International Journal of Information Technology and Knowledge Management
July-December 2008, Volume 1, No. 2, pp. 303-309

A FRAMEWORK TO PREDICT SUITABLE PERIOD FOR MUSTARD PLANT CONSIDERING EFFECT OF WEATHER PARAMETERS USING FACTOR AND PRINCIPAL COMPONENT ANALYSIS

Satyendra Nath Mandal, J. Pal Choudhury,
S. R. Bhadra Chaudhuri & Dilip De

Abstract

The plants of different corn are planted in different month within a year. Some are suitable in summer; some are winter and some rainy season. One common question is that can plants of all corn are planted throughout the year? The growth of plant is dependent on environment parameters like nature of soil, temperature (maximum and minimum), humidity (maximum and minimum), rainfall, soil moisture and sun rays etc. The nature of soil is overall West Bengal, India, morel less same. But other parameters vary from place to place. Using fuzzy logic, artificial neural network and genetic algorithms has been already tested to predict the growth of shoot length without considering the effect of environmental parameters [1]. In this paper, an effort has been made to predict the suitable time period within a year for mustard plant by considering the total effect of maximum and minimum temperature, rainfall, maximum and minimum humidity, soil moisture, and sun shine using the method of factor analysis and principal component analysis for estimation of growth of shoot length.

Keywords: Environment parameters, Plant parameter, Factor Analysis, Principal Component Analysis, Total Effect.

1. Introduction

The growth of mustard plant without considering environmental parameters is predicted by Satyendra Nath Mandal, J. Pal Chaudhury, S.R Bhadra Chaudhuri and Dilip De [1] using fuzzy logic and neural network. The goal in Principal Component's analysis ([2],[9]-[11]). is to construct linear combinations of the original variables that account for a large part of the total variation. In this paper, the effect of all environmental parameters has also been considered using the method of factor analysis and principal component analysis for estimation of growth of shoot length. The effect will predict suitable period of planting of muster plant. Research have been done ([2-[10]) in different field using Principal Component Analysis and Factor Analysis to solve different types of problem. The work related to the factor and principal component analysis has not been done in agriculture field so far, that is the reason for making this effort in this paper.

2. METHODOLOGY

2.1. Factor Analysis

In order to factor *S*, the special decomposition is *S* = CDC' where *C* is an orthogonal matrix constructed with normalized vectors ($c'c = 1$) of *S* as columns and *D* is a diagonal matrix with eigen values $\theta_1, \theta_2, \ldots \theta_p$ of S on the diagonal:

$$D = \begin{matrix} \sqrt{\theta}_1 & 0 & 0 \ldots 0 \\ 0 & \sqrt{\theta}_2 & \ldots 0 \\ & . & \\ 0 & 0 \ldots & \sqrt{\theta}_p \end{matrix}$$

The notation θ_i is used for eigenvalues.

2.2. Principal Component Analysis

PCA components analysis transforms the original set of variables into a smaller set of linear combination that account for most of the variance of the original set. The principal components are extracted so that the first principal component, denoted here by $PC_{(1)}$, accounts for the largest amount of the total variation in the data. It is possible to extract as many principal components as there are original variables; however, the goal in most principal components applications is to account for most of the total variation with as few principal components as possible.

2.3. Data Used in this Paper (Time Instances vs. Shoot Length of Different Mustard Plant and Environmental Parameters)

A Statistical survey has been conducted by a group of certain agricultural scientists on different mustard plants under the supervision of Prof. Dilip De, Bidhan Chandra Krishi Viswavidayalay West Bengal, India. In this paper, the value of shoot length measured at initial stage (after an interval of 7 days, 14 days, 21 days, 28 days) after plantation and environmental data from this survey has been taken (between November to February) as input data of a particular type of mustard plant furnished in table 1 and 2. The objective

Table 1
Time Instances versus Shoot Length of Mustard Plant

Time Instances	*Shoot Length*
1	17
2	22
3	27
4	31

is finding out the effect the environmental parameters in prediction of time period of mustard plant.

Table 2
Data Related Environmental Parameter

Maxi Temp.	*Min Temp.*	*Rain Fall*	*Max. Humidity*	*Min Humidity*	*D1*	*D2*	*D2*	*Sun/ Shine*
27.4	15.6	28.8	95.75	58.35	16.9	21.96	28.13	7.77
25.9	14.15	19.2	96.68	59.48	14.83	19.63	26.1	8.11
24.4	12.7	9.6	97.6	60.6	13.23	17.7	23.76	8.32
24	12.33	29.68	97.78	60.73	11.86	16.33	21.63	7.44
23.6	11.95	49.75	97.95	60.85	18.86	14.9	19.76	7.18
23.2	11.58	69.83	98.13	60.98	11.03	15.03	19.3	7.62
22.8	11.2	89.9	98.3	61.1	10.76	14.93	19.13	8.48
24.08	12.38	74.73	97.93	59	10.8	14.83	18.96	7.32
25.35	13.55	64.55	97.55	56.9	9.8	13.8	17.96	7.14
26.63	14.73	51.88	97.18	54.8	9.3	13.1	17.03	6.05
27.9	15.9	39.2	96.8	52.7	8.13	12.36	16.03	7.48
28.53	16.38	79.3	96.43	51.83	7.86	11.86	15.46	9.42
29.15	16.85	119.4	96.05	50.95	6.96	10.13	14.33	7.22

3. Implementation

3.1. Factor Analysis

Step 1: After the plantation the values of environmental parameters have been collected for the entire period (weekly) of growing stage of the plant i.e. up to harvest period (for approximately 90 to 95 days) furnished in table 2. Now it is necessary to see the total effect from these weather parameters affect on the growth of shoot length. The correlation matrix ([10]) has been calculated of all the elements from the table 2.

Step 2: The eigen values and eigen vectors have been calculated using Mathlab using correlation matrix. The eigen values are 0.0000 0.0000 0.0023 0.0106 0.2363 0.4298 1.0022 2.5762 4.7426 respectively. The eigen vectors corresponding eigen values are as follows:

0.07	0.84	0.10	0.04	-0.06	-0.00	0.04	-0.34	-0.38
0.75	-0.38	-0.0	-0.09	-0.05	-0.01	0.05	-0.36	-0.37
0.08	0.03	0.10	-0.09	0.25	0.82	-0.18	0.33	-0.27
0.44	0.24	0.40	-0.43	-0.09	-0.24	-0.01	0.50	0.25

0.46	0.24	-0.42	0.54	0.13	0.10	-0.04	0.11	0.44
0.00	0.00	0.05	-0.09	-0.75	0.45	0.13	-0.23	0.36
-0.02	0.11	-0.34	-0.65	0.40	0.10	-0.02	-0.37	0.35
0.02	-0.08	0.71	0.25	0.36	0.10	-0.01	-0.39	0.34
0.00	-0.00	0.00	-0.00	-0.18	-0.10	-0.96	-0.13	0.01

Step 3: the total effect has been calculated using cumulative effect furnished in table 4

Table 4
Total Effect Value Incorporating All Items

Maxi Temp.	*Min Temp.*	*Rain Fall*	*Max. Humidity*	*Min Humidity*	*D1*	*D2*	*D2*	*Sun/ Shine*	*Total Effect*
27.4	15.6	28.8	95.75	58.35	16.9	21.96	28.13	7.77	182.40
25.9	14.15	19.2	96.68	59.48	14.83	19.63	26.1	8.11	182.90
24.4	12.7	9.6	97.6	60.6	13.23	17.7	23.76	8.32	178.92
24	12.33	29.68	97.78	60.73	11.86	16.33	21.63	7.44	179.13
23.6	11.95	49.75	97.95	60.85	18.86	14.9	19.76	7.18	180.81
23.2	11.58	69.83	98.13	60.98	11.03	15.03	19.3	7.62	181.63
22.8	11.2	89.9	98.3	61.1	10.76	14.93	19.13	8.48	184.18
24.08	12.38	74.73	97.93	59	10.8	14.83	18.96	7.32	181.68
25.35	13.55	64.55	97.55	56.9	9.8	13.8	17.96	7.14	177.95
26.63	14.73	51.88	97.18	54.8	9.3	13.1	17.03	6.05	174.47
27.9	15.9	39.2	96.8	52.7	8.13	12.36	16.03	7.48	168.87
28.53	16.38	79.3	96.43	51.83	7.86	11.86	15.46	9.42	172.46
29.15	16.85	119.4	96.05	50.95	6.96	10.13	14.33	7.22	178.75

Step 4: The Growth of shoot length has been measured weekly for four weeks. The weather records have also been taken during that time. That is the reason; table 5 contains only four rows for total effect and shoot length.

Table 5
Growth of Shoot Length Corresponding to Total Effect Value

Total Effect Value	*Growth in Shoot Length*
182.40	17
182.90	36
178.92	54
179.13	67

Step 5: The value of total effect corresponding to the weather parameters related to any period outside the range of time selected for plantation of mustard seed is furnished table 6.

Table 6
Total Effect Value for Different Time Excluding the Planting and Growing Time

Date	*Maxi Temp.*	*Min Temp.*	*Rain Fall*	*Max. Humidity*	*Min Humidity*	*D1*	*D2*	*D3*	*Sun/ Shine*	*Total Effect*
April 14,	34.03	23.22	79.9	94.4	62	15.9	17.96	20.13	9.77	218.72
May 6,	35.6	25.1	61.5	89.96	62.4	10.83	15.63	20.1	8.11	220.86
June 5	33.68	24.4	18.47	96.57	78.33	11.23	15.7	21.76	8.11	246.46
July 2	32.1	25.5	10.17	97.42	78.85	10.86	13.33	19.63	6.44	247.97
Aug 5	32.84	26.1	16.02	97.42	76.71	11.86	14.9	18.76	7.18	246.55
Sep 6	32.7	25.5	274	97.56	79.82	10.03	13.03	17.3	6.62	226.11
Oct 5	33.41	24.67	17.77	99	68	8.76	12.93	17.13	7.48	237.92

3.2. Principal Component Analysis

The weather parameters (maximum and minimum temperature, rainfall and maximum and minimum humidity, soil moisture in different depth(D1,D2,D3) and sun shine of table 2 have been used. The covariance matrix[10] has been calculated of all the elements(maximum and minimum temperature, rainfall and maximum and minimum humidity, soil moisture in different depth and sun shine) of the table 2. The eigen values and eigen vectors have been calculated using Matlab. The contribution of eigen value among all the eigen values also have been calculated. Using all the eigen values belong to each component (maximum temperature, minimum temperature, rainfall, maximum humidity, minimum humidity, soil moisture and sun shine), total effect has been calculated. The growth in shoot length corresponds to total effect value are furnished in table 7.

Table 7
Growth of Shoot Length Corresponding to Total Effect Value

Total Effect Value	*Growth in Shoot Length*
2635.31	17
2748.54	36
2839.64	54
2814.77	67

In Table 8, furnished total effect of weather parameters relating to other period.

Table 8
Total Effect Value for Different Time Excluding the Planting and Growing Time

Date	*Max Temp.*	*Min Temp.*	*Rain Fall*	*Max. Humidity*	*Min Humidity*	*D1*	*D2*	*D3*	*Sun/ Shine*	*Total Effect*
April	34.03	23.22	79.9	94.4	62	15.9	17.96	20.13	9.77	2980.45
May	35.6	25.1	61.5	89.96	62.4	10.83	15.63	20.1	8.11	2989.67
June	33.68	24.4	18.47	96.57	78.33	11.23	15.7	21.76	8.11	2427.42
July	32.1	25.5	10.17	97.42	78.85	10.86	13.33	19.63	6.44	2200.87
Aug	32.84	26.1	16.02	97.42	76.71	11.86	14.9	18.76	7.18	1968.55
Sep	32.7	25.5	274	97.56	79.82	10.03	13.03	17.3	6.62	1888.20
Oct	33.41	24.67	17.77	99	68	8.76	12.93	17.13	7.48	1789.67

4. Results

It has also been observed that using factor analysis the value of total effect that has been calculated during planting period table 6 are not tally with any value of total effect of table 5 which contains the data related to the plantation period of mustard (from third week of November up to end of February of next year). In the next section, the value of total effect in table 8 does not tally with any value of total effect of table 7 (which contains the data related to the plantation period of mustard i.e. from third week of November up to end of February of next year). Now conclusion can be drawn that mustard can be planted during the end of November. If the mustard be planted during other time of year, it can not be grown and proper harvest can not be executed which tallies with real situation.

5. Conclusions and Future Work

The effect of weather parameters on the growth of shoot length has also been examined using factor analysis and principal component analysis. It has been observed that that total effect value combining all the weather parameters is not same range in the time excluding time of plantation, and growing time of mustard plant as compared to the plantation and growing time of mustard plant. The total effect value during month of mid November up to February may be suitable for growth of shoot length. That is the reason the mustard is being planted during the month of November of any year and other time of the year is not suitable for growth of mustard plant.

Fuzzy–PCA and Neuro-PCA will be used in future to predict the shoot length at maturity considering all plant parameters and environmental parameters. Finally, a particular mustard plant will be predicted based on its maximum productivity at mature stage.

References

[1] *Satyendra Nath Mandal, J. Pal Choudhury, Dilip De and S. R. Bhadrachaudhuri* , " A Framework for Development of Suitable Method to find Shoot Length at Maturity of Mustard Plant Using Soft Computing Model", *International journal of Computer Science and Engineering* **2**, (3)

[2] Jon Shlens, "A Tutorial on Principal Component Analysis Derivation, Discussion and Singular Value Decomposition", *http://www.jonshlens@ucsd.edu*, **1**, , (25 March 2003) 1–16.

[3] Changwon Suh, Arun Rajasgopalan, Xiang Li and Krishna Rajan, "The Application of Principal Component Analysis to Materials Science Data", 19–26

[4] Tutorial, "Principal Components and Factor Analysis", *http://www.startsoft.com/textbook/stfacan.html*

[5] M. Iqbal Saripan, M. Haikal Saripan, Emilia Udin, "Database Compression Using Principal Component Analysis Method", *2nd World Engineering Congress*, Sarawak Malaysis, 22–25, (July 2002), 500–512.

[6] Tapani Raiko, A Laximander llin and Juha Karhumen, "Principal Component Analysis for Large Scale Problems with Lots of Missing Values", ECML 2007, LNAI 4701, 691–669,@ Springer – Verlag Berlin, Heidelberg (2007).

[7] Patrik O Hoyer, Aapo Hyviarinen, "Independent Component Anslysis Applied to Feature Extraction from Colour and Stereo Images Network. Computation in Neural System", 191–210 (August 2000) 1–11.

[8] Neil Lawrence, "Probabilistic Non-linear Principal Component Analysis with Gaussian Latent Variable Methods", *Journal of Machine Learning Research* **6**, (2005), 1783–1816.

[9] Geosoft Technical Note on "Principal Component Analysis (PCA) and Factor Analysis in Oasis montaj", *www.geosoft.com*

[10] Alvin C Rencher & Willey Intersciences, "Methods of Multivariate Analysis", John Wiley & Sons Inc Publication, USA.

[11] Lindsay I Smith, "A Tutorial on Principal Components Analysis", February 26, 2002.

Satyendra Nath Mandal & J. Pal Choudhury
Dept. of I.T
Kalyani Govt. Engg College, Kalyani
Nadia (W.B.), India
E-mail: satyen_kgec@rediffmail.com; jnpc193@yahoo.com

S. R. Bhadra Chaudhuri
Dept. of ETE
Bengal Engineering & Science University
Howrah (W.B), India
E-mail: prof_srbc@yahoo.com

Dilip De
Bidhan Chandra Krishi Viswavidayalay
Kalyani
Nadia (W.B), India
E-mail: dilipde.bckv@gmail.com