Indian Journal of Agricultural Sciences **90** (10): 1952–8, October 2020/Article

Weather based forecasting of sterility mosaic disease in pigeonpea (*Cajanu cajan*) using machine learning techniques and hybrid models

RANJIT KUMAR PAUL[1], S VENNILA[2], SATISH KUMAR YADAV[3], M N BHAT[4], M KUMAR[5], P CHANDRA[6], A K PAUL[7] and M PRABHAKAR[8]

ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110 012, India

Received: 16 Janaury 2020; Accepted: 17 February 2020

ABSTRACT

Modelling incidence of sterility mosaic disease (SMD) on pigeonpea [*Cajanus cajan* (L.) Millsp.] for four locations [S K Nagar (Gujarat), Gulbarga (Karnataka), Rahuri (Maharashtra) and Vamban (Tamil Nadu)] was carried out based on field data sets generated during six *kharif* seasons [2011-16]. Mean seasonal incidence amongst all locations was on the decline during recent periods (0.5-5.3%) over past decades (9.8-12.8%). Correlation analyses of SMD incidence with weather parameters lagged one and two weeks indicated spatial differences for the variables besides their significance. While Max T (ºC) lagged by one week alone was significantly positive with SMD at Gulbarga (KA), Vamban (TN) had negative significance of rainfall (mm/week) and rainy days. S K Nagar (GJ) and Rahuri (MH) had shown opposite effects of both morning and evening RH (%) of both one and two lagged weeks. Support vector regression (SVR), artificial neural network (ANN) models and their combination with autoregressive integrated moving average (ARIMA) models applied for prediction of SMD incidence across locations revealed performance of hybrid models in general to be better based on the evaluation criteria of root mean square error (RMSE). ARIMA-SVR>ARIMA-ANN>SVR>ANN was the order of prediction accuracies at S K Nagar (GJ), Gulbarga (KA), and Vamban (TN). At Rahuri (MH), individual models performed better over their hybrids with ARIMA. While application of hybrid model of SVR-ARIMA is applicable under situations of SMD seasonal mean severity exceeding 1%, SVR model proves better for mean seasonal disease incidence in decimal values less than one.

Key words: ANN, ARIMA, Pigeonpea, SMD, SVR, Weather variables

Pigeonpea (*Cajanus cajan* (L.) Millsp.) is an important pulse crop of semi-arid and subtropic regions, *viz.* Asia and Africa (Van Der Maesen 1990). Pigeonpea is popularly known by different names such as redgram, *tur* or *arhar*. Pigeonpea is primarily grown for its protein (20 to 30%) particularly in developing countries including India. India is leading in pigeonpea cultivation with an area of 3.81 million ha with an annual production of 3.02 million tonnes. Indian states, *viz.* Maharashtra, Karnataka, Madhya Pradesh, Andhra Pradesh, Gujarat, Tamil Nadu and Uttar Pradesh grow pigeonpea. Although India leads the world in both area and production, lower pigeonpea productivity over global average is attributed to various abiotic and biotic constraints. Nene *et al.* (1981) listed about 50 diseases occurring in mild to severe form on pigeonpea which includes fungi, bacteria, viruses, *phytoplasmas* and nematodes. Of these *Fusarium* wilt, SMD and *Phytophthora* blight are major constraints to pigeonpea production. About 15 viruses are reported to infect pigeonpea (Kumar *et al*. 2008) and SMD caused by pigeonpea sterility mosaic virus (PPSMV) (Jones *et al.* 2004) is economically the most important viral disease in India causing an estimated annual loss of more than US$300 million (Reddy *et al.* 1998).The population of *A. cajani* and incidence of SMD were found to be positively correlated (Lakshmikantha and Prabhuswamy 2002).

Weather prevalent during cropping season also influence the incidence of SMD with its influence through the vector as well as on the virus. Relative humidity and average temperature of about 20-30°C was found to be congenial for the multiplication of mite that transmitted SMD (Kaushik *et al*. 2013). Deducing weather based relations and predictions of SMD can help farmers in taking decisive action on management strategies. Establishing pest-weather relations needs precise and reliable statistical tools. While multiple linear regressions (MLR) are suitable for short or intermediate term, ARIMA model (Box and Jenkins 1970) is a forecasting technique that projects future

Present address: [1]Senior Scientist (ranjitstat@gmail.com), ICAR-IASRI; [2]Principal Scientist (svennila96@gmail.com), [3]Research Associate (akash@gmail.com), [2]Principal Scientist (mnbhat.ncipm@icar.org.in), [5,6]Senior Research Fellow (murari.kumar14@gmail.com, pchausali28@gmail.com), ICAR-NCIPM; [7]Principal Scientist (amrit.paul@icar.gov.in) ICAR-IASRI, New Delhi. [8]Principal Scientist (prab249@gmail.com), ICAR-CRIDA.

values of a series based entirely on its own inertia. Various studies exist in literature for forecasting crop yields with linear and nonlinear techniques but the prominent ones of linear and nonlinear category are ARIMA and ANN, respectively (Zhang 1998). ARIMA and ARIMA with exogenous variables (ARIMAX) have been applied for forecasting agricultural prices and yield predictions (Paul and Das 2010; Paul *et al*. 2014ab; 2013). Arya *et al.* (2015) applied ARIMAX model for predicting pest population. In the present investigation, an attempt has been made to apply the combinations of wavelets with regression and ANN to forecast SMD and test for their accuracy of prediction in respect of selected Indian locations.

MATERIALS AND METHODS

Study area, surveillance and sampling plans

Locations involved in studies on pigeonpea SMD were from four different agro climatic zones and four-agro eco regions (Table 1) of India. Specific locations were part of a network implementing electronic pest and weather surveillance under National Innovations in Climate Resilient Agriculture (NICRA) project operated since 2011 until 2016 (Table 1).

Surveillance plan consisted of selection of 10 pigeonpea fields from 10 different villages located within a vicinity of 30-50 km radius of the research experimental station's meteorological observatory of each location. Surveillance was conducted during *kharif* seasons of 2011-16 with sampling for SMD done on weekly basis during each crop season. In each of the fields, SMD was examined at five different randomly selected spots with two plants/spot thus constituting 10 plants sampled per field. Mean incidence (%) of SMD was worked out for each of the weekly observation periods on standard meteorological week (SMW) basis in respect of each field with 10 cases (from 10 fields) contributing to data sets/week. Weather data on maximum and minimum temperature (Max T & Min T) (^{0}C), morning and evening humidity (RHM & RHE.) (%), sunshine hours (SS) (h/day), wind velocity (Wind) (km/h), total rainfall (RF) (mm), and rainy days (RD) on SMW basis also collected for study locations. The data entry, upload to centralised server and reporting were part of electronic pest surveillance with data base on pests including SMD maintained and accessed through web at : http://www.ncipm.res.in/nicra2015/NICRAPanel2012Onwards/rvLogin.aspx for study locations. For current investigation, data on SMD incidence (%) and of weather factors extracted on SMW basis for individual seasons of 2011-16 of each location was extracted and used.

Statistical analyses

Analysis of variance (ANOVA)

SMD incidence across seasons in respect of each location was analysed using one-way analysis of variance (ANOVA) after arcsin transformation with mean comparisons made through Duncan multiple range test (DMRT). Pearson correlation coefficients were worked out between SMD incidence and weather variables lagged by one and two weeks over all seasons (2011-16) for a given location. Analyses were carried out using *R* software (R Core Team 2013).

Support Vector Regression (SVR)

SVR formulated as an optimization problem by first defining a convex ε-insensitive loss function to be minimized and finding the flattest tube that contains most of the training instances (Awad and Khanna 2015) was used. A multi-objective function was constructed from the loss function and the geometrical properties of the tube. Then, the convex optimization that has a unique solution was solved using appropriate numerical optimization algorithms. The hyper plane represented the support vectors which are training samples that lie outside the boundary of the tube. The support vectors are the most influential instances that affect the shape of the tube, and the training and test data are assumed to be independent and identically distributed (iid), drawn from the same fixed but unknown probability distribution function in a supervised-learning context. The general form of nonlinear SVR estimating function is:

$$f(x) = w^{T}\varphi(x) + b, \tag{1}$$

where, $\varphi(.)$: $R^{n} \longrightarrow R^{nh}$ is a nonlinear mapping function from original input space into a higher dimensional, feature space, which can be infinitely dimensional, $w \in R^{nh}$ is weight vector, b is bias term and superscript T indicates transpose. The coefficients w and b are estimated from data by minimizing the following regularized risk function:

Table 1 Details of study locations

Location and state	Agro ecological region	Agro climate zone	GPS co-ordinates
SK Nagar (GJ)	Western plain, kachhh and part of Kathiawar peninsula,hot arid ecoregion	Gujarat Plains and Hills Region	24:19'N72:18'E
Gulbarga (KA)	Eastern ghat, TN upland and deccan plateau hot semi-arid ecoregion	Southern Plateau and Hills Region	17:21'N76:48'E
Rahuri (MH)	Deccan plateau Aravalli's hot semi-arid ecoregion	Western Plateau and Hills Region	19:21'N74:39'E
Vamban (TN)	Eastern ghat, TN upland and deccan plateau hot semi-arid ecoregion	East Coast Plains and Hills Region	10:21'N78:54'E

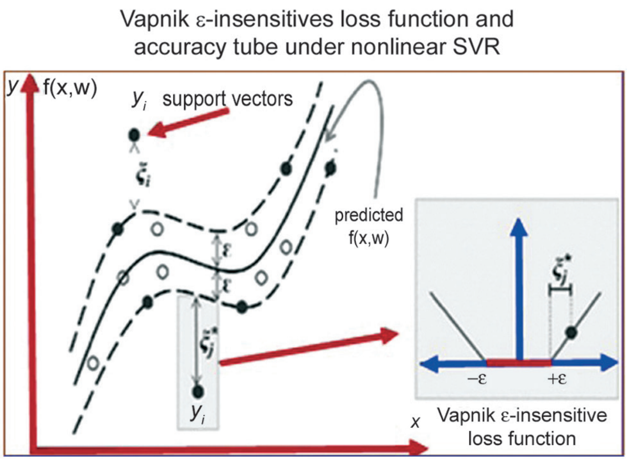

Fig 1 Schematic representation of nonlinear SVR model.

$$R(\theta)=\frac{1}{2}\|w\|^2+c\left[\frac{1}{N}\sum_{i=1}^{N}L\varepsilon\left(y_i,f(\mathrm{x}_i)\right)\right]. \tag{2}$$

This regularized risk function minimizes both empirical error and regularized term simultaneously and implements structural risk minimization principle to avoid under and over fitting of training data. Here, ½||w||^2 is called ‘regularized term’ and

$$\frac{1}{N}\sum_{i=1}^{N}L_\varepsilon\left(y_i,f(x_i)\right)$$

is called ‘empirical error’ estimated by Vapnik ε-insensitive loss function (Fig 1). In equation (2), C referred to as regularized constant determines trade-off between empirical error and regularized term. The value ε is called as tube size equivalent to approximation accuracy in training data. Both C and ε are user-determined hyper-parameters. Only those data points located on or outside the ε-tube are penalized and served as support vectors. Two positive slack variables ξ_i and ξ_i^* are introduced for representing the distance from actual values to corresponding boundary values of the ε-tube. A detailed description of above methodology can be found in Vapnik (2000).

Artificial neural network (ANN)

ANN, a powerful and self-adaptive approach for modeling nonlinear data was applied to datasets where the underlying data relationship was unknown. A general neural network consists of an input layer that accept external information, one or more hidden layers that provide non-linearity to the model and output layer that provides the target value. Each layer consists of one or more nodes. All the layers are connected through acyclic arc. Each input node in the input layer is associated with its corresponding weight. To compute the output, its activation function is applied to the weighted sum of the inputs. The activation function is either the identity function or sigmoidal function. Most commonly used ANN is multi-layer perceptron (MLP), a class of feed forward neural network. MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. An application of this approach can be found in Paul and Sinha (2016). A graphical presentation of MLP is given in Fig 2.

Hybrids of ARIMA-ANN and ARIMA-SVR

A hybrid approach proposed by Zhang (2003) that considered the time-series y_t as a combination of both linear and nonlinear components was also applied. That is,

$$y_t = L_t + N_t \tag{3}$$

where, L_t and N_t represent the linear and nonlinear part present in the given data respectively. Steps involved in hybrid method of combining forecasting are

(1) First, a linear time-series model, say, ARIMA is fitted to the data

(2) At the next step residuals are obtained from the fitted linear model. The residuals will contain only the non-linear components. Let e_t denotes the residual at the time t from the linear model, then

$$e_t = y_t - \hat{L}_t \tag{4}$$

where, $\hat{N}_t$ is the forecast value for the time t from the estimated linear model.

(3) Diagnosis of residuals is done to check if there is still linear correlation structures left in the residuals. The residuals are tested for nonlinearity by using BDS test developed by Brock, Dechert and Scheinkman.

(4) Once the residuals confirm the nonlinearity, then the residuals are modeled using machine learning tech-

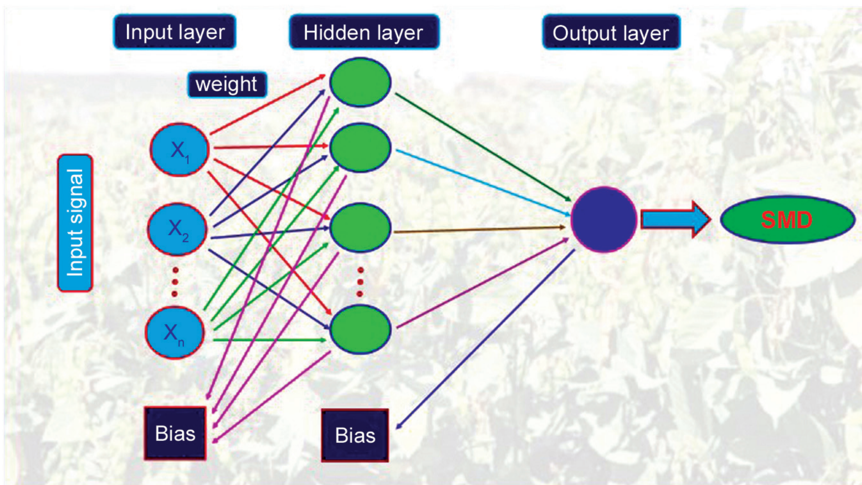

Fig 2 A multilayer perceptron (MLP) architecture with one hidden layer

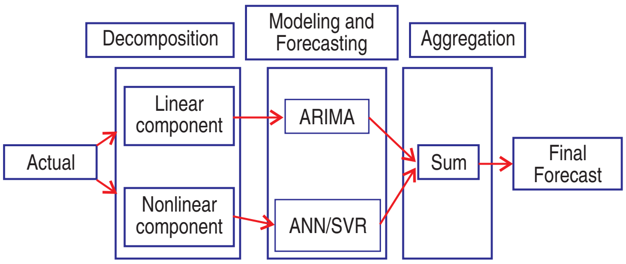

Fig 3 Schematic representation of ARIMA-ANN/SVR hybrid methodology.

niques, *viz.* ANN and SVR. The forecast values $\widehat{N}_t$ for the residual series are also obtained.

(5) Finally, the forecasted linear and nonlinear components are combined to obtain the aggregated forecast values as

$$\widehat{y}_t = \widehat{L}_t + \widehat{N}_t \quad (5)$$

An application of this methodology can be found in Mitra and Paul (2017). The schematic representation of the hybrid algorithm can be seen in Fig 3.

Validation of forecasts

While 90% of the datasets in each location on SMD and weather were used for model development, remaining 10% were utilized for validation. Comparative assessement of prediction performance of different models namely SVR, ANN, ARIMA-SVR and ARIMA-ANN models was carried out through statistical measure of root mean square error (RMSE) based on the following formluae:

$$\text{RMSE}=\sqrt{1/h\sum_{i=1}^{h}\left\{y_{t+i}-\widehat{y}_{t+i}\right\}^2}$$

where, h dentoes the number of observations for validation, y_i is the observed value and $\widehat{y}_i$ is the predicted one.

RESULTS AND DISCUSSION

Seasonal dynamics and status of SMD

Epidemics of diseases have increased in recent years due to climate change and there is a need to understand the impact of climate change on host pathogen interaction to outline appropriate management strategies (Chowdappa 2010). Studies on SMD in pigeonpea carried out for six consecutive *kharif* seasons (2011-16) at four locations namely S K Nagar (GJ), Gulbarga (KA), Rahuri (MH), and Vamban (TN) showed the commencement of infestation from second week of August with peak incidence between third week of October and November. SMD incidence was higher during 2011 at all locations with Gulbarga (5.30%) >S K Nagar (4.73%), Rahuri (2.08%) and Vamban (1.33%). Lowest SMD was during 2016 at both S K Nagar (0.03%) and Gulbarga (0.68%). Lowest incidence of SMD at Rahuri (0.11%) was in 2012.The seasonal variations in occurrence of SMD in pigeonpea across four studied locations is graphically represented in Fig 4.

Table 2 Comparative analysis of SMD occurrence across the years

Location	2011	2012	2013	2014	2015	2016
S K Nagar (GJ)	4.73^a	3.81^a	4.49^a	4.27^a	0.62^b	0.03^b
Gulbarga (KA)	5.30^a	1.06^b	1.19^b	4.90^a	1.34^b	0.68^b
Rahuri (MH)	2.08^a	0.11^b	0.68^b	0.39^b	-	-
Vamban (TN)	1.33^a	1.35^a	4.05^a	1.21^a	-	-

* Means followed by the superscript of same at p<0.05 based on DMRT

Comparisons of SMD incidence for levels across seasons for each location carried out using DMRT are presented in Table 2. Both S K Nagar (GJ) and Gulbarga (KA) had mean incidence significantly lower in 2016 as compared to 2011 that was on par with other seasons. Significantly, higher incidence in 2011 and lower on par incidence during three consecutive seasons (2012-14) was noted at Rahuri (MH). SMD incidence did not vary across seasons at Vamban (TN). Kannaiyan *et al*. (1984) reported higher incidences over the present study at Tamil Nadu (12.8%), Gurajat (12.2%) and Karnataka (9.8%) indicating the overall reduction of SMD in recent periods.

Association of SMD incidence with weather

The relationships between population size/pest incidence and climatic variables are analyzed based on simple correlations or using the weather variables as an additive covariate in statistical models (Stenseth *et al.* 2002). Current study used, Pearson's correlation analysis to find out the significant weather variables influencing the occurrence of SMD in pigeonpea (Table 4). Correlation analyses of SMD were performed with weather parameters lagged by one and two weeks considering the the periods before symptom manifestation. Results indicated significant and positive influence of maximum temperature of two lags and sunshine hour of one and two lags on SMD at S K Nagar (GJ). Significant and positive influence of maximum temperature of one lag week at Gulbarga was observed indicating the spatial variations of the effects of weather. Significant and positive influence of minimum temperature, morning and evening humidity of one and two lag weeks and wind speed at two lag weeks was noted for Rahuri (MH). At Vamban, positive influence of morning humidity and sunshine and negative influence of rainfall and rainy day of one and two lag weeks were significant on SMD. Not many studies are

Table 3 Comparative analysis of SMD occurrence across the locations

	S K Nagar (GJ)	Gulbarga (KA)	Rahuri (MH)	Vamban (TN)
SMD	2.92^a	2.34^a	0.60^c	1.37^b
N	90	161	83	139

* Means followed by the superscript of same at p<0.05 based on DMRT

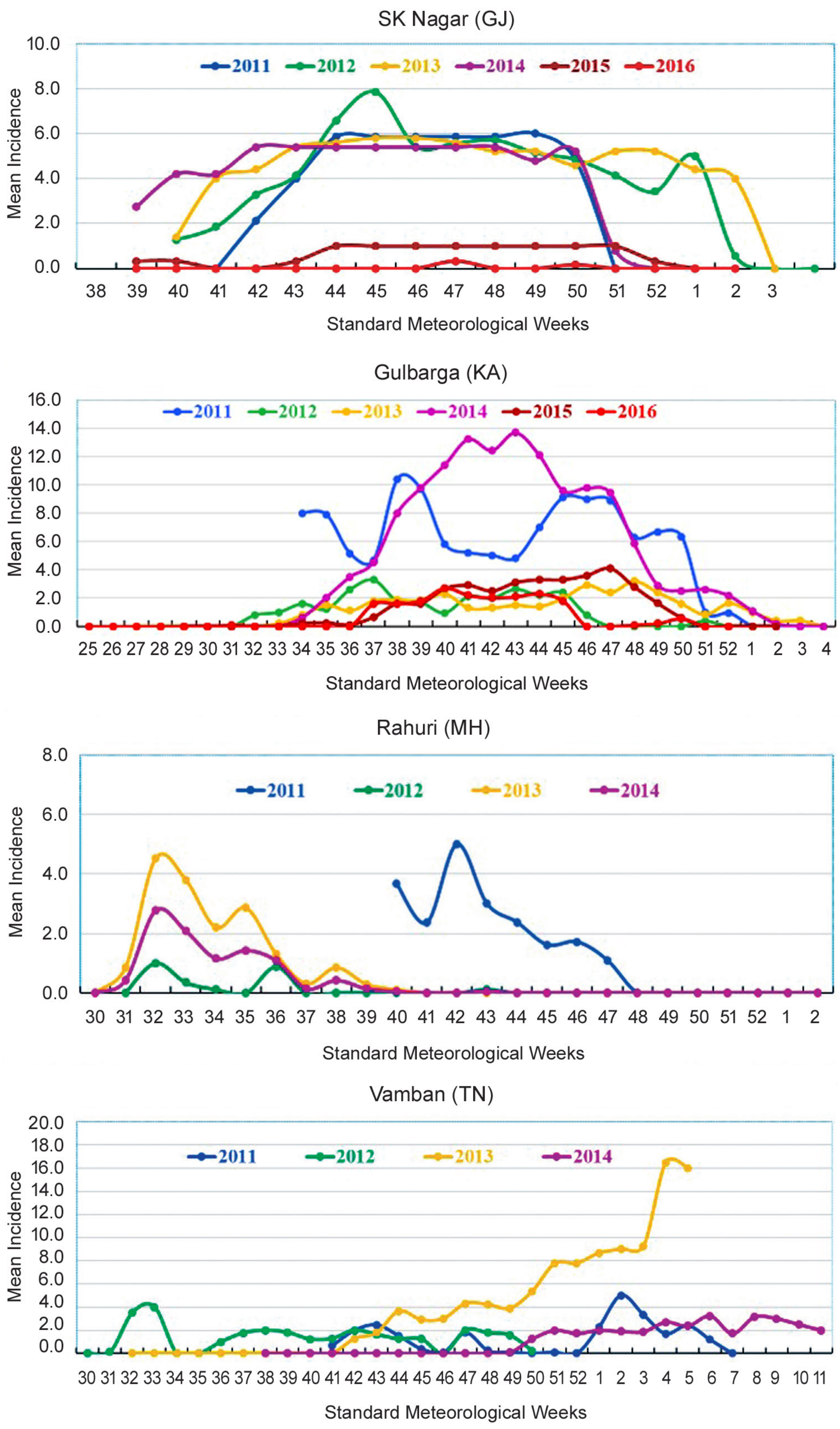

Fig 4 Seasonal variations in occurrence of SMD across studied locations

available on the effect of weather on SMD. Abhijit *et al.* (2014) reported positive effect of lower temperature and rainfall positively influencing manifestation of SMD which was obvious across different locations. On application of the multiple regression model for prediction of incidence of SMD in different locations, the R^2 values are obtained as 0.55, 0.35, 0.27 and 0.29 in SK Nagar (Gujarat), Gulbarga (Karnataka), Rahuri (Maharashtra) and Vamban (Tamil Nadu) respectively. It clearly indicates that models are not good fit for the data under consideration.

Prediction of SMD incidence through conventional and Hybrid models

Regression and neural network methods have become competing model-building methods in the recent days. ANN and SVR are considered to be the best techniques for extracting information from imprecise and non-linear data (Caselli *et al.* 2009). SVR and ANN being nonparametric in nature could model the non-normal variables more precisely and could capture the nonlinearity present in dataset. Cheng and Titterington (1994) carried out complete analysis and comparison of different network techniques with traditional statistical techniques. Li *et al.* (2007) carried out a study to develop a new methodology using an ANN to estimate and predict corn and soybean yields on a county-by-county basis, in the "corn belt" area in the Midwestern and Great Plains regions of the United States. Mitra and Paul (2017) analyzed agricultural commodity prices using hybrid models. Paul *et al.* (2018) applied SVR technique for prediction of SMD of pigeonpea for Banaskantha region of Gujarat and found it to be better over ANN and ARIMAX. When such approaches were applied across locations, the results of (Table 5) indicated that performance

Table 4 Correlation coefficients between SMD with weather factors lagged by one and two weeks#:

	S K Nagar (GJ)	Gulbarga (KA)	Rahuri (MH)	Vamban (TN)
$MaxT_{-1}$	0.11	0.16*	0.11	-0.03
$MinT_{-1}$	-0.07	0.00	0.34**	-0.12
RHM_{-1}	-0.24*	-0.03	0.29*	0.32**
RHE_{-1}	-0.51**	0.03	0.26*	-0.06
RF_{-1}	0.03	-0.03	-0.11	-0.20*
SS_{-1}	0.22*	-	-0.27*	0.22*
$Wind_{-1}$	0.09	-0.12	0.21	0.01
RD_{-1}	-0.03	-0.04	-0.01	-0.23*
$MaxT_{-2}$	0.21*	0.15	-0.09	-0.01
$MinT_{-2}$	-0.03	0.02	0.34**	-0.11
RHM_{-2}	-0.22*	-0.03	0.32**	0.37**
RHE_{-2}	-0.43**	0.05	0.32**	-0.05
RF_{-2}	0.00	-0.03	-0.03	-0.20*
SS_{-2}	0.30**	-	-0.34**	0.17*
$Wind_{-2}$	-0.08	-0.11	0.25*	-0.07
RD_{-2}	-0.09	-0.04	0.07	-0.24*

#The suffix 1 and 2 denotes the lag in weeks of weather relating to disease severity considered for correlations. ** significant at $p< 0.01$; * significant at $p< 0.05$

varied across locations. Residual diagnostics carried out to check the adequacy of fitted model revealed absence of autocorrelations at all locations. The hybrid model of ARIMA-SVR performed better than the other three models in all other locations but for Rahuri. In Rahuri, SVR model was the best in predicting the occurrence of SMD.

The order of performance of models predicting SMD at locations other than Rahuri (MH) was: ARIMA&SVR> ARIMA&ANN> SVR> ANN. Hybrid models proving to be better in forecasting SMD at many locations over individual model approaches ranging from multiple regression, ARIMA, ANN, SVR and ARIMAX. SVR model predictions was noted. The very low levels of SMD incidence between 2012 and 2014 could be one of the reasons for the under performance of hybrid models. It could also be said that hybrid models would perform better when mean seasonal incidence of SMD is greater than 1%, although further confirmations are advocated. SVR model performing better over other conventional models, *viz.* MLR, ARIMA, ARIMAX and ANN have already been documented for SMD using datasets of S K Nagar (Banaskantha, Gujarat) (Paul *et al.* 2018). Present investigation added additional dimension of application of hybrids to increase SMD predictions.

Conclusion

Adverse effects of climate change in pulse growing arise largely due to wider fluctuations in temperature and aberrant rainfall patterns. Assessment of seasonal incidence of diseases in relation to weather variations over

Table 5 RMSE values of SVR, ANN and hybrids with ARIMA predicting SMD incidence

Location	SVR	ANN	Hybrid ARIMA and ANN	Hybrid ARIMA and SVR
S K Nagar (GJ)	1.8	1.96	0.75	0.26
Gulbarga (KA)	1.09	1.66	0.78	0.74
Rahuri (MH)	0.12	0.22	0.34	0.35
Vamban (TN)	0.54	1.02	0.83	0.48

a considerable number of seasons is of significance to understand the effect of changing climate. Present study revealed declined mean incidence of SMD on pigeonpea across all study locations. Significant influence of a single variable of MaxT lagged by a week at Gulbarga (KA) to many variables (RHM, RHE and SS) lagged by one as well as two weeks at S K Nagar (GJ) and Rahuri (MH) were noted with nature of significant associations differing spatially. Vamban (TN) the humid eco-region over other locations had only the negative relationship of rainfall with SMD significant. Irrespective of the differential influences of weather variables, statistical approaches brought out the overall better performance of hybrid models of SVR and ANN with ARIMA over the individual models when mean seasonal incidence worked out to be greater than one percent. Ability of the nonlinear models to predict SMD better when combined with linear ARIMA offer scope of higher prediction accuracies and hence improved strategies for disease management. The approaches implemented for forecast of SMD could be verified with additional diseases and insects such that field of pest forewarning could be a robust component of integrated pest management in the days to come.

ACKNOWLEDGEMENTS

The authors are thankful to Central Research Institute for Dryland Agriculture, Hyderabad (CRIDA), India for the financial support to undertake this study through National Innovations in Climate Resilient Agriculture (NICRA).

REFERENCES

Abhijit Daspute, Fakrudin B, Shivarudrappa, Bhairappanavar B, Kavil S P, Narayana Y D, Muniswamy, Anil Kaumar, Krishnaraj P U, Abid Yerimani, and Khadi B M. 2014. Inheritance of pigeonpea sterility mosaic disease resistance in pigeonpea. *Plant Pathology Journal* **30**(2): 188–194.

Arya P, Paul R K, Kumar A, Singh K N, Sivaramne N and Chaudhary P. 2015. Predicting pest population using weather variables: an ARIMAX time series framework. *International Journal of Agriculture and Statistical Sciences* **11**: 381-386.

Awad M and Khanna R. 2015. Support Vector Regression. (In) *Efficient Learning Machines*. Apress, Berkeley, CA.

Box G and Jenkins G. 1970. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA.

Caselli M, Trizio L, Gennaro G D and Ielpo P. 2009. A simple feed forward neural network for the PM10 forecasting: comparison with a radial basis function network and a multivariate linear regression model, *Water Air Soil Pollution*. **201**: 365–377.

Cheng B and Titterington D M. 1994. Neural networks: A review from statistical perspective. *Statistical Science* **9**: 2-54.

Chowdappa P.2010. Impact of climate change on fungal diseases of Horticultural crops. (*In) Challenges of Climate Change-Indian Horticulture,* pp 144-151. H P Singh, J P Singh and S S Lal (Eds). Westville publishing house, New Delhi.

Jones A T, Kumar P L, Saxena K B, Kulkarni N K, Muniyappa V and Waliyar F. 2004. Sterility mosaic disease-the '*green plague'* of pigeonpea: advances in understanding the etiology, transmission and control of a major virus disease. *Plant Disease* **88**: 436–445.

Kaushik D, Srivastava S, Singh J P, Chauhan V B and Singh R N. 2013. Evaluation of pigeonpea genotypes against Sterility Mosaic Disease (SMD) caused by the mite vector, *Aceria cajani. Journal of Mycopathological Research* **51**: 225-233.

Kumar P L, Jones A T and Waliyar F. 2008. Virus diseases of pigeonpea.(*In*) *Characterization, Diagnosis and Management of Plant Viruses. Vegetable and Pulse Crops,* 3, pp 235-258. Rao G P, Kumar P L and Holguin-Pena R J (Eds). Texas, USA: Studium Press.

Lakshmikantha B P and Prabhuswamy H P. 2002. Reaction of pigeon pea varieties to mite (*Aceria cajani* Channabasavanna) transmitting sterility mosaic disease. *Plant Protection Bulletin* **54**: 6-7.

Li A, Liang S, Wang A and Qin J. 2007. Estimating crop yield from multi-temporal satellite data using multivariate regression and neural network techniques. *Photogrammetric Engineering & Remote Sensing* **73**: 1149–1157.

Mitra D and Paul R K. 2017. Hybrid time-series models for forecasting agricultural commodity prices. *Model Assisted Statistics and Applications* **12**: 255–264.

Nene Y L, Kannaiyan J and Reddy M V. 1981. Pigeonpea diseases: resistance-screening techniques. Information Bull. No.9. Patancheru, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics, p 19.

Paul R K and Das M K. 2010. Statistical modelling of inland fish production in India. *Journal of Inland Fisheries Society of India* **42**: 1-7.

Paul R K, Alam W and Paul A K. 2014a. Prospects of livestock and dairy production in India under time series framework. *Indian Journal of Animal Sciences* **84**:130-134.

Paul R K, Ghosh H and Prajneshu. 2014b. Development of out-of-sample forecast formulae for ARIMAX-GARCH model and their application. *Journal of the Indian Society of Agricultural Statistics* **68**: 85-92.

Paul R K, Prajneshu and Ghosh H. 2013. Statistical modelling for forecasting of wheat yield based on weather variables. *Indian Journal of Agricultural Sciences* **83**:180-183.

Paul R K, Vennila S, Singh N, Chandra P, Yadav S K, Sharma O P, Sharma V K, Nisar S, Bhat M N, Rao, M S and Prabhakar M. 2018. Seasonal dynamics of sterility mosaic of pigeonpea and its prediction using statistical models for Banaskantha region of Gujarat, India. *Journal of the Indian Society of Agricultural Statistics* **72**: 213-223.

Paul R K and Sinha K. 2016. Forecasting crop yield:a comparative assessment of ARIMAX and NARX model. *RASHI* **1**: 77-85.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reddy M V, Rayu T N and lanne J M. 1998. Diseases of Pigeonpea. (*In*) The Pathology of Food Pasture Leguenes, pp 517-58. Allen D J. and lenne J M (eds). Wallingford oxfordshire: CAB International.

Stenseth N C, Mysterud A, Ottersen G, Hurrell J W, Chan K S and Lima M. 2002. Ecological effects of climate fluctuations. *Science* **297**: 1292-1296.

Van Der Maesen L J G. 1990. Pigeonpea: Origin, history, evolution and taxonomy. (*In*) *The Pigeonpea*, pp 15-46. Nene Y L, Hall S D, and Sheila V K (Eds). CAB International, Wallingford, UK.

Vapnik V N. 2000. *The Nature of Statistical Learning Theory.* Springer-Verlag, New York.

Zhang G, Patuwo B and Hu M Y. 1998. Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting* **14**:35-62.

Zhang G P. 2013. Time series for coming using a hybrid ARIMA and neural network model. Neuro computing 50: 159-175.