Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 26 (3) : 377-379 (September - 2024)
https://doi.org/10.54386/jam.v26i3.2445
https://journal.agrimetassociation.org/index.php/jam

Short Communication

Comparative analysis of wheat yield prediction through artificial intelligence, simulation modelling and statistical analysis in central Punjab

K. K GILL[1], KAVITA BHATT[1], AKANSHA[1], BALJEET KAUR[2]* and S. S. SANDHU[1]

[1]*Department of Climate Change and Agricultural Meteorology, Punjab Agricultural University, Ludhiana, Punjab, India*
[2]*Krishi Vigyan Kendra, Nurmahal, Jalandhar, Punjab, India*
**Corresponding author email: bchahal57@gmail.com*

Recognizing the paramount importance of wheat in agricultural landscapes, crop yield forecasting assumes a pivotal role in various aspects of farming planning and management. Beyond its implications for domestic food supply, accurate yield forecasts contribute to informed decision-making in international food trade, facilitate ecosystem sustainability efforts, and play a crucial role in the broader context of agricultural practices (Pandey and Sinha, 2006). Integration of artificial intelligence including ARIMA, LASSO, ANN, ELNET, MLP, ELM (Gomez *et al*., 2021) and simulation modelling like CERES-wheat model (Ritchie *et al.,* 1988), reveals a high level of accuracy in predicting wheat yield, providing valuable insights into the relationships between crop yield and weather parameters. These simulation models enhance predictive precision, facilitating informed decision-making and sustainable agricultural practices. Researchers have enhanced the precision of predicting wheat yield in Ludhiana, by developing different models (Singh *et al.,* 2021).

The daily weather data of different parameters (maximum temperature, minimum temperature, morning relative humidity, evening relative humidity, rainfall and bright sunshine hour) for the years 2001 to 2022 were collected from the Agrometeorological Observatory of Department of Climate Change & Agricultural Meteorology, Punjab Agricultural University, Ludhiana. Correspondingly, wheat yield data for Ludhiana district was collected from statistical abstracts of Punjab for the years 2001-2022. The weekly weather variables were generated using daily data by averaging the daily maximum temperature (Tmax) and daily minimum temperature (Tmin), averaging the morning relative humidity (RH1) and evening relative humidity (RH2), and summing up the rainfall (RF). These weekly aggregated variables were utilized for subsequent analysis and for generating the weather indices following the method described by Ghosh *et al.,* (2014).

The stepwise multiple linear regression (SMLR), artificial neural network (ANN), elastic net (ELNET), multiple layer perceptron (MLP), extreme learning machine (ELM) were used in the study as described by Krithikha and Velammal (2022) and Ajith *et al.,* (2023). The already calibrated CERES-wheat model (Gill *et al.,* 2018) was also used to predict the wheat yield. Out of total dataset of 22 years (2001-2022), the training dataset (2001 to 2017) was used for calibration of all the listed techniques and testing dataset (2018-2022) was used for validation.

The performance of different yield forecasting models was tested by computing coefficient of determination (R^2), root mean square error (RMSE), normalised root mean square error (nRMSE), mean bias error (MAE) and Pearson correlation coefficient (r) by using the below-mentioned formulae:

$$R^2 = \frac{1}{n}\left\{\frac{\sum_{i=1}^{n}(Mi - \bar{M})(Oi - \bar{O})}{\sigma_M \sigma_O}\right\}$$

$$MBE = \frac{1}{n}\left[\sum_{i=1}^{n}(Si - Oi)\right]$$

$$RMSE = \sqrt{\left(\sum_{i=1}^{n}(Si - Oi)2\right)/n}$$

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \times \frac{100}{\bar{O}}$$

Article info - DOI: *https://doi.org/10.54386/jam.v26i3.2445*
Received: 15 November 2023; Accepted: 22 May 2024 ; Published online : 01 September 2024

Table 1: Statistical evaluation of various modelling techniques for different parameters during calibration period

Statistical parameters	Models					
	SMLR	ANN	ELNET	MLP	ELM	DSSAT
R^2	0.79	0.98	0.99	0.68	0.90	0.56
RMSE (kgha^{-1})	242.0	68.02	60.19	344.2	301.2	346.1
NRMSE (%)	5.1	1.3	1.2	6.8	5.9	10.2
MBE (kgha^{-1})	206.0	49.5	18.2	-305.1	-252.6	-223.0

Evaluation of different models

The statistical evaluation of different models carried out using 17 years dataset are presented in Table 1. The stepwise multiple linear regression (SMLR) technique yielded a coefficient of determination (R^2) of 0.79, suggesting that approximately 79 % of the variability in the observed yield can be explained by the model (Table 1). The root mean square error (RMSE) was recorded at 242.0 kg ha^{-1} and the normalized RMSE (NRMSE) was 5.1%, indicating a relatively low percentage of error relative to the range of observed yield values. Furthermore, the Mean Bias Error (MBE) was determined to be 206.0 kg ha^{-1}, representing the average difference between the predicted and observed yields. A positive MBE suggests a tendency for the model to slightly overestimate yields on average. This study corroborates the findings of Kumar *et al.,* (2019), emphasizing the significance of SMLR in providing valuable insights for accurate crop yield forecasting.

ANN and ELNET models demonstrate exceptional predictive capabilities, as evidenced by the perfect R^2 value of 0.98 and 0.99, respectively, indicating a precise fit to the observed yield data (Table 1). The root mean square error (RMSE) was remarkably low for both models, with ANN at 68.02 kg ha^{-1} and ELNET at 60.19 kg ha^{-1}. The normalized RMSE (NRMSE) percentages were also low, highlighting the accuracy of these models (1.3% for ANN and 1.2% for ELNET), respectively. Moreover, the mean bias error (MBE) values for ANN (49.5 kg ha^{-1}) and ELNET (18.2 kg ha^{-1}) demonstrate minimal average differences between predicted and observed yields. In contrast, the MLP and DSSAT models exhibited a low R^2 value (0.68 and 0.56, respectively), suggesting poor predictive performance or potential model over fitting. The MBE value of -305.1 kg ha^{-1} for MLP and -223.0 kg ha^{-1} for DSSAT indicated a slight underestimation bias. The ELM model, with an R^2 value of 0.90, presented a moderate fit to the data. While the RMSE (301.2 kg ha^{-1}) and NRMSE (5.9%) values were relatively low, the MBE value of -252.6 kg ha^{-1} indicated a substantial underestimation bias.

Comparison of different technique

Based on the per cent deviation of error values determined during calibration and validation of the model, which ranged from 3.78 to 16.70 being highest for DSSAT and lowest for ELNET model (Table 2). Based on the percent deviation values and ranking of different approaches the ELNET approach emerged as the most effective method for predicting wheat yield in Ludhiana. Pearson correlation coefficient analysis revealed that, ANN, ELNET, SLMR, MLP and ELM were all statistically significant at the 1% level, with correlation coefficients ranging from 0.70 to 0.86. Conversely, the DSSAT model exhibited a correlation coefficient of 0.22, which was significant at the 5 % level of significance. The accuracy of wheat yield prediction, based on different approaches ranked in the following order: ELNET>ANN>ELM>SLMR>MLP>DSSAT (Table 2).

Table 2: Performance of different models during validation period

Model	Per cent deviation	Correlation coefficient
ANN	7.26	0.81*
ELNET	3.78	0.86*
MLP	13.58	0.70*
ELM	7.67	0.75*
SMLR	10.70	0.72*
DSSAT	16.70	0.22**

*Significant at p<0.01, **significant at p<0.50

Thus, based on the statistical evaluation of different statistical, machine learning and simulation approaches during training and testing period it can be concluded that the ELNET & ANN approaches emerge as the most effective for wheat yield prediction in Ludhiana, underscoring the significance of selecting appropriate modelling techniques for accurate and reliable agricultural forecasts. This research expands the scope of agricultural forecasting by demonstrating the potential of machine learning methods to significantly enhance predictive accuracy in crop yield estimation. Simulation models could not perform better than any of the statistical model tested in the study.

ACKNOWLEDGMENT

Authors thank the Department of Climate Change and Agricultural Meteorology, Punjab Agricultural University, Ludhiana and India Meteorological Department for providing meteorological weather data.

Source of funding: No source of funding

Conflict of interest statement: The author(s) declare(s) that there is no conflict of interest.

Data availability statement: Can't be shared.

Author's contribution: **K. K. Gill**: Methodology, Conceptualization of research; **K. Bhatt**: Paper writing & analysis; **Akansha**: Analysis of data; **B. Kaur**: Data tabulation; **S. S. Sandhu:** Editing

REFERENCES

Ajith, S., Manoj, K. D., Deb, S.G., Pradip, B., Subhendu, B., Shyamal, K. and Ragini, H.R. (2023). Comparative evaluation of penalized regression models with multiple linear regression for predicting rapeseed-mustard yield:

Weather-indices based approach. *J. Agrometeorol., 25*(3): 432–439. https://doi.org/10.54386/jam.v25i3.2185

Gill, K. K., Sandhu, S. S., Divya and Mishra, S. K. (2018). Pre-harvest wheat yield prediction using CERES-wheat model for Ludhiana district, Punjab, India. *J. Agrometeorol.,* 20(4): 319-21. https://doi.org/10.54386/jam.v20i4.574

Gomez, D., Salvador, P., Sanz, J. and Casanova, J. L. (2021). Modelling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico. *Agric. For. Meteorol., 300*, 1-8.

Ghosh, K., R. Balasubramanian, S. Bandopadhyay, N. Chattopadhyay, K.K. Singh and L. S. Rathore. (2014). Development of crop yield forecast models under FASAL- a case study of kharif rice in West Bengal. *J. Agrometeorol.,* 16(1): 1–8. https://doi.org/10.54386/jam.v16i1.1479

Krithikha, S. S. and Velammal B. (2022). A comprehensive approach on predicting the crop yield using hybrid machine learning algorithms. *J. Agrometeorol.,* 24(2):179-185. https://doi.org/10.54386/jam.v24i2.1561

Kumar, S., Attri, S. D. and Singh, K. K. (2019). Comparison of LASSO and stepwise regression technique for wheat yield prediction. *J. Agrometeorol.,* 21*(*2): 188-92. https://doi.org/10.54386/jam.v21i2.231

Pandey, S. N. and Sinha, B. K. (2006). *Plant physiology* (4th ed.). Vikas Publishing House Pvt. Ltd, New Delhi (pp. 485-487).

Ritchie, J. T., Godwin, D. C. and Singh, U. (1988). Soil and weather inputs for the IBSNAT crop models. In *Proceedings of the IBSNAT symposium: Decision support system for Agrotechnology Transfer Part I* (pp. 31-45). University of Hawaii, Hoolulu, Hawaii, USA.

Singh, M. C., Pal, V., Singh, S. P. and Satpute S. (2021). Wheat yield prediction in relation to climatic parameters using statistical model for Ludhiana district of central Punjab. *J. Agrometeorol.,* 23(1):122-26. https://doi.org/10.54386/jam.v23i1.97