International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Extemporize Agriculture Yield with Predictions Based on Water and Soil Properties using Multivariate Analytics and Machine Learning Algorithm

V. Sudha, S. Mohan

Abstract – ***Big data have rapidly developed in agriculture that increase the extensive attention of farmers to get extraordinary ideas for future crop development based on soil and water is an essential key factor of agriculture to predict data. In future farmers apply the extracts PCA (Principal component analysis) in the agriculture for crops to find best yield. The level of data is analyzed using PCA and PLS (Projection to latent structure) datasets like Crop data, Soil properties and Water properties such as Linear Regression, Multi Linear Regression, Discriminate Analysis, Partial Least Square, Least squares algorithm is used instead of Multivariate Curve Resolution. Predictive analytics can be used to make a modular decision in farming by observation of actual time data on crops, soil and water data received from different agriculture sources would contain multi-dimensions, the entire content is needed for performing analysis. Multivariate data Analysis based on Partial Least Square Programming model that identifies the cropping pattern to getting maximum yield correlated. The motive of the work is to compare various techniques which give the maximum accuracy of crop.***

Keywords: PCA; PSL; Support Vector Machine; Multivariate; Correlate; Data Analytics.

I. INTRODUCTION

The prediction about the yield is very important to the farmers to get highest Productivity. The prediction of yield can be extracted from the PCA (Principal Component Analysis) and PLS (Partial Least Squares) through our researchers the sharp observation of data analytics that has to be increased the yield so that we have a complicate problem to the big data for better observation. The major plan of principal component analysis (PCA) is to reduce the originality of a dataset having many variables correlated with each other, either heavily or slightly, from the remaining variation provide in the dataset, up to the maximum extent. Projection to Latent Structures regression observious and result may be useful to first review this information in Elementary concepts. The varieties of samples analyzed in this title are also characterized in common linear sample, Generalized Linear Models, and General systematic Regression. In Cuddalore district agriculture maintains to be the top most group in the economic growth of Crop Paddy, Cumbu, Ragi, Cholam, Redgram, Gingerly, Groundnut, Greengram, Cashew nuts, Sugarcane, Coconut.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

V. Sudha*, Research Scholar, Department of CIS, Annamalai University.

Dr. S. Mohan, Assistant Professor, Department of Computer Science and Engineering, Annamalai University.

II. RELATED WORK

Sirisha Adamala presented and implemented in this study papers with big data applications and water Resources engineering by using machine learning, research [1] the water managers, is showing big promise in many water related application such as planning optimum water systems, detecting ecosystem changes through big remote sensing and geographical information system, forecasting, predicting, detecting natural and man-made calamities, scheduling irrigations, mitigating environmental pollution, analyzing atmosphere alterations of impacts etc.

S. P. P. Kolhe, [2] Information technology is mostly used in modern agricultural irrigation systems to improve the food production. In this paper, we present the material for various purposes about modern irrigation system with the use of information technology in irrigation system at irrigation sector of agriculture, such as weather parameter, water availability, and water requirements of crops to create water supply schedules. Management of water availability in irrigation scheduling, crop management automated irrigation.

S. Kumawat, Sensor Based Automatic Irrigation System and Soil pH Detection using an Image Processing Development of agriculture is necessary [3]. Lack of rain and water scarcity is the major problems that are faced by the farmers. The main objective of this paper is providing an automatic irrigation system, thereby saving the time, money & power of the farmer. The traditional farmland irrigation techniques require manual intervention can be minimized. Nutrient availability and plant growth are affected by the acidity and basicity of the soil, which is determined by soil pH property. Haryana, Evaluation of ground water quality for irrigation [4] the quality of groundwater, water samples was collected from all the villages. Water sample were collected and analyzed for pH, EC, soluble cations (CA, mg, Noah and k) and anions.The values of sodium adsorption ration (SAR) and residual sodium carbonate (RSC). Spatial maps of EC, pH, RSC, SAR, Water quality of groundwater, Following details are used for irrigation.

The analysis of various parameters indicated that quality of groundwater is not safe for irrigation purpose.

The techniques of data mining are [5] Though data mining techniques in agriculture is extremely popular, its application in agricultural soil knowledge sets field is very young to wear data processing system product and domain Specific data processing application software are Obtainable however; data processing in agricultural soil Knowledge set may be a comparative a young analysis field. The information from the dataset is extracted and transformed to an understandable structure, which is the ultimate goal of data-mining. This paper offers internet based answer for soil testing laboratories yet as free messages to farmers that contains soil testing code important chemicals required for the crop.

The standard statistical analysis [6] they are time consuming and expensive. Agriculture research has been profited by technical advance such as automated data mining. Data mining can be used for solving complex data, which in turn can be used for improving accuracy and effectiveness in classification of large soil. Data mining process involves extracting information from dataset and transforming into an understandable structure that can be used for further studies. The soil test is to analyze the nutrient content, composition and other characteristics. Soil classification deals with the categorization of soil into different soil classes as very Low, low, medium, high and very high of nutrient found in the soil and the basis of these classes fertilizers are recommended for a soil sample.

Soil samples have been collected from [7] the samples of the locations spread uniformly over the study area in order to examine the infiltration rate of soils and its impacts on the overall crop production process. Double ring infiltration was used to carry out the experimental study. Infiltration rates were taken at 0 to 70 minutes of 10 minute intervals. The assessment of infiltration rate was processed by laboratory analysis of soil samples for the particle size distribution the recommended values for crop production. This will assists to develop the structure and maintain the soil potentials.

The paper by the author Abdul-Rahman, S., & Mohamed, A. (2010) Explains "Support Vector Machine" [8] based on the classification of the soil types. Soil Classification includes steps like image acquisition, image preprocessing, and feature extraction and classification. The texture features soil image is extracted using the low pass filter, Gabor filter and using the color quantization technique. Mean amplitude, Standard deviations is taken as the statistical parameter. Classification of soil is the dissolution to soil sets to particular group having a like characteristics and similar manners. Higher agricultural products are very much depends on the soil characteristics, identification and classification is very much important. Soil type helps to avoid agricultural product quantity loss.

J.Gholap presented and implemented information system Performance tuning of Algorithm for prediction of soil fertility. [9].Soil Test is the analysis of a soil sample to determine nutrient content, composition and other characteristics. Soil dataset containing soil test result has been used to apply various classification techniques in data mining.

H. U. Leena, B. G. Premasudha, and P. K. Basavaraja,[10] "Sensible approach for soil fertility management using GIS cloud, "presented an overview of the current state of the art of web-based tools and technologies for processing large dataset. These latest technologies allows to observe and survey a wider collection of agricultural related data such as different crops, soil type, variety of seeds, climatic conditions, and a mixture of socio economic variables concurrently and accurately. Integration of GIS technology into cloud together emerges as a crucial model for capturing uploading and maintaining the new promising web applications in agricultural sector.

K. Arunesh and V. Rajeswari, [11] has presented an efficient strategy for Lime soil status analysis using techniques. Agriculture is one of the most important research fields because variety of data is available in this field for researches. Techniques of data analysis including natural trees, statistical machine learning and other analysis methods.

M. N. Chiranjeevi and R.B.Nadagoudar developed a framework [12] Farmer must find which crop is suitable to cultivate in a particular soil sample. The farmers can get it from data mining Techniques.V. A. Bharadi "Analysis and Prediction in Agricultural Data Using Data Mining Techniques, [13] suggested the agricultural yield primarily depends on environmental factors such as rainfall, temperature and geographical topology of the particular region. These factors along with some other influence the crop cultivation.

N. Sneha and J. Majumdar, "Translated copy of Tank of [14] presented an efficient strategy for crops cultivation of crop yield prediction. It is the most important factor where the farmers need some prior information about the crop yield before sowing seeds in their fields with available requirements. Agriculture will face many challenges due environmental changes and natural calamities. To maximize the particular crop provided the prior information about the crop for the farmers can be done by analyzing historical crop data.

M. P. K. Devi, "Enhanced Crop Yield Prediction and Soil Data Analysis Using Data Mining [15] suggested a crop yield prediction and soil data analysis using data. Productivity of crop is one of the main issues in agriculture. The farmer is interested to know how much yield is about to expect. In earliest days, the yield prediction was carried out only by the experience of the farmers. As technology improved many ways that lead to know the productivity of crop. Crop models and decision tools are increasingly used in agricultural field to improved production competence. Crop yield predict analysis requires a model of how crops respond to soil factors.

K. Samundeeswari and K. Srinivasan, "Data Mining Techniques [16] presented an efficient techniques can be developed and tailored for solving complex soil data sets using data mining to improve the effectiveness and accuracy of the classification of large soil datasets.

III. PROPOSED WORK

A. Variation of Crop Yield with Soil

The soil of the Cuddalore district is classified as the black, red, ferruginous and arenaceous. They are again subdivided into clays, loam and sands. Black soils are observed. The soil is a major factor of agriculture Production. Dissimilarity of soil cans affected the crop production.

Table 3.1 General Statement for Soil Status

Elements	Very low	Low	Medium	High	Very high
pH	<5.0	5.1-6.5	6.6-7.5	7.6-8.0	>8.0
Organic carbon(OC)in %	<0.25	0.50-0	0.51-0.75	0.76-1.00	>1.00
Nitrogen (N)in kg/ha	<150	151-250	251-400	401-600	>600
Phosphorous(P)in kg/ha	<5	06-10	11-20	21-40	>40
Potassium(K) in kg/ha	<200	201-250	251-400	401-600	>600
Zinc(Zn) in kg/ha	<0.30	0.31-0.60	0.61-1.20	>1.20	Not Defined
Iron(Fe) in kg/ha	Not Define	<4.50	4.51-9.0	>9.0	Not Defined
Copper(Cu) in kg/ha	Not Defined	<0.20	0.21-4.0	>0.40	Not Defined
Manganese(Mn)in	<1.0	1.0-2.0	2.1-4.0	>4.0	Not Defined
Sulphur(S) in kg/ha	<10	11-20	21-30	31-40	>40

Table 3.2 Soil Attributes

SYMBOL	DESCRIPTION
S	Sand
Sicl	Silty Clay Loam
Sic	Silty Clay
C	Clay
Sl	Sandy loam
Cl	Clay Loam
Sil	Silty Loam
L	Loam
Ls	Loamy Sand
Scl	Sand Clay Loam
Sc	Sand Clay

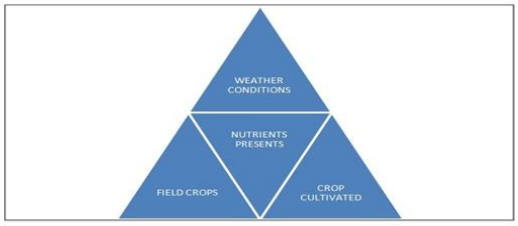


Figure 3.1 Soil Type

The figure 3.1 shows the chemical compositions of soil and its ratio of mixing with soil. The chemicals above are disturbs the growth of the crops depending upon its mixing ratio with soil.

Table 3.3 Soil PH Range

Attribute	Description
Ultra Acidic	<3.5
Extremely Acidic	3.5-4.4
Very Strongly Acidic	4.5-5.0
Strongly Acidic	5.1-5.5
Moderately Acidic	5.6-6.0
Slightly Acidic	6.1-6.5
Neutral	6.6-7.3
Slightly Alkaline	7.4-7.8
Moderately Alkaline	7.9-8.4
Strongly Alkaline	8.5-9.0
Very Strongly Alkaline	>9.0

Table 3.4 PH Level

ACIDIC	NETURAL	ALKALINE
LESS 6.0	7.5 – 8	8.6 – 9.0

Table 3.5 Soil Key Parameters

Attribute	Description
pH	pH Value of soil data
EC	Electrical Conductivity
OC	Organic
N	Nitrogen
P	Phosphorus
S	Sulphur
Zn	Zinc

B. Variation of Crop Yield with Water

The Geological processes role and significant role in cultivation. The humidity is one of the most considerable factors of geological processes.

Table 3.6 Water Range

Class	C1	C2	C3	C4	C5
EC	<1.5	1.5-3	3.5	5-10	>10
Water Quality	Normal Waters	Low Salinity Waters	Medium Salinity Waters	Saline Water	High Salinity Water

C. Principal Component Analysis

PCA is a familiar strategy who explore important concept of variation in a model of random curves. This major source of variation are represented by functional principal component analysis.PCA can be used within huge reports is a bilinear interpretable overview of the important message had a multidimensional table, projection method because instructions carried by the original variables and projects them onto a smaller number of latent variables. Matrix is representation a model with a given number of components equations.

$$X = TP^T + E$$

T ->Score Matrix
P-> Load Matrix
E-> Error Matrix

D. Partial Least Squares

Projection of Latent Structure both the X-and Y-matrixes at the same time to find the least variables in X that will best to predict the least variables in Y. The PLS components are parallel to principal components they are referred to as Factors.PLS magnifies the equivalent in centre of X and Y.

E. Support Vector Machines

SVM is a machine learning (Supervised Learning) technique. Which means nonlinear attributes and it is used for classification as well as regression problem. Linear classification is anyway easy to implement as analyse to nonlinear classification.

F. PLS Based Field Crop Yield

PLS is a statistical method actual choice of relation to Principal component regression, alternatively finding hyper plane of maximal variance between the reaction and impartial variables, it initiate a linear regression model by predicted variables. This method of work is used to realise classification regression problems. When the dataset is complex and large then we frequently go for linear version.

IV. PERFORMANCE EVALUATION

Descriptive statistics provides some simple and valuable plotting to medium size datasets. We receive and get a data from various farmers of villages and create the sequential datasets. Actually data of cultivation about field Crops, Soil and Water.

Generated From – The reports of farmers to the Government.

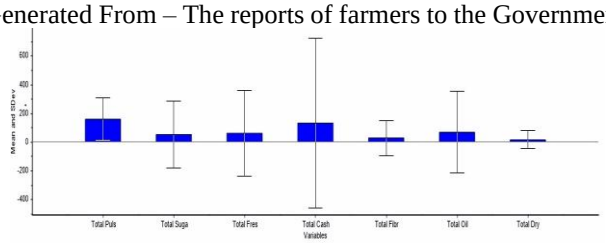


Figure 4.1 Various Crops Yield

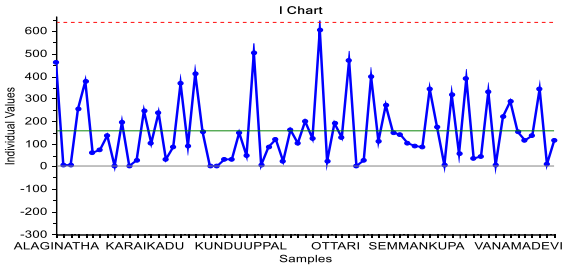


Figure 4.2 Various Soil Samples and Crop

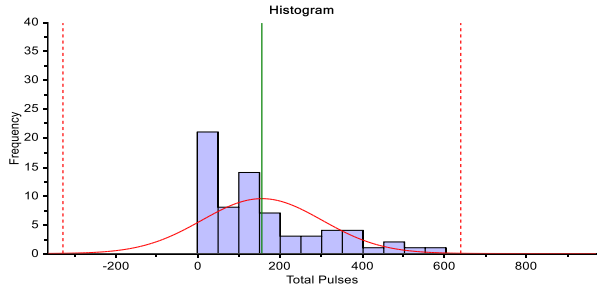


Figure 4.3 Histogram of Pulses

A. PLS based Soil

Obviously the PLS method was better than various models for the plan of maximum yield from soil status it would be a valuable tool to conform with sample is need future analysis. The character of PLS factor was identified using cross validation during the time of soil or water.

Table 4.1 PLS Statement of Soil

Soil		EC	pH	N	P	K	Zn	Cu	Fe	Mn	B	ca	mg	S
		1	2	3	4	5	6	7	8	9	10	11	12	13
ALAGINAT...	1	0.6400	7.1000	67.0000	14.0000	83.0000	0.5900	1.5700	17.5500	10.6500	0.5000	1560.0000	732.0000	6.2500
ANNAVELI	2	0.4600	7.6000	55.0000	17.0000	105.0000	0.8600	0.8200	17.3100	7.2600	0.5000	786.0000	226.0000	6.2500
ARISIPERI...	3	0.1000	7.6000	50.0000	13.0000	40.0000	0.8600	0.9100	13.9400	8.0100	0.0000	655.0000	241.0000	0.0000
C.N.PALAY...	4	0.3000	6.8000	42.0000	16.0000	63.0000	1.1000	1.3800	16.9800	8.3200	0.5000	570.0000	278.0000	12.5000
CHELLANC...	5	0.1100	7.5000	53.0000	9.0000	35.0000	0.8800	0.8100	13.7400	6.9200	0.0000	648.0000	292.0000	0.0000
CINNAKAM...	6	0.7200	6.8000	63.0000	14.0000	95.0000	0.9600	1.4100	19.7200	10.7200	0.5000	800.0000	332.0000	6.2500
CUDDALO...	7	0.3200	8.1000	35.0000	19.0000	133.0000	1.0500	1.3900	17.7200	8.7200	0.5000	1080.0000	252.0000	12.5000
CUDDALO...	8	0.1000	5.4000	56.0000	11.0000	43.0000	0.9300	0.8200	14.3300	7.4600	0.0000	648.0000	276.0000	0.0000
DEVANAM...	9	0.3000	7.5000	59.0000	9.0000	40.0000	1.1200	0.9900	13.3600	6.3200	0.0000	716.0000	312.0000	0.0000
ERAIKADY...	10	0.2200	6.9000	59.0000	20.0000	88.0000	0.4800	1.6000	17.5500	10.6400	0.5000	906.0000	732.0000	6.2500
GENGEKU...	11	0.1100	6.3000	83.0000	16.0000	30.0000	0.7400	0.6200	15.0100	7.6700	0.0000	786.0000	342.0000	0.0000
KALAIYUR...	12	0.5200	7.9000	59.0000	17.0000	85.0000	0.7900	0.8200	15.2800	7.5900	0.5000	756.0000	292.0000	6.2500
KARAIKADU...	13	0.7400	7.5000	73.0000	15.0000	33.0000	0.9800	0.9200	14.1100	7.0200	0.0000	745.0000	222.0000	0.0000
KARAIMEDU...	14	0.7000	6.7000	62.0000	13.0000	63.0000	18.6100	0.5800	9.7500	1.3600	0.0000	800.0000	348.0000	6.2500
KARAIYUR...	15	0.2700	7.1000	53.0000	13.0000	130.0000	1.3000	1.3600	20.3900	8.2300	0.0000	1120.0000	420.0000	6.2500
KARAMAN...	16	0.8100	6.8000	59.0000	14.0000	85.0000	0.7100	1.2600	20.1600	10.3400	0.0000	900.0000	384.0000	6.2500
KARANAPP...	17	0.6000	6.7000	52.0000	11.0000	63.0000	0.9500	1.7200	18.4200	9.4500	0.0000	740.0000	180.0000	6.2500
KARUPPA...	18	0.6200	7.4000	18.0000	16.0000	30.0000	0.9200	0.8300	14.2100	6.9300	0.0000	698.0000	326.0000	0.0000
KILAZCHIR...	19	0.4200	6.5000	63.0000	7.0000	20.0000	0.8600	0.8100	15.8600	5.8600	0.0000	671.0000	231.0000	0.0000
KILINCHIK...	20	0.8700	7.7000	60.0000	23.0000	133.0000	0.9600	1.3800	17.6600	9.1000	0.5000	440.0000	264.0000	6.2500
KILKUMAR...	21	0.2800	5.9000	73.0000	14.0000	15.0000	0.9200	0.8500	14.7300	7.1400	0.0000	618.0000	276.0000	0.0000
KONDURUN...	22	0.3500	7.2000	63.0000	20.0000	78.0000	0.7100	1.2100	17.1600	7.1600	0.0000	840.0000	384.0000	6.2500
KUDIKADU...	23	0.0600	6.5000	69.0000	31.0000	55.0000	0.7600	1.4100	21.3800	11.3500	0.0000	780.0000	384.0000	12.5000
KUMARAP...	24	0.4100	7.2000	59.0000	21.0000	85.0000	0.7800	0.9400	18.2600	7.4100	0.5000	922.0000	376.0000	6.2500
KUNJAMAN...	25	0.2400	7.4000	69.0000	11.0000	90.0000	1.1700	0.9800	16.4700	8.4300	0.0000	738.0000	344.0000	0.0000
KUNDUUP...	26	0.3800	7.2000	65.0000	19.0000	285.0000	1.1200	1.2100	12.7480	7.5200	1.0000	1390.0000	288.0000	12.5000
MALAYAPE...	27	0.1500	6.5000	65.0000	13.0000	90.0000	1.1800	1.8200	17.3500	8.3800	0.0000	960.0000	408.0000	6.2500
MANJAKKU...	28	0.1100	6.4000	31.0000	16.0000	25.0000	0.9400	0.8500	15.8200	7.8600	0.0000	601.0000	301.0000	0.0000
MARUTHA...	29	0.1000	6.5000	42.0000	14.0000	63.0000	0.8100	0.7900	15.7200	6.7100	0.0000	642.0000	283.0000	0.0000
MATHALAP...	30	0.3800	7.5000	70.0000	21.0000	83.0000	0.9300	0.8200	15.2200	8.2100	0.5000	690.0000	376.0000	6.2500
MAYADIPA...	31	0.5200	7.4000	64.0000	13.0000	75.0000	1.0900	1.6200	19.8200	10.8200	0.0000	780.0000	108.0000	6.2500
MELAKKUP...	32	0.3100	7.4000	56.0000	14.0000	115.0000	0.9100	0.7700	16.5400	8.3200	0.5000	837.0000	298.0000	6.2500
MELAZHIN...	33	0.4400	6.8000	64.0000	26.0000	100.0000	1.2140	1.5500	13.3400	6.2400	0.0000	845.0000	626.0000	0.0000

Table 4.2 PLS Statement of Water

Statistics		EC	pH	N	P	K	Zn	Cu	Fe	Mn	B	ca	mg	S
		1	2	3	4	5	6	7	8	9	10	11	12	13
# of Missing	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mean	2	0.3204	7.0403	57.5942	15.1159	71.5217	1.1688	1.0542	16.3148	7.8297	0.1678	789.5217	308.2029	21.8387
Max	3	0.8700	8.1000	84.0000	31.0000	285.0000	18.6100	1.8200	21.3800	11.4200	1.0000	1560.0000	732.0000	625.0000
Min	4	0.0600	5.4000	18.0000	6.0000	15.0000	0.3500	0.4700	9.7500	1.3600	0.0000	440.0000	84.0000	0.0000
Range	5	0.8100	2.7000	66.0000	25.0000	270.0000	18.2600	1.3500	11.6300	10.0600	1.0000	1120.0000	648.0000	625.0000
Std Deviation	6	0.2033	0.5583	14.3151	5.7583	41.2106	2.1380	0.3618	2.3149	1.5276	0.2682	192.1049	113.5096	105.0499
Variance	7	0.0413	0.3117	204.9211	33.1275	1698.3120	4.5710	0.1309	5.3587	2.3335	0.0719	36904.2800	12884.4300	11035.4800
RMS	8	0.3787	7.0621	59.3215	16.1537	82.3957	2.4230	1.1137	16.4758	8.0754	0.3147	812.2278	328.1566	106.5480
Skewness	9	1.0316	0.0044	-0.3799	0.6673	2.1389	8.2172	0.5252	0.0751	-0.4480	1.2674	1.8871	1.4879	5.7275
Kurtosis	10	0.2491	0.1128	0.2227	0.0338	9.4972	68.0029	-0.9412	-0.3758	4.3393	0.7229	4.8462	5.1075	31.7783
Median	11	0.2700	6.9000	59.0000	14.0000	70.0000	0.9400	0.9100	15.9200	7.5800	0.0000	732.0000	292.0000	6.2500
Q1	12	0.1900	6.6800	52.0000	11.0000	40.0000	0.8600	0.8200	14.3600	7.1400	0.0000	698.0000	264.0000	0.0000
Q3	13	0.4100	7.4900	67.0000	18.0000	90.0000	1.0100	1.3600	18.0600	8.4500	0.5000	845.0000	348.0000	6.2500

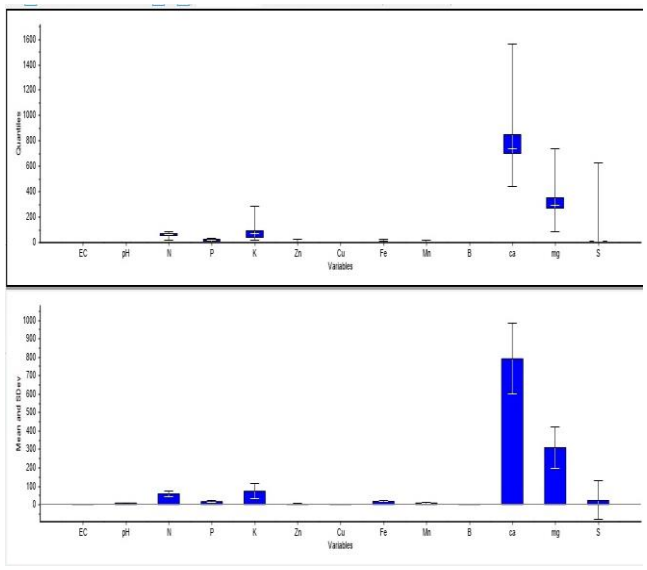


Figure 4.4 Graphical view of Water

V. RESULT AND DISCUSSION

The above data sets are finalized in this paper is achieved from the agriculture soil Testing Laboratory, Cuddalore District, Tamilnadu, India. Field sampling gives us the primary data for soil survey.

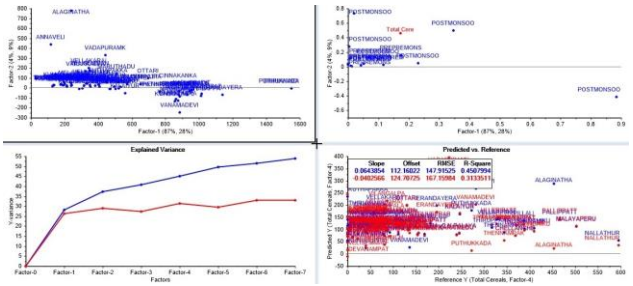
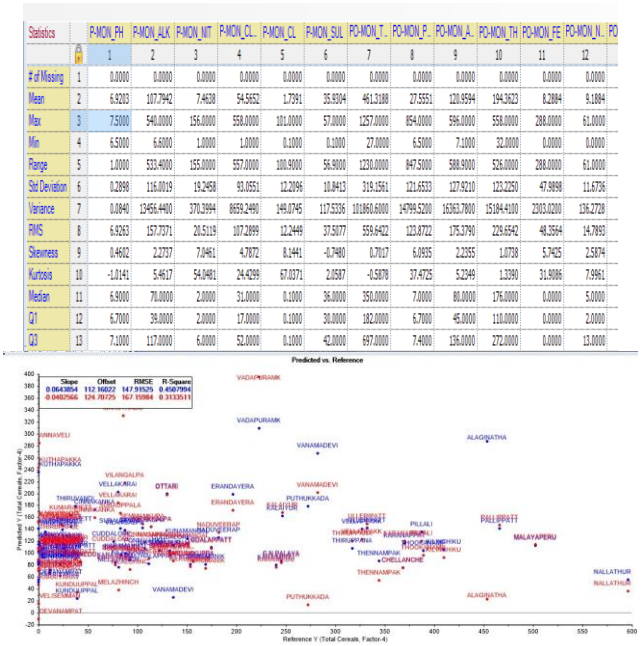


Figure 5.1 Graphical view of PCS Range

Dataset has 13 attributes such village name as Ec, pH, N, P, K, Zn, Cu, Fe, Mn, B, Ca, Mg, and S. As before the primary data for water and Field Crop also acquired, observed and collected directly from the farmers. The following dataset has 15 attributes of water. In this work we have proposed and analysis of the soil, water and field crop data using classification techniques and prediction techniques to predict the status for maximized yield.

We have reported comprehensive study of various classification Algorithms with the Principal component Analysis, Partial Least Squares Regression, Descriptive statistics must perform efficiently. In future we decided to purpose a system to predict the suitable Field Crop for a particular soil and water.



Statistics		F-MON_PH	F-MON_ALK	F-MON_NIT	F-MON_CL..	F-MON_CL	F-MON_SUL	PO-MON_T..	PO-MON_P..	PO-MON_A..	PO-MON_TH	PO-MON_FE	PO-MON_N..	PO
	A	1	2	3	4	5	6	7	8	9	10	11	12	
# of Missing	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
Mean	2	6.9203	107.7942	7.4638	54.5652	1.7791	35.9304	461.3188	27.5551	120.9594	194.3623	8.2884	9.1884	
Max	3	7.5000	540.0000	156.0000	558.0000	101.0000	57.0000	1257.0000	854.0000	596.0000	558.0000	288.0000	61.0000	
Min	4	6.5000	6.8000	1.0000	1.0000	0.1000	0.1000	27.0000	6.5000	7.1000	32.0000	0.0000	0.0000	
Range	5	1.0000	533.4000	155.0000	557.0000	100.9000	56.9000	1230.0000	847.5000	588.9000	526.0000	288.0000	61.0000	
Std Deviation	6	0.2898	116.0019	19.2458	93.0551	12.2096	10.8413	319.1561	121.6533	127.9210	123.2250	47.9898	11.6736	
Variance	7	0.0840	13456.4400	370.3994	8659.2490	149.0745	117.5336	101860.6000	14799.5200	16363.7800	15184.4100	2303.0200	136.2728	
RMS	8	6.9263	157.7571	20.5119	107.2899	12.2449	37.5077	559.6422	123.8722	175.3790	228.6542	48.3564	14.7893	
Skewness	9	0.4602	2.2737	7.0461	4.7872	8.1441	-0.7480	0.7017	6.0935	2.2355	1.0738	5.7425	2.5874	
Kurtosis	10	-1.0141	5.4617	54.0481	24.4299	67.0371	2.0587	-0.5878	37.4725	5.2349	1.3390	31.9086	7.9961	
Median	11	6.9000	70.0000	2.0000	31.0000	0.1000	36.0000	350.0000	7.0000	80.0000	176.0000	0.0000	5.0000	
Q1	12	6.7000	38.0000	2.0000	17.0000	0.1000	30.0000	182.0000	6.7000	45.0000	110.0000	0.0000	2.0000	
Q3	13	7.1000	117.0000	6.0000	52.0000	0.1000	42.0000	697.0000	7.4000	136.0000	272.0000	0.0000	13.0000	

Figure 5.2 Overall Statistical Statements of Water and Field Crop

VI. CONCLUSION

The intention of this paper is to intend the multivariate analytics and Machine Learning Algorithms; we forecast a statement to improve the agriculture yield. For that prediction we gather reports from farmers of each village based on water and soil properties. This will also offer the well-organized information regarding the soil and water level like, **pH, Ec, N, P, K, Zn, Cu, Fe, Mn, B, Ca, Mg,** and **S** along with the proper implication.

Finally, it is accomplished that, with this projected system one can save manpower and water to extemporize the production.

REFERENCES

1. Adamala, "An Overview of Big Data Applications in Water Resources Engineering," vol. 2, no. 1, pp. 10–18, 2017. S. P. R. Kolhe, S. P. P. Kolhe, and M. H. Tharkar, "ROLE OFCR INFORMATION TECHNOLOGY IN (An emerging dimension in IT)," vol. 2, no. 10, pp. 1–3, 2014.
2. S. P. R. Kolhe, S. P. P. Kolhe, and M. H. Tharkar, "ROLE OF INFORMATION TECHNOLOGY IN (An emerging dimension in IT)," vol. 2, no. 10, pp. 1–3, 2014.
3. S. Kumawat, M. Bhamare, A. Nagare, and A. Kapadnis, "Sensor Based Automatic Irrigation System and Soil pH Detection using Image Processing.," pp. 3673–3675, 2017.
4. D. Haryana, V. K. Singh, R. Prakash, M. A. Bhat, and G. Deep, "Evaluation of groundwater quality for irrigation in Kaithal block (Kaithal District) Haryana," no. February, 2018.
5. Dr.S.Padmavathi, S. (n.d.). Performance of SVM Classifier for Image Based Soil Classification. International conference on Signal Processing. SCOPUS.
6. Bhushan Naib, B. (2013). Soil Classification and Fertilizer R Soil Recommendation using WEKA using IJCSMS International Journal of Computer Science & Management Studies (Vol. 13). Retrieved from.
7. Ganesh, S. H., Jayasudha, M., & Scholar, M.P. (2015). An Enhanced Technique to Predict the Accuracy of Soil Fertility in Agricultural Mining. International Journal of Advanced Research in Computer and Communication Engineering,
8. Mutalib, S., Fadhlun Jamian, S. N., Abdul-Rahman, S., & Mohamed, A. (2010). Soil classification: An application of self organizing map and k-means. In Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications,ISDA'10.http://doi.org/10.1109/ISDA.2010.5687224
9. J.Gholap "Performance Tuning of J48 Algorithm for Prediction of Soil Fertility," arXivPrepr. ArXiv, vol. 1208, p. 3943, 2012.
10. H. U. Leena B. G. Premasudha, and P. K. Basavaraja, "Sensible approach for soil fertility management using GIS cloud," 2016 Int. Conf. Adv. Computer Communication Informatics, ICACCI 2016, pp. 2776–2781, 2016.
11. P. R. No, "Big Data Frame work To Identify Agriculture Crop and Aquaculture Fish Based Diseases and Recommendation of Solution Project Reference No: 40S _ BE _ 0460 Introduction : Objective : Methodology :K.Arunesh and V.Rajeswari, "Agricultural Soil, Lime Status Analysis Using Data Mining Classification PP : 174–182.
12. M.N.Chiranjeevi and R.B.Nadagoudar, "Analysis of Soil Nutrients by using Data Mining Techniques," Int. J. Recent Trends Eng. Res., vol.4, no.7, PP. 103-107, 2018.
13. V. A. Bharadi, P. P. Abhyankar, R. S. Patil, S. S. Patade, T. U. Nate, and A. M. Joshi, "Analysis and Prediction in Agricultural Data Using Data Mining Techniques," Int. J. Res. Sci., no. 7, pp. 386–393, 2017.
14. E.Manjula and S. Djodiltachoumy, "Available Online at www.ijarcs.info DATA MINING TECHNIQUE TO ANALYZE SOIL NUTRIENTS BASED ON HYBRID CLASSIFICATION," vol. 8, no. 0976, pp. 505–510, 2017.
15. M. P. K. Devi, "Enhanced Crop Yield Prediction and Soil Data Analysis Using Data Mining," vol. 4, no. 6, 2016.
16. K. Samundeeswari and K. Srinivasan, "Data Mining Techniques in Agriculture Prediction of Soil Fertility," vol. 8, no. 4, pp. 45–51, 2017.