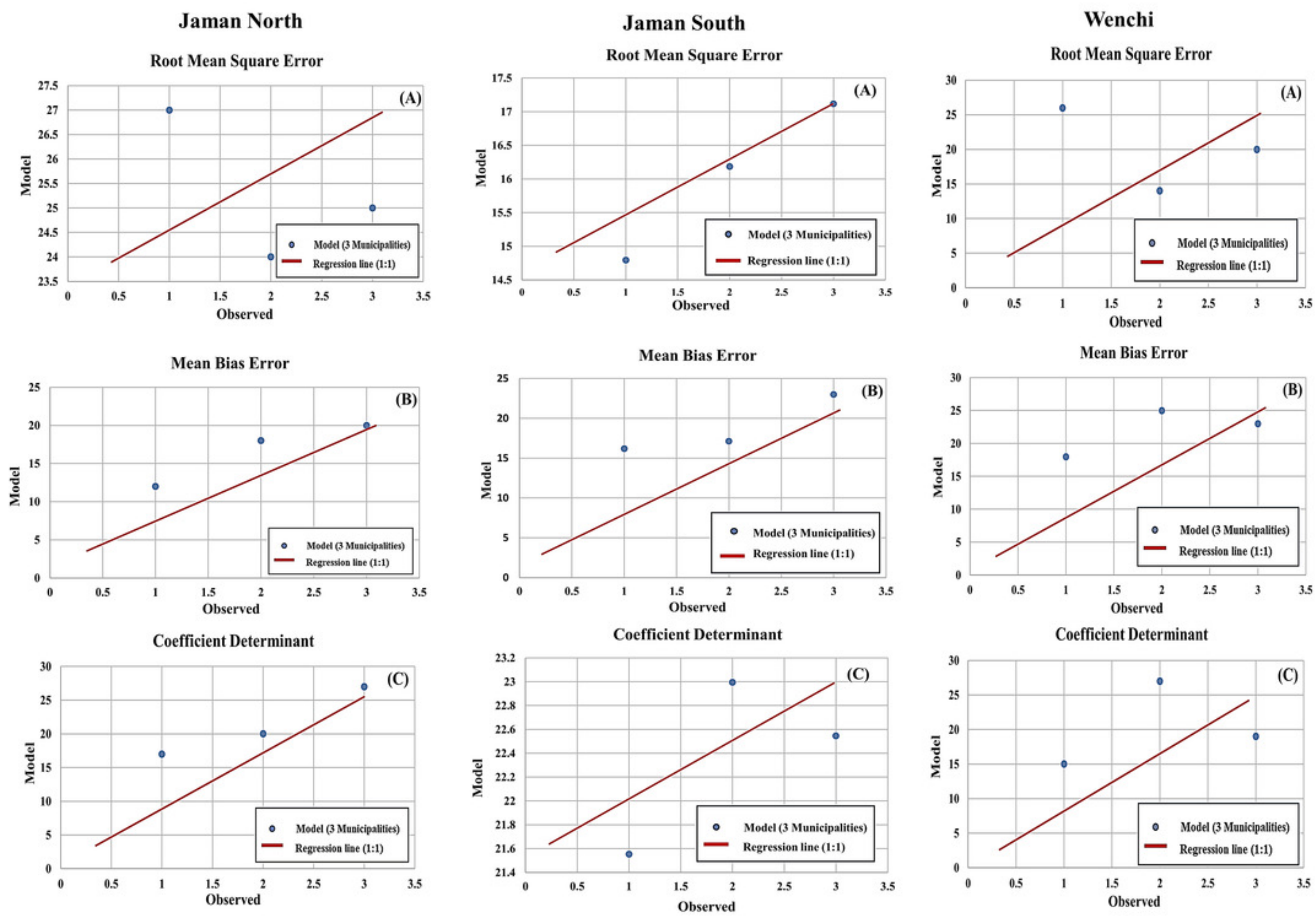
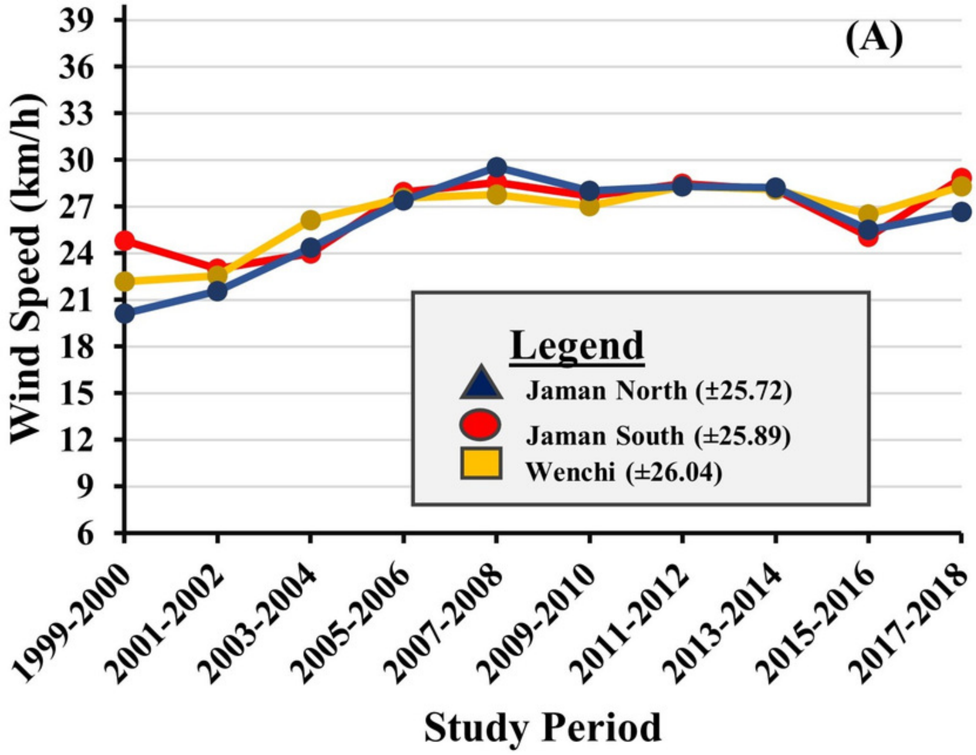


Figure 5

Study Area Performance on Wind Speed and Wind Direction Results Line Chart


(A)
Wind Speed (km/h)
39
36
33
30
27
24
21
18
15
12
9
6
Legend
Jaman North (±25.72)
Jaman South (±25.89)
Wenchi (±26.04)
1999-2000
2001-2002
2003-2004
2005-2006
2007-2008
2009-2010
2011-2012
2013-2014
2015-2016
2017-2018
Study Period

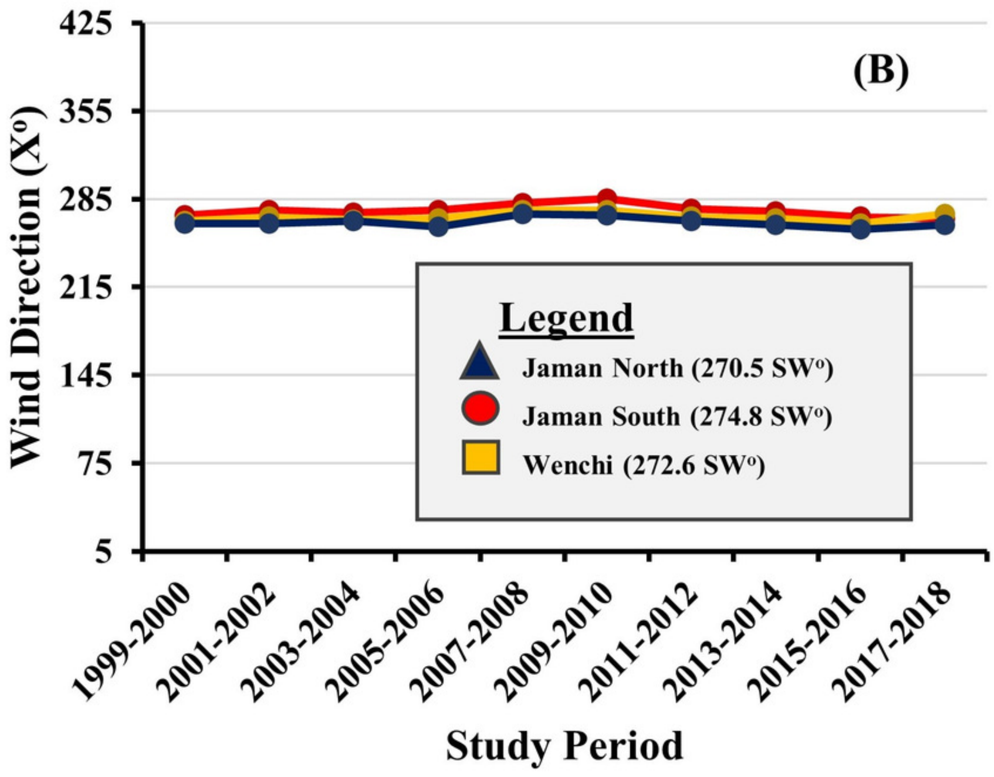

(B)
Wind Direction (X°)
425
355
285
215
145
75
5
Legend
Jaman North (270.5 SW°)
Jaman South (274.8 SW°)
Wenchi (272.6 SW°)
1999-2000
2001-2002
2003-2004
2005-2006
2007-2008
2009-2010
2011-2012
2013-2014
2015-2016
2017-2018
Study Period

Table 1 (on next page)

Regressor Matrix Score for the model ablation of the cashew crop-growing municipalities

MBE/RMSE can range from *0* to (*n*), where the closer the score is to 0 the better performing the model is. ***R2*** 0.75 - 1 a substantial amount of variance simplified

Table 1. Evaluation Metrics Score for the overall model performance and of the cashew crop-growing municipalities

Municipalities	MBE	RMSE	R^2
Jaman North	0.231	0.802	0.742
Jaman South	0.22	**0.883**	**0.835**
Wenchi	**0.212**	0.746	0.702
Overall Model			**0.847**