

Journal of Agrometeorology

ISSN : 0972-1665
Vol. No. 24 (1) : 50-54 (March 2022)

<https://journal.agrimetassociation.org/index.php/jam>



Research Paper

Extreme learning machines for weather-based modelling of silk cocoon production

PRAMIT PANDIT, BISHVAJIT BAKSHI* and SHILPA M.

Department of Agricultural Statistics, Applied Mathematics and Computer Science, University of Agricultural Sciences, Bengaluru-560065, Karnataka, India

*Corresponding author email : bishvajitb93@gmail.com

ABSTRACT

In spite of the immense popularity and sheer power of the neural network models, their application in sericulture is still very much limited. With this backdrop, this study evaluates the suitability of neural network models in comparison with the linear regression models in predicting silk cocoon production of the selected six districts (Kolar, Chikballapur, Ramanagara, Chamarajanagar, Mandya and Mysuru) of Karnataka by utilising weather variables for ten consecutive years (2009-2018). As the weather variables are found to be correlated, principal components are obtained and fed into the linear (principal component regression) and non-linear models (back propagation-artificial neural network and extreme learning machine) as inputs. Outcomes emanated from this experiment have revealed the clear advantages of employing extreme learning machines (ELMs) for weather-based modelling of silk cocoon production. Application of ELM would be particularly useful, when the relation between production and its attributing characters is complex and non-linear.

Keywords: BP-ANN, cocoon production, ELM, principal component regression, silk, weather-based modelling

Silk production is a trait unique to the arthropods with approximately 98% of the Lepidoptera species producing some form of silk (Craig, 1997). Silks are developed inside but performed outside the body, necessitating specialisations in both internal processing and environmental effects. Since silkworms are ectothermic, the environment during spinning has a significant impact on the silk production (Offord *et al.*, 2016). Other than temperature, humidity is another environmental factor known to influence the insect behaviour and metabolism (Manoj and Anil, 2016; Sharma *et al.*, 2017). A sudden rise in temperature or humidity can cause abrupt thinning or thickening of the silk filament (Ramachandra *et al.*, 2001).

Sericulture in India is a farm-based, labour-intensive commercial enterprise that is well-suited to rural farmers offering high employment and lucrative returns. Multivoltine races have dominated the Indian sericulture industry with 92% of the total silk production (Shilpa, 2018). Keeping the economic importance of sericulture as well as the impact of weather variables in its production scenario in mind, several researchers have given painstaking efforts in developing prediction models based on environmental factors. Rahmathulla (2012) has emphasised the influence of temperature and humidity on post cocoon parameters. Saikia *et al.* (2016) have

provided an extensive review on the factors affecting the muga silkworm production. Sisodia and Gaherwal (2017) have discussed the temperature and humidity thresholds for the cocoon traits.

Of late, neural network models are drawing substantial attention in the area of prediction modelling, where regression and other related statistical models have traditionally been applied. Panda *et al.* (2010) have employed neural network techniques for corn yield prediction in the Oakes Irrigation Test Area research site of North Dakota, USA. Yang *et al.* (2009) have employed BP-ANN (back propagation-artificial neural network) methodology to predict pest outbreaks while utilising principal components (PCs) of weather variables as inputs. Sajjadi *et al.* (2016) have observed that extreme learning machine (ELM) can produce good generalisation performance in most cases and can learn thousands of times faster than the conventional popular learning algorithms. Li and Wang (2019) have evaluated the performance of the multiple linear regression and ANN models in characterising concrete dam deformation under environmental loads. The superiority of ELM models in providing better prediction has been evident in their study.

Even though regression and neural network modelling are well-recognised as competitive model-building strategies in the

Article info - DOI : <https://doi.org/10.54386/jam.v24i1.946>

Received: 3 April 2021; Accepted: 6 January 2022; Published online: 11 February 2022

This work is licenced under a Creative Common Attribution 4.0 International licence @ Author(s), Publishing right @ Association of Agrometeorologists

literature, the applications of neural networks are still very much limited in sericulture. With this backdrop, an effort has been made in this paper to evaluate the suitability of neural network models in comparison with the linear regression models in predicting silk cocoon production of Karnataka by utilising weather variables.

MATERIALS AND METHODS

Study location and datasets

Karnataka is the highest silk producing state of India, accounting for approximately 30% of its exports. As sericulture in Karnataka is primarily concentrated in the southern region, Kolar, Chikballapur, Ramanagara, Chamarajanagar, Mandya and Mysuru districts are selected for the current investigation. Ten years' (2009-2018) monthly data on silk cocoon production ('00 MT) of bivoltine and multivoltine races pertaining to the selected districts have been collected from the Department of Sericulture, Government of Karnataka. Weekly data on five weather parameters, viz., maximum temperature (°C), minimum temperature (°C), morning relative humidity (%), evening relative humidity (%) and rainfall (mm) for the same period have been obtained from the All India Coordinated Research Project on Agrometeorology, Bengaluru centre, Gandhi Krishi Vigyana Kendra (GKVK), Bengaluru.

Model development

As the first step towards model development, all the weather data are standardised. Secondly, the correlation between monthly production data (bivoltine, multivoltine and total, respectively) and instar-wise weather observations are computed. Correlation between monthly production data and monthly (averaged) weather data are also obtained and compared with the instar-wise results. As a significantly higher correlation is observed for the latter one in all cases, monthly (averaged) weather data are utilised for the analysis.

Principal component analysis

Principal component analysis (PCA) is a statistical procedure that utilises an orthogonal transformation to convert a set of observations of possibly correlated variables into a (hopefully, smaller) set of values of linearly uncorrelated variables called principal components. In regression analysis, if the number of explanatory variables is relatively large or if there are substantially higher correlations among the explanatory variables, PCA is usually carried out to result a better test and more stable estimates of the regression coefficients. Bartlett's sphericity test, which compares a correlation matrix with the identity matrix to check for the possible presence of redundancy between the variables, is utilised to adjudge the applicability of PCA.

For a better interpretation of the factor loadings, PCs are obtained by varimax rotation in this study. As the retention criterion, principal components with Eigen values higher than one are considered for the regression analysis.

Principal component regression

Multiple linear regression is basically the extension of

simple linear or ordinary least squares (OLS) regression by allowing more than one explanatory variable to rely on the mean function $E(Y)$ (Fumo and Biswas, 2015). A usual representation of multiple linear regression is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

where, Y and X_i ($i = 1, 2, \dots, k$) represent the response and explanatory variables, respectively. β_i ($i = 0, 1, \dots, k$) are model parameters and ε represents the random error term.

In this study, as the principal component regression (PCR) model building strategy, the retained PCs are used as the explanatory variables of the full model. Variable selection and model building are then carried out simultaneously as per the stepdown procedure.

Back propagation-artificial neural network

ANN can be depicted as the abstraction and simulation of the human brain with the ability of complex parallel information processing. The most widely used ANN is BP-ANN, which is basically a multi-layer feed-forward network based on an error back-propagation (BP) learning algorithm. The usual form of a BP-ANN with p input nodes, q hidden nodes and one output node is represented as:

$$y_i = g\left\{\alpha_0 + \sum_{k=1}^q \alpha_k f\left(\beta_{0k} + \sum_{j=1}^p \beta_{jk} x_j\right)\right\}$$

where, $i = 1, 2, \dots, n$ and y_i denotes the i^{th} output. Activation function at hidden and output layer is denoted by f and g , respectively. β_{jk} denotes the weight attached to the connection between j^{th} input node and the k^{th} hidden node and similarly, α_k denotes the weight attached to the connection from k^{th} hidden node to the output node. β_{0k} and α_0 represent the bias at the input and hidden layer, respectively. Neural network models with a single hidden layer are considered for this study as per the universal approximation theorem (Karsoliya, 2012).

Extreme learning machine

For the past few years, unsatisfactory learning speed has remained as the major bottleneck in the application of feed-forward neural networks. Slow gradient-based learning algorithms and iterative tuning of the network parameters are thought to be the key reasons behind this downside. Huang *et al.* (2006) have proposed a new learning algorithm called ELM, which,

Table 1: Results of Bartlett's sphericity test

District	χ^2 statistic	P value
Kolar	192.89	<0.001
Chikballapur	226.32	<0.001
Ramanagara	175.37	<0.001
Chamarajanagar	197.66	<0.001
Mandya	207.28	<0.001
Mysuru	187.20	<0.001

Table 2: Loadings of the selected principal components

District	Principal component	T _{max}	T _{min}	Morning RH	Evening RH	Total rainfall	% of variance explained
Kolar	PC ₁	-0.45	-0.19	0.93	0.92	0.22	39.93
	PC ₂	0.72	0.92	-0.26	-0.28	-0.02	30.12
	PC ₃	0.29	-0.19	0.16	0.21	0.95	21.78
Chikballapur	PC ₁	-0.62	-0.22	0.88	0.88	0.31	41.99
	PC ₂	0.55	0.95	-0.23	-0.24	-0.09	26.42
	PC ₃	0.41	-0.15	0.34	0.35	0.88	24.08
Ramanagara	PC ₁	-0.26	-0.23	0.92	0.92	0.25	37.05
	PC ₂	0.87	0.88	-0.29	-0.27	0.01	33.66
	PC ₃	-0.02	0.01	0.19	0.24	0.97	20.58
Chamarajanagar	PC ₁	-0.03	-0.27	0.98	0.97	0.10	39.76
	PC ₂	0.92	0.86	-0.16	-0.14	-0.03	32.67
	PC ₃	-0.12	0.07	0.05	0.11	0.99	20.28
Mandya	PC ₁	-0.17	-0.11	0.97	0.98	0.12	39.29
	PC ₂	0.87	0.89	-0.14	-0.16	-0.08	32.09
	PC ₃	-0.12	-0.01	0.12	0.07	0.99	20.19
Mysuru	PC ₁	-0.17	-0.30	0.94	0.92	0.24	38.22
	PC ₂	0.92	0.84	-0.22	-0.29	0.03	33.54
	PC ₃	-0.12	0.18	0.19	0.19	0.96	20.84

The values marked in bold indicate the high correlation between the monthly (averaged) weather variables and the corresponding PCs

Table 3: Parameter estimates of the selected PCR models

District	Explanatory variable	Multivoltine		Bivoltine		Total	
		Parameter estimate	P value	Parameter estimate	P value	Parameter estimate	P value
Kolar	PC ₂	-0.323	0.005	-0.266	0.030	-0.325	0.005
	PC ₃	-0.427	0.000	-0.312	0.012	-0.425	0.000
Chikballapur	PC ₁	-0.265	0.025	-0.252	0.045	-0.265	0.025
	PC ₂	-0.234	0.047	-	-	-0.229	0.042
	PC ₃	-0.344	0.004	-0.239	0.049	-0.341	0.005
Ramanagara	PC ₁	-0.328	0.004	-0.291	0.014	-0.329	0.004
	PC ₂	-0.232	0.037	-	-	-0.226	0.042
	PC ₃	-0.404	0.001	-0.383	0.002	-0.406	0.001
Chamarajanagar	PC ₁	-0.339	0.003	-0.307	0.014	-0.340	0.003
	PC ₂	-0.306	0.007	-	-	-0.303	0.009
	PC ₃	-0.314	0.006	-0.250	0.042	-0.295	0.010
Mandya	PC ₁	-0.266	0.025	-0.339	0.004	-0.271	0.022
	PC ₃	-0.387	0.001	-0.388	0.001	-0.388	0.001
Mysuru	PC ₃	-0.307	0.012	-0.302	0.011	-0.308	0.012
	PC ₃	-0.222	0.036	-0.342	0.004	-0.226	0.046
	PC ₃	-0.240	0.047	-0.214	0.026	-0.240	0.047

unlike the conventional implementations, selects hidden nodes randomly and determines the output weights of single hidden layer feed-forward neural networks (SLFNs) analytically. There are several empirical evidences (Albadra and Tiun, 2017) suggesting that ELM can provide satisfactory generalisation performance at an extremely fast learning speed. For fixed input weights (w_i ; $i=1,2,\dots,\tilde{N}$) and the hidden layer biases (b_i ; $i=1,2,\dots,\tilde{N}$), training an SLFN is analogous to finding a least squares solution ($\hat{\beta}$) of the linear system, $H\hat{\beta}=T$:

$$\|H(w_1, w_2, \dots, w_{\tilde{N}}, b_1, b_2, \dots, b_{\tilde{N}})\hat{\beta} - T\|$$

$$= \min_{\beta} \|H(w_1, w_2, \dots, w_{\tilde{N}}, b_1, b_2, \dots, b_{\tilde{N}})\beta - T\|$$

where, H represents the hidden layer output matrix.

Measures of model accuracy

The predictive ability of the fitted models under study are evaluated in terms of coefficient of determination (R^2), root mean square error (RMSE) and mean absolute percentage error (MAPE) (Stangierski *et al.*, 2019).

Table 4: Comparative results of the prediction models

District	Model	Multivoltine			Model	Bivoltine			Model	Total		
		R ²	RMSE	MAPE		R ²	RMSE	MAPE		R ²	RMSE	MAPE
Kolar	PCR	0.72	0.643	20.532	PCR	0.74	0.014	8.821	PCR	0.73	0.651	16.884
	BP-ANN <3,4,1>	0.97	0.360	10.924	BP-ANN <3,6,1>	0.99	0.007	5.471	BP-ANN <3,6,1>	0.98	0.315	7.352
	ELM <3,7,1>	0.99	0.130	4.078	ELM <3,11,1>	0.99	0.003	1.549	ELM <3,8,1>	0.99	0.134	3.990
Chikballapur	PCR	0.78	0.618	23.627	PCR	0.73	0.014	358.41	PCR	0.85	0.652	21.187
	BP-ANN <3,2,1>	0.97	0.354	12.917	BP-ANN <3,5,1>	0.89	0.007	259.01	BP-ANN <3,4,1>	0.97	0.315	9.502
	ELM <3,4,1>	0.99	0.126	4.725	ELM <3,8,1>	0.92	0.003	60.706	ELM <3,6,1>	0.99	0.134	4.848
Ramanagara	PCR	0.70	0.649	17.476	PCR	0.61	0.016	40.865	PCR	0.74	0.729	19.502
	BP-ANN <3,7,1>	0.90	0.379	11.627	BP-ANN <3,5,1>	0.81	0.009	24.382	BP-ANN <3,6,1>	0.98	0.376	11.697
	ELM <3,11,1>	0.94	0.148	4.084	ELM <3,7,1>	0.88	0.004	10.845	ELM <3,10,1>	0.99	0.141	3.974
Chamarajanagar	PCR	0.76	0.017	22.289	PCR	0.73	0.003	17.725	PCR	0.75	0.018	22.520
	BP-ANN <3,3,1>	0.96	0.009	11.123	BP-ANN <3,6,1>	0.98	0.002	13.631	BP-ANN <3,5,1>	0.97	0.009	13.300
	ELM <3,5,1>	0.99	0.003	4.420	ELM <3,9,1>	0.99	0.001	4.253	ELM <3,8,1>	0.99	0.004	4.459
Mandya	PCR	0.71	0.865	18.493	PCR	0.66	0.037	23.369	PCR	0.63	0.993	22.847
	BP-ANN <3,4,1>	0.96	0.548	12.586	BP-ANN <3,8,1>	0.92	0.019	12.606	BP-ANN <3,6,1>	0.93	0.514	10.712
	ELM <3,8,1>	0.98	0.181	4.038	ELM <3,11,1>	0.95	0.005	3.439	ELM <3,10,1>	0.99	0.178	3.986
Mysuru	PCR	0.69	0.882	17.709	PCR	0.73	0.025	18.077	PCR	0.75	0.720	14.618
	BP-ANN <3,5,1>	0.91	0.494	11.132	BP-ANN <3,8,1>	0.94	0.013	10.499	BP-ANN <3,7,1>	0.94	0.533	11.144
	ELM <3,6,1>	0.96	0.193	4.213	ELM <3,10,1>	0.98	0.003	2.629	ELM <3,9,1>	0.97	0.197	4.056

RESULTS AND DISCUSSION

The results of the Bartlett's sphericity test, as presented in the Table 1, confirm the appropriateness of PCA application to the monthly (averaged) weather data. In the next step, PCA is carried out by utilising the five monthly (averaged) weather parameters for feature extraction of the data. In order to evaluate the sensitivity of the attributing parameters, principal component loadings are computed. These loadings in effect reflect the correlations among the principal component scores and the attributing parameters. Table 2 provides the loadings of the selected PCs. For all the districts, the selected PCs cumulatively account for more than 90 percent of the variations of the respective input data.

The principal component score values are obtained by multiplying the Eigen vectors with the standardised weather variables in a linear fashion and these scores are further utilised for predicting the cocoon production. The stepdown regression procedure identifies the best model to have a significant linear relationship towards the cocoon production in each case. The obtained variance inflation factors (VIFs) for the selected best PCR models have also justified the use of PCA by indicating the absence of multicollinearity problem. Table 3 provides the parameter estimates of the selected PCR models.

Considering the non-linear pattern of the cocoon production, BP-ANN and ELM are employed in this study. For both the models, 12 out of 120 data points are kept for testing and validation purposes while utilising the rest for training. The input nodes are varied from 1 to 3 PCs. The number of hidden nodes, with the aid of basic cross-validation method, is varied from 2 to 14. Parsimony has been given the due weightage other than training and testing performance while selecting the best model in each case.

A model with a higher R² and smaller RMSE and MAPE values is considered to be the best. The comparative assessment of the prediction models under study, as presented in Table 4, has clearly revealed that even though both non-linear models have performed substantially better than the respective PCR model, ELM model has outperformed its counterparts in all cases. Fig. 1 presents the scatter plot of the observed and ELM-predicted silk cocoon production over all the areas under study along with R², RMSE and MAPE values.

The PCR model has the ability to extract only the linear pattern, whereas BP-ANN as well as ELM can capture the non-linear relationship between cocoon production and weather variables. It is noteworthy to mention at this juncture that in terms of R², the performance of BP-ANN and ELM is almost a tie, whereas

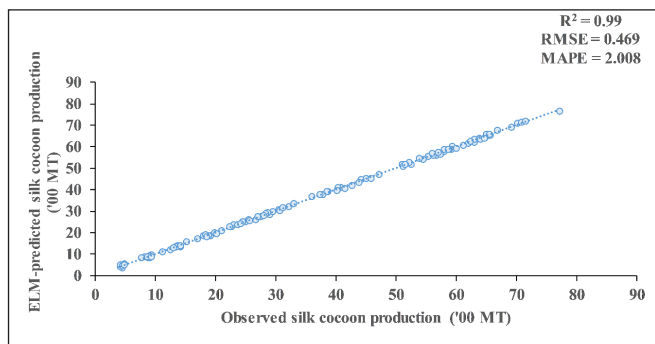


Fig. 1 : Scatter plot of the observed and ELM-predicted silk cocoon production over the study area during 2009-2018

the superiority of ELM over BP-ANN has been evident in terms of RMSE and MAPE values. Hence, the comparative assessment of the prediction models crucially depends on the accuracy measure as well.

CONCLUSION

Cocoon production in all the selected districts has exhibited a non-linear relationship with the weather variables. The ability of linear and non-linear models to predict cocoon production is compared in this study by employing PCR, BP-ANN and ELM. ELM has provided uniformly better results for all the cases considered. Application of ELM would be particularly useful, when the relation between production and its attributing characters is complex and non-linear.

ACKNOWLEDGEMENTS

Authors are thankful to University of Agricultural Sciences, Bengaluru for providing the facilities to carry out the investigation.

REFERENCES

- Albadra, M.A.A., and Tiun, S. (2017). Extreme learning machine: a review. *Int. J. Appl. Eng. Res.*, 12(14): 4610-4623.
- Craig, C. (1997). Evolution of arthropod silks. *Ann. Rev. Entomol.*, 42: 231-267.
- Fumo, N., and Biswas, M.R. (2015). Regression analysis for prediction of residential energy consumption. *Renew. Sustain. Energy Rev.*, 47: 332-343.
- Huang, G.B., Zhu, Q.Y., and Siew, C.K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3): 489-501.
- Karsoliya, S. (2012). Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *Int. J. Eng. Trends Technol.*, 3(6): 714-717.
- Li, M., and Wang, J. (2019). An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modelling. *Math. Probl. Eng.*, 19: 1-13.
- Manoj, K., and Anil, G. (2016). Effect of weather variables on whitefly (*Bemisia tabaci* Gennadius) population in development of potato apical leaf curl virus disease. *J. Agrometeorol.*, 18(2): 288-291.
- Offord, C., Vollrath, F., and Holland, C. (2016). Environmental effects on the construction and physical properties of *Bombyx mori* cocoons. *J. Mater. Sci.*, 51(24): 10863-10872.
- Panda, S.S., Ames, D.P., and Panigrahi, S. (2010). Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sens.*, 2(3): 673-696.
- Rahmathulla, V.K. (2012). Management of climatic factors for successful silkworm (*Bombyx mori* L.) crop and higher silk production: a review. *Psyche*, 12: 1-12.
- Ramachandra, Y.L., Bali, G., and Rai, S.P. (2001). Effect of temperature and relative humidity on spinning behaviour of silkworm (*Bombyx mori* L.). *Indian J. Exp. Biol.*, 38: 87-89.
- Saikia, M., Ghosh, K., and Peigler, R.S. (2016). Factors affecting on quality muga silkworm (*Antheraea assamensis* Helfer) seed crop production: a review. *J. Entomol. Zool. Stud.*, 4(6): 806-10.
- Sajjadi, S., Shamshirband, S., Alizamir, M., Yee, L., Mansor, Z., Manaf, A.A., Altameem, T.A., and Mostafaiepour, A. (2016). Extreme learning machine for prediction of heat load in district heating systems. *Energy Build.*, 122: 222-227.
- Sharma, S., Kooner, R., Sandhu, S.S., Arora, R., Kaur, T., and Kaur, S. (2017). Seasonal dynamics of insect pests of sugar beet under sub-tropical conditions. *J. Agrometeorol.*, 19(1): 81-83.
- Shilpa, M. (2018). Statistical Evaluation of Climate on Silk Cocoon Production in Karnataka. (Master's dissertation, University of Agricultural Sciences, Bengaluru).
- Sisodia, N.S., and Gaherwal, S. (2017). Effects of temperature and relative humidity on commercial product of silkworm (*Bombyx mori* L.) in Indore region of (MP) India. *Int. J. Zool. Stud.*, 2(5): 52-55.
- Stangierski, J., Weiss, D., and Kaczmarek, A. (2019). Multiple regression models and Artificial Neural Network (ANN) as prediction tools of changes in overall quality during the storage of spreadable processed Gouda cheese. *Eur. Food Res. Technol.*, 245(11): 2539-2547.
- Yang, L.N., Peng, L., Zhang, L.M., Zhang, L.L., and Yang, S.S. (2009). A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on back propagation artificial neural network and principal components analysis. *Comput. Electron. Agric.*, 68(2): 200-206.