Agriculture Association of Textile Chemical and Critical Reviews Journal (2025) 532-536

AATCC Review

04 May 2025: Received
22 June 2025: Revised
30 June 2025: Accepted
01 August 2025: Available Online

https://aatcc.peerjournals.net/

Original Research Article

Open Access

Forecasting of rice yield of India through non linear growth models

Mahesh Kumar*[1], Aarti Kumari[2], Thakur Bajrang Kumar Singh[1] and Ajeet Kumar[3]

[1]*Department of Statistics and Computer Applications, College of Basic Sciences and Humanities (CBS&H), Dr Rajendra Prasad Central Agricultural University, Samastipur, India*
[2]*Department of Plant Pathology and Nematology, Post Graduate College of Agriculture, Dr Rajendra Prasad Central Agricultural University, Samastipur, India*
[3]*Department of Microbiology, College of Basic Sciences and Humanities (CBS&H), Dr Rajendra Prasad Central Agricultural University, Samastipur, India*

ABSTRACT

*This research paper entitled **"Forecasting of rice yield of India through non linear growth models"** is based on secondary data. Data was collected for the years 1963 to 2021 from the official sites of indiastat. The data from 1963 to 2016 were analyzed through R- Software and five years of data 2017 and 2021 were kept for model validation of yield forecasting of rice in India. For forecasting rice yield in India, three different nonlinear models namely Monomolecular, Gompertz and Logistic, were used. All three non-linear models were fitted to data by using Statistical software R. For validation of assumptions of residuals i.e., randomness and normality of residuals, Run's test and Shapiro wilk's tests were employed respectively while for goodness of fit and validation of models, Chi-square test and eight steps ahead forecasting were done. For getting best best-fitted models for forecasting rice yield, models are compared by seven different statistics R^2, RSS, MAPE, MAE, MSE, RMSE,RSE So, after analysing the data, Logistic model is found better for forecasting of rice yield in India with FE% of 6.25% and 5.02 % for the year 2020 and 2021 respectively. Forecasted rice yield for the years 2023 and 2024, calculated by the logistic model and found 2.43 t/h and 2.67 t/h respectively. Forecasting model of rice yield for india is best fitted model (i.e. Logistic) as below.*

$\hat{Y} = 4.0048/(1+(4.0048/0.8612-1)*\exp(-0.0321*t))$

Keywords: *Nonlinear growth model, Forecasting, Run's test, Shapiro wilk's test.*

Introduction

Rice (Oryza sativa L.) is the most important crop of India [16]. It is grown on approximately 3.16 million hectares' area under rice in India. In India rice is grown in 43.86 million ha, the production level is 104.80 million tones and the productivity is about 2390 kg/ha [1]. India is the world's second-largest producer of rice and the largest exporter of rice in the world [2, 3, 4]. Production increased from 53.6 million tons [5] in 1980 to 120 million tons in 2020-21. Approximately 90% of the world population in Asia has the consumption of rice in its meal planning. In India, rice is cultivated in a large portion of the states, with West Bengal leading the way in production, followed by Uttar Pradesh, Andhra Pradesh, Punjab, Tamil Nadu, and Bihar [18]. Rice is a major food grain in India, where the yield is emulous with China, with more than 11% of the global production rate. Rice production has increased 3.5 times during the last 55 years, after the Green Revolution was imposed in India. Various traditional statistical methods were employed to predict the rice yield based on highly influential parameters, such as the area under cultivation, production and ultimately yield, that still resulted in a gap in measuring the accurate information [19]. Non linear techniques make it possible to implement means of predicting the rice yield by overcoming the limitations of traditional techniques [20, 21] and forecasting methods for current needs. The Non-linear techniques are playing a prominent role in dealing with such complex situations and making wise decisions in support of farmers as well as decision-makers.

Corresponding Author: **Mahesh Kumar*

DOI: https://doi.org/10.21276/AATCCReview.2025.13.03.532

Review and Literature

Most researchers have focused on developing traditional and advanced regression models in linear and nonlinear situations. Starting with the traditional multiple linear regression to predict the crop yield in Andhra Pradesh [11], kernel ridge, lasso, and elastic net regression models considering parameters such as the state, district, season, area, and year have been used to estimate the particular crop yield in India [9]. Applications of non-linear techniques play a vital role in handling rice yield forecasting. Based on accurate predictions by these techniques, farmers can plan how much area to take for particular crop production, as well as the yields of crops. A study intended to forecast the rice yield through non-linear techniques was taken care of in this study. Research was attempted to compare the performance of a non-linear model based on leaf data of maize [12]. Some non-linear models i.e. Richards model, Logistic model, Weibull model, MMF model and Gompertz model were used to fit with leaf data. Coefficient of determination(R^2), sum of squares error (SSE), root mean square error (RMSE), and mean relative error (MRE) were used to compare models [7]. The result showed that to determine leaf growth of maize Richards, Logistic and Gompertz model were better than other models under consideration.

The study was carried out on forecasting of production of groundnut using time series data from the period 1950-51 to 2011-12. Analysis was carried out using three different nonlinear models. Gauss- Newton algorithm was used to estimate the parameters of model and monomolecular and logistic model was found better then gompertz model [10, 22]. The Scientist [15] was conducted to compare six different non-linear growth models Logistic, Gompertz, Monomolecular, Richards, Quadratic and Reciprocal model fitness to data based on cotton in India. The data was collected from period of 1980-2013 for fitting models and R^2, RSS and RMSE were used check goodness of fit. To check the randomness and normality of error term "Run Test" and "Shapiro- Wilks" test were used respectively.

Materials And Methods

Forecasting of rice yield of India through non- linear growth models is based on the secondary data. Data was collected for the years 1963 to 2021 from the official sites of indiastat. For achieving objective, data from 1963 to 2016 were analysed through R- Software and five years of data 2017 and 2021 were kept for model validation of yield forecasting of rice in India. For forecasting rice yield in India, three different nonlinear models namely Monomolecular, Gompertz and Logistic, were used. All three non-linear models were fitted to data by using Statistical software R. For validation of assumptions of residuals i.e., randomness and normality of residuals, Run's test and Shapiro Wilk's tests were employed respectively and eight steps ahead forecasting were done. For getting best best-fitted models for forecasting rice yield, models are compared by different statistics $R^{2,}$RSS, MAPE, MAE, MSE, RMSE,RSE were used.

NON-LINEAR MODEL

Every statistical inquiry where principles from some body of knowledge enter into the analysis most likely will lead to a Nonlinear model [14]. These nonlinear models play a crucial role in getting complex relationship between variables. A model which exhibits nonlinearity for at least one parameter is called a nonlinear model. Occurrence of growth leads to the model needed for a particular time. The process involves making assumptions about the type of growth that is occurring, highlighting differential equations that represent those assumptions and then solving those equations to frame an appropriate model. The nonlinear models used in this study are; Monomolecular model; Gompertz model and Logistic model.

I) Monomolecular model

The model describes the progress of a growth at any time is proportional to the resources yet to be achieved. Differential equation of this model is given by

$\frac{dy}{dt}$= r (k-y)........................(i)

While its solution is given by ; Y(t)=k-(k-y0) *exp(-r*t) ....(ii)

II) Gompertz model

Differential equation of this model is expressed as

$\frac{dy}{dt}$= r*y*ln(k/y) ....(iii)

On integrating it, solution of equation will be; $Y_{(t)}$ =k*exp(ln(y0/k) *exp(-r*t) ....(iv)

III) Logistic model

The logistic model is appropriate in situation when growth rate is directly proportional to the product of the present size and the further amount of growth. it is given by

$\frac{dy}{dt}$= r*y(1-y/k) ....(v)

On integrating it, Y_t = k/(1+(k/y0-1)*exp(-r*t) ....(vi)

PARAMETER ESTIMATION

Let's consider nonlinear model represented by

$Y_t = f(x_t, \beta) + u_t$; t = 1, 2 ..., n ....(vii)

Y_t = a independent variable; X_t= explanatory variables; β = unknown parameters; u_t are error terms which are unobservable and identical, independent and follows normal distribution with mean zero and constant variance. The non-linear least-squares (NLLS) estimator is that estimator which minimize the value of sum of squared residuals and denoted by $\hat{\beta}$.

$$Sn(\beta)=\sum_{t=1}^{n}[Y_t - f(X_t, \beta)]^2$$

The differentiation is to be made of equation (3.10) with respect to β which will give normal equations and by solving them, we get values of β.

The normal equation can be expressed as

$$\sum_{t=1}^{n}\left[y_t - f\left(X_t, \hat{\beta}\right)\right]^2 \left(\frac{\partial f(X_t, \beta)}{\partial \beta_1}\right)_{\beta=\hat{\beta}} = 0$$

There doesn't exist any explicit formula to get nonlinear least squares estimator $\hat{\beta}$, so to minimize equation. Some iterative method to be used.So, some nonlinear estimation procedure like Gauss-Newton algorithm, steepest descent algorithm and Levenberg-Marquardt algorithm, can be employed to fit models. The detailed description of these algorithm is given by Scientist [5].Gauss-Newton algorithm is used in this study for fitting models and analysis is diner using R software. Validated the models for getting best fitted models for forecasting yield, models are compared by nine different statistics $R^{2,}$RSS, MAPE, MAE, MSE, RMSE The residuals are examined by run test [13, 17] and Shapiro wilk's test for randomness and normality assumptions respectively.

RESULTS AND DISCUSSION

For forecasting rice yield for India, three different nonlinear models have been used. The models have been fitted on rice yield data from period of 1963 to 2021 from the official sites of indiastat. For achieving objective, data from 1963 to 2016 were analysed through R- Software and five years data 2017 and 2021 were kept for model validation of yield forecasting of rice in India.

Fitting of different nonlinear model

Initial values for parameters of nonlinear model are required while doing estimation by fitting nonlinear model. The values are estimated by using Draper and Smith method [5] and used to nonlinear models. The estimated initial values are used in R software for fitting different nonlinear models, i.e., Monomolecular, Logistic and Gompertz model. The results of the model are given below.

Table- 1. Summary statistics for Monomolecular model of rice yield (t/ha) in India

Parameters	Estimate	Std. Error	t-value	Pr(>|t|)
K(Carrying capacity)	-2.564157	2.428073	-1.056	0.2959
y0(Initial value at time=0)	0.852225	0.039747	21.441	<2e-16 ***
r(Intrinsic growth rate)	-0.007136	0.004147	-1.721	0.0914 .

It is revealed from the above table that among three parameter y0 are statistically significant at 1% level of significance while r and k are not.

Table- 2. Summary statistics for Gompertz model of rice yield (t/ha) in India

Parameters	Estimate	Std. Error	t-value	Pr(>|t|)
K(Carrying capacity)	7.277683	3.664979	1.986	0.05245 *
y0(Initial value at time=0)	0.856159	0.035612	24.041	< 2e-16 ***
r(Intrinsic growth rate)	0.012491	0.004276	2.921	0.00519 **

It is revealed from the above table that among three parameters k is significant at 5 % . y0 and r are statistically significant at 1% level of significance.

Table 3. Summary statistics for Logistic model of rice yield (t/ha) in India

Parameters	Estimate	Std. Error	t-value	Pr(>|t|)
K(Carrying capacity)	4.004809	0.745862	5.369	1.96e-06 ***
y0(Initial value at time=0)	0.861205	0.032429	26.557	< 2e-16 ***
r(Intrinsic growth rate)	0.032104	0.004527	7.091	3.91e-09 ***

It is revealed from the above table that among three parameters k y0 and r all are statistically significant at 1% level of significance.

Criteria for Selecting the Good Fitted Model

To choose the best fitted model among three different nonlinear model used in study, seven different statistics were compared which are R^2 (coefficient of determination), Residual sum of squared (RSS), Mean absolute percentage error (MAPE), Mean absoluteerror (MAE), Mean square error (MSE), Root mean square Error (RMSE).

Table 7. Comparison of Monomolecular, Gompertz and Logistic models for rice productivity (t/ha) in India

Statistics	Monomolecular	Gompertz	Logistic
R^2	0.9965	0.9966	0.9966
RSS	0.4837	0.4761	0.4703
MAPE	5.574	5.524	5.4791
MAE	0.0785	0.0777	0.0772
MSE	0.0095	0.0093	0.0092
RMSE	0.0974	0.0966	0.0960
RSE	0.0983	0.0974	0.0968

From table -7, it can be concluded that Logistic model is equal to R^2 for two models (Gompertz model and Logistic) and higher than monomolecular . The value of RSS, MAPE, MAE,MSE are also low for Logistic model. So, it can be concluded that Logistic model is best fitted for rice yield in India.Fig.-1 shows the graph of actual v/s fitted yield for Logistic model where time in year (t) in x-axis and yield (t/ha) in y-axis. However, assumptions regarding randomness and normality of residuals are also checked using run test and Shapiro wilk's test respectively and showed in table- 4 and table- 5 respectively.

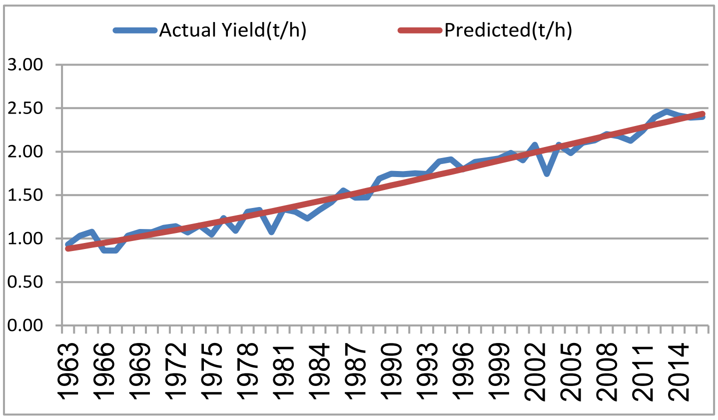

Fig -1. Actual yield vs Fitted yield of rice (Logistic model)

Examiantion of Residuals

Although R^2 is considered as effective statistic for adequacy of model [8] but sole reliance on it may not reveal important characteristics. So, one should always go through residuals or error examination while fitting any model. The residuals are the unexplained part of information which can expressed as difference between the observed and predicted values of fitted model. Usually, there are two assumption which are made for residuals:

I) Residuals are random

ii) Residuals are normally distributed

So, in the way to select model it must be cleared that residuals should follow the required assumptions. Run test and Shapiro-wilks (W) test are used for randomness and normality of residuals respectively. Given table shows the result of respective test performed on residuals.

Table 4. Randomness test for the residuals (Run test)

Models	Standard normal value	p-Value
Monomolecular	-1.560	0.965
Gompertz	-0.520	0.603
Logistic	-0.011	0.991

The table- 4 shows that z values for each applied model are not significant which reveals that we accept null hypothesis and conclude that residuals are random for all three model.

Table 5. Normality test for the residuals (Shapiro-wilks test)

Models	W (Statistic)	p-Value
Monomolecular	0.9323	0.0061
Gompertz	0.9386	0.0108
Logistic	0.9423	0.0151

On the basis of p-value showed in the above table it can be concluded that residuals are normally distributed which support the assumptions made regarding residuals.

Fitting of model for rice yield forecasting in India

All three model are compared based on different statistics for forecasting productivity of rice in India. It is found that the Logistic model is best suited for forecasting among three model. As per run test it is cleared that residuals are following random distribution while Shapiro wilk test revealed that residuals are not normally distributed. The parameter estimates for all the models for estimating the productivity of rice is given below.

For the Logistic model:

$\hat{Y}$= k / (1+ (k/y0-1)* exp(-r *t)).

$\hat{Y}$ =4.0048/ (1 + (4.0048/0.8612-1) *exp. (-0.0321*t))

Table 8. Forecasting yield (t/ha) using Logistic model for India

Year	Logistic model		
	Actual yield(t/h)	Forecasted yield(t/h)	% FE
2017	2.49	2.47	0.08
2018	2.58	2.50	3.10
2019	2.64	2.53	4.16
2020	2.72	2.56	5.88
2021	2.72	2.59	4.77
2022	------------	2.61	---------
2023	------------	2.64	----------
2024	--------------	2.67	--------------

In table- 8. Rice forecasted for India for year is 2024 i.e. 2.67 t/h. As computed values of Z- statistics for each model are not significant hence null hypothesis of randomness of errors is not rejected at 5% level of significance. Even W- statistics of Shapiro wilk's test are not significant for each model. Hence, it can be concluded that residuals follow normality.

Conclusions

The low values of RSS, MAPE, MAE, MSE, RMSE,RSE and almost equal value of R-Square are good inferences for forecasting rice yield in India. Eight years ahead forecasting from 2017 to 2024 for Rice yield was calculated using Nonlinear (Logistic) model. For getting best fitted models for forecasting rice yield, models are compared by different statistics R^2 , RSS, MAPE, MAE, MSE, RMSE,RSE So, after analysing the data, Logistic model is found better for forecasting rice yield in India with minimum FE% of 0.08 % and maximum 5 .88% for year 2017 and 2020 respectively. Forecasted rice yield for the years 2023 and 2024, calculated by the logistic model and found 2.64 t/h and 2.67 t/h respectively. For India out of three nonlinear models viz., Monomolecular, Logistic and Gompertz models Logistic is best one. To obtain best fitted model among three, different statistics viz, R^2,RSS, MSE, MAE, RMSE, MAPE, RSE are compared. Based on comparison of statistics , Logistic model is found best fitted model for India . By Logistic model, Rice yield is forecasted for upcoming years i.e for the eight year ahead yield forecasting of Rice in India for the year 2017 to 2024 are 2.47t/ha, 2.50 t/ha, 2.53 t/ha, 2.56 t/ha 2.59 t/ha 2.61, 2.64 t/ha,2.67t/ha respectively with forecast error of minimum 0.08% and maximum 5.88%.

Future scope of the study

The present study provides a solid foundation for forecasting rice yield in India using non-linear growth models, and offers multiple avenues for future exploration. Integrating real-time data sources such as satellite imagery, weather patterns, and soil health metrics through IoT and remote sensing can significantly enhance the accuracy and timeliness of yield predictions. Expanding the model's application from national to regional or district levels could support localized planning and agricultural interventions. Incorporation of advanced machine learning algorithms like neural networks, random forests, or LSTM could be explored for comparative performance enhancement over traditional non-linear models. Additionally, coupling yield forecasts with relevant climatic, agronomic, and economic variables in multivariate frameworks would allow for more robust and holistic predictions. The study's outputs can also be effectively used in policy formulation, food security planning, and agricultural insurance schemes. Evaluating the model's resilience under various climate change scenarios may provide insights into long-term agricultural sustainability. Developing user-friendly decision support systems or mobile applications based on the best-fit model could aid stakeholders such as farmers, planners, and policymakers. Furthermore, the methodology can be extended to other crops, enabling cross-crop comparisons and improved resource management. Continuous model refinement through incorporation of new data will be vital to maintaining the reliability and applicability of forecasts over time.

Acknowledgment

The authors express their gratitude to the Dean, College of Basic Sciences and Humanities (CBS&H), Rajendra Prasad Central Agricultural University, Pusa, for approving the research work. We are also deeply appreciative of the technical staff of the Department, whose invaluable assistance greatly contributed to the success of this study. Their unwavering support and commitment were integral to the completion of this research.

Conflict of interest

The authors declare that they have no conflict of interest

References

1. Food grain production estimated to reach new high of 150 million tonnes in 2021-22: Centre". *The Indian Express. 22 September 2021.* Archived *from the original on 22 October 2021. Retrieved 12 October 2021*

2. Kumar, Ajeet (2016) Soil Health and Soil Quality: - The Fundamental Keys to a Sustainable and Productive Agriculture; *Pg. 48-54;* Souvenir *of International Symposium on Management of Rice based agricultural system under stress prone environment;* organized by Rajendra Agricultural University, Pusa, Samastipur, Bihar during 17 - 19 March, 2016. https://hal.science/hal-04972899.

3. India becomes world's largest exporter of Rice: List of top exporters. *India Today. 28 January 2016.* Archived *from the original on 29 October 2021.* Retrieved 12 October 2021, Agricultural Statistics at a glance- 2015

4. Kumar A., Sinha S.K., Singh S.K., Rana L., Singh A.K., Kumari S., Kumar A., Singh H. and Paswan S. (2025) Influence of Intercropping and Planting Techniques on Sugarcane Yield and Nutrient Absorption in the North West Alluvial Plains of Bihar. *AATCC Review*, 13 (1): 144-153. https://doi.org/10.21276/AATCCReview.2025.13.01.144. ; https://hal.science/hal-04962582.

5. Draper, N.R. and Smith, H. (1998).Applied Regression rd Analysis.3 Edn.John Wiley & Sons. U.S.A., pp. 543-49.

6. Heitzman, James; Worden, Robert L., eds. (1996). "Food-Crop Production". India: a country study (5th ed.). Washington, D.C.Federal Research Division, Library of Congress. pp. 406-407. ISBN 0-8444-0833-6. OCLC 34598209.

7. Karadavut, U., Palta, Ç.,Kokten, K., &Bakoğlu, A. (2010). Comparative study on some non-linear growth models for describing leaf growth of maize. *International Journal Of Agriculture & Biology,*12,227–230

8. Kvålseth, T. O. (1985). Cautionary note about R 2. *The American Statistician*, *39*(4),279-285.

9. Nishant, P.S.; Venkat, P.S.; Avinash, B.L.; Jabber, B. Crop Yield Prediction based on Indian Agriculture using Machine Learning. In Proceedings of the 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 5-7 June 2020; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2020; pp. 1–4

10. Pal, S., and Mazumdar, D. (2015). Forecasting groundnut production of India using nonlinear growth models.*Journal Crop and Weed*, 11, 67-70.

11. Ramesh, D.; Vardhan, B.V. Analysis of crop yield prediction using data mining techniques. Int. J. Res. Eng. Technol. 2015, 4, 47–473.

12. Kumar, Ajeet and Mukul Kumar (2017) Effect of different levels of phosphorus on the yield and yield components of maize under agro-climatic zone- II of Bihar. *International Journal of Agricultural Sciences.* 13 (2): 266-270. https://hal.science/hal-04676729

13. Ratkowsky, D. A. (1990). Handbook of nonlinear regression models, parcel Dekker. Inc., New York and Basel.

14. Seber, G. A., & Wild, C. J. (2003).Nonlinear regression.hoboken. New Jersey: John Wiley & Sons, 62(63), 1238.

15. Sundar Rajan, M., &Palanivel, M. (2017). Estimation of non-linear models for fitting growth curves of cotton in India.*International Journal of Statistics and Applied Mathematics*; 2(4): 86-93.

16. Paswan, D., Arun, A., Kumar, A. and Choudhary, C.S. (2014) Effect of NPK uptake at different growth stages of wheat (*Triticum aestivum* L.) for yield maximization. *An Asian Journal of Soil Science* (An International Refereed Research Journal). 9 (2): 265-270. https://hal.science/hal-04676702.

17. Kumar, A., Singh, S. K., Meena, S. K., Sinha, S., Rana, L., Kumari, A., Singh, A.K., & Kumar, A. (2024). Human Health Risk Assessment of Nitrate-Contaminated Groundwater in North East Alluvial Plains of India. In Book: Geography, Earth Science and Environment. *Geography, Earth Science and Environment: Research Highlights Vol. 2*, 1–26. https://doi.org/10.9734/bpi/geserh/v2/1424

18. Singh, S.K., Tagung, T., Kumar, A., Singh, P., Singh, H., Kashiwar, S. R., Tiwari, S., Singh, A.K., Kumari, S. and Singh. Y.V. (2024). Geo Spatial Assessment and Micronutrient Mapping of Calcareous Soils in Muzaffarpur District, Bihar, India. *Journal of Scientific Research and Reports*; 30 (11):155-65. https://doi.org/10.9734/jsrr/2024/v30i112543.

19. Kumar, A., Meena, S. K., Singh S. K. and Sinha S.K. (2025) Revolutionizing Agriculture through Advances in Biofertilizer Formulations and Innovative Application Systems. *New Era Agriculture Magazine,* 3 (11): 156-158. https://neweraagriculture.com/wp-content/uploads/2025/05/Article-26.pdf; https://hal.science/hal-05062703.

20. Kumar A., Bairwa M., Meena S. K., Singh S. K., Sinha S.K., Rana L., Singh Y. V., Kumar A., Singh H. and Paswan S. (2025) Comparative Analysis of Macronutrient, Soil Sulphur Availability and Micronutrient Dynamics across Diverse Land Use Types of Indo-Gangetic plains of India. AATCC Review; 13 (01): 467-472. ; https://hal.science/hal-05081536; https://doi.org/10.21276/AATCCReview.2025.13.01.467.

21. Sinha, S.K., Ajeet Kumar, Brajesh Kumar , Amrita Kumari and A.K. Singh (2024) Influence of Organic-Inorganic Interaction on Nutrient Use Efficiency and Sugarcane Yield in Calcareous Soils of the Indo-Gangetic Plains, India; In Book: Current Research Progress in Agricultural Sciences. Volume.3. 22 August 2024, Page 1-22. https://doi.org/10.9734/bpi/crpas/v3/1426.

22. Kumar A and Singh, A. K. (2022) Effect of climate change on soil microorganism in perspective of living soils and plant-microbe interaction. Pg No. 89-101. In Book: Agriculture and Climate Change-Impacts and Solutions, edited by A. Kumar, N. Kumar, G. Kumar, G. Kumari, L. Rana, S.K. Sahoo (2022). ISBN: 978-93-92725-46-3, Brillion Publishing 22 B/ ground floor, Desh Bandhu Gupta Road, Karol Bagh, New Delhi-110005; https://hal.science/hal-04881411.

