Optimal Ensemble Technology for Precision Agriculture (OETPA)

Dr. Sangeeta Ahuja
Scientist (S.S.)
IASRI (CAR), New Delhi-12, INDIA.

Abstract

Precision Agriculture is a thumb rule for increasing the production and productivity of different crops. The tools, techniques and technologies which can implement for precision Agriculture have to design and develop so that with minimum resources, the production and productivity of different crops will be maximized. Keeping above issues in mind, the tools and technologies have been designed for achieving the optimum solution. The advanced methodologies of optimality with ensemble techniques have been implemented in order to get the precision agriculture. The optimum variety of crop, dose of treatment, fertilizer, irrigation, soil etc. have been selected by optimality solution. The results of the methodologies have been tested also statistically at different level of significance. All possible levels of various treatments have been computed with inclusions of treatment means and standard errors. The advanced statistical computation of R-Square, RMSE, CV and treatment critical difference have also been computed and utilized for varying the different parameters. Rigorous testing, experimentation and critical evaluation and its implementation, results shows efficient improvement for various crops production and productivity and completely fulfulling the targets of achieving the Precision Agriculture.

Keywords: Precision Agriculture, Optimality, Irrigation, CV, RSquare, Efficiency etc.

I Introduction

OETPA (Optimal Ensemble Technology for Precision Agriculture) is novel algorithm for Cluster ensemble. Cluster ensembles have emerged as a technique for overcoming problems with clustering algorithms. It is well known that off-the-shelf clustering methods may discover different behavior in a given set of data. High throughput data technologies allow the production and analysis of agricultural data to address critical questions related to selection of minimum resources with the maximum benefit. The principle behind the use of precision agriculture is basically to enhanced product quality, increased agricultural profitability and sustainability, protecting the environment, optimized use of agricultural pesticides, fertilizers, seeds, water, energy and other crop amendments so the selection of the resources leads to optimum. Keeping above issues in mind, the methodologies with technology have been designed for achieving the optimum solution. An Optimal Ensemble Technology for Precision Agriculture (OETPA) which generates the best suitable grouping of different treatment combinations, fertilizers doses, selection of varieties which is based on their effectiveness/performance towards the optimum productivity.

Motivation

The Agriculture sector is important assets for any country. The land, the soil content, the perfect combination of fertilizers, treatment combinations, pesticides, the right amount of irrigation level etc. are very much needed for maintaining the precision agriculture. To optimize the yields, suitable treatment combinations are to be identified and standardized for a particular set of agricultural components. In view of the above, need arises of an computer tool based algorithm which gives the best combination of all the resources by understanding the performance of treatment complex combinations by analyzing huge datasets in fraction of seconds. There was a need to develop these kinds of algorithms for Precision Agriculture to optimize the productivity. OETPA is also beneficial for all fields of Agriculture. It also gives comparison of all the efficient groups. The treatment combinations of doses, fertilizers, variety of crops and their spacing i. e. geometrical arrangements, canopy manipulations, crop harvest intervals, irrigation schedules etc. are standardized specifically to develop different agricultural Models.

II Methodology of Cluster Ensemble in OETPA

OETPA is based upon the concept of cluster ensemble [1,2,3].

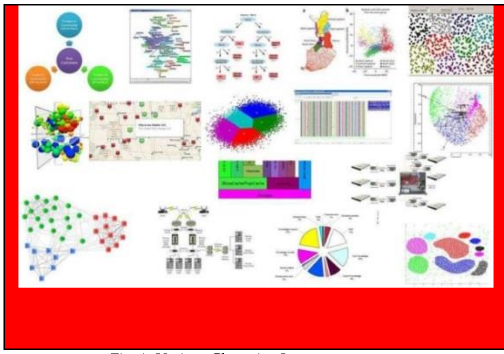

Fig. 1: Various Clustering Images

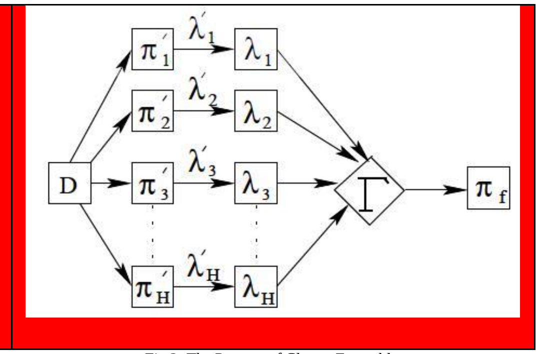

Fig 2: The Process of Cluster Ensemble

The process of cluster ensemble in general diagrammatically shown as in Fig. 2. OETPA aims at improving robustness and quality of clustering scheme or Efficient Group, particularly in Agriculture sector which in turn enhance the production and productivity of any crop. It operates in multiple phases. In the initial phase generates the multiple clustering schemes and gives the allocation with relabellings. In the next phases Efficient groups have been computed by implementing different statistical and engineering methods by varying with different threshold level by reaching to the optimum. Efficient Groups are determined in descending order of optimum results i. e. Efficient Group 1 gives the maximum yield /best dose of fertilizer/best treatment combination, survival percentage etc. followed by other Efficient Groups respectively. The process is diagrammatically shown as in Fig. 3.

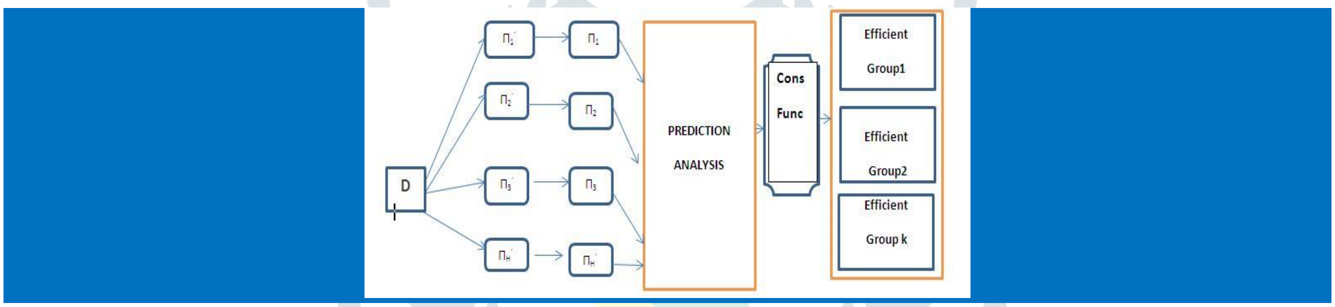

Fig. 3: OETPA

Experimentation

Extensive experimentation [4-8] has been done using Agriculture data by varying the number of partitions and clusters in cluster ensemble. Different Efficient Groups are achieved by using this technique that segregates the various resources in order to achieve the optimum production. The first Efficient Group 1 refers to the selection of the resources (treatments, spacing, diameters and heights) which gives the maximum yields and survival depending upon the context. Similarly Efficient Group 2 gives the next best solution given by the algorithm and so on and so forth. Furthermore, we investigate in depth the about the quality, accuracy and stability of results by using different Efficient Groups by utilizing the various quality and diversity measures viz., Purity, Normalized Mutual Information and Adjusted Rand Index. Then the comparison of OETPA is made with various traditional clustering algorithms in terms of the quality and stability measures. Except this, further testing is done by computing the Mean, Standard Deviation and Coefficient of Variation of each Efficient Group. The results are further verified by SAS and computation of ANOVA (Analysis of Variance) for Treatment Effects, Root Mean Square Error, R Square, Coefficient of Variation, all pair wise treatment comparisons are determined for each character in each data set. The result is also statistically tested at different level of significance and Critical Differences were also obtained. Experimental results show that the proposed techniques (OETPA) are capable of producing a partition that is far better than the best individual clustering. It gives the optimum output with the combination of minimum input resources and leads to the precision agriculture.

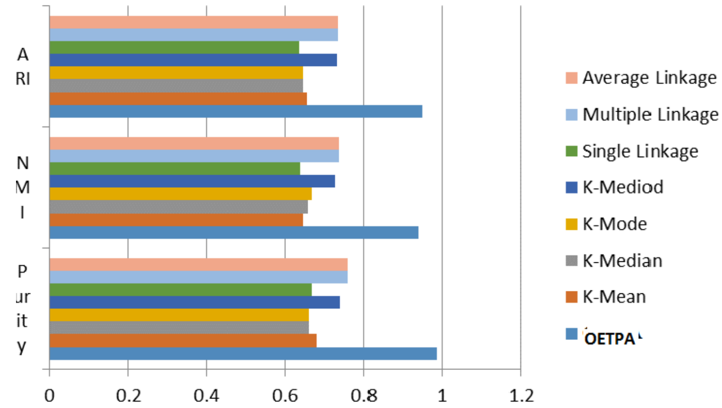

Fig. 4: Comparisons of Quality and Stability Measures for OETPA and Traditional Clustering Algorithms

OETPA is compared with traditional clustering schemes in terms of quality and stability measures viz., (Purity, Normalized Mutual Information (NMI) and Adjusted Rand Index(ARI)) (See Fig.4). The histogram clearly shows that improvement is achievement in terms of all the quality measures i.e. purity, NMI and ARI in the case of OETPA as compared to other traditional clustering scheme viz., Average Linkage, Multiple Linkage, Single Linkage, K-Mediod, K-Mode, K-Mean respectively. So, all the further computations and analysis have been done on all the data sets by using OETPA. After that the Mean, Standard Deviation and Coefficient of Variation of each Efficient Group is determined. The results are further tested and verified by SAS (Statistical Analysis System) and computation of Root Mean Square Error (RMSE), General Mean and R Square is also obtained. The computation of ANOVA (Analysis of Variance) for Treatment Effects for each character and for each data set is determined. Moreover, the statistical significance of the result is tested by computing the Critical Differences of treatments at different levels of significance.

In these cases, the treatments performance is considered as cluster in OETPA and the results from analysis through the algorithm are shown in all cases as given below. Results by the Algorithm OETPA of all Cases clearly show improvement in terms of yield and survival percentage in the Efficient Groups as compared to Control. OETPA gives each Efficient Groups in descending order of yield percentage. The first Efficient Group 1 refers to the selection of the resources (treatments, spacing, diameters and heights) which gives the maximum yields and survival depending upon the context. Similarly Efficient Group 2 gives the next best solution given by the algorithm. The result is further statistically tested for the accuracy and stability. The detailed analysis and interpretation of the results obtained by OETPA are as follows

CASE 1: Growth Performance of Bamboo seedlings

The tree growth parameter like culm length, collar diameter and number of new shoot emergence were recorded and data is presented in TABLE-1 below.

Table-1: Growth parameter of *B. balcooa* seedling under various spacing trials.

Treatment	Spacing	Culm length (cm)		Collar diam (cm)		No. of new shoot	
		2008	2009	2008			2008
Control		118	141. 2	1. 23	Control		118
T1	3x3	160. 7	185. 7	1. 40	T1	3x3	160. 7
T2	3x4	166. 6	193. 1	1. 64	T2	3x4	166. 6
T3	4x4	187. 3	217. 6	1. 89	T3	4x4	187. 3
T4	5x5	314. 6	377. 7	2. 3	T4	5x5	314. 6

Result by OETPA

Efficient Groups	Year	Treatment	Yield
Efficient Group – 1	Y2	T3	426. 5667
	Y2	T4	528. 0667
Efficient Group – 2	Y1	T1	228. 5667
	Y2	T1	261. 4667
	Y1	T2	228. 5667
	Y2	T2	267. 1

	Y2	T2	228. 9
	Y1	T3	271. 8
	Y2	T3	283. 8
	Y1	T4	277. 6333
	Y2	T4	324. 5667
	Y1	T4	305. 5667
Efficient Group – 3			
	Y1	T1	139. 6333
	Y2	T1	192. 2
	Y1	T2	139. 6333
	Y1	T3	210. 5333
Efficient Group – 4			
	Y2	T1	26. 66667
	Y2	T2	40. 5
	Y2	T3	58. 86667
	Y1	T4	44. 86667
	Y2	T4	68. 9
Efficient Group – 5			
	Y1	T1	4. 31
	Y2	T1	4. 356667
	Y1	T1	1. 433333
	Y2	T1	1. 756667
	Y1	T1	22. 33333
	Y1	T2	4. 343333
	Y2	T2	4. 486667
	Y1	T2	1. 46
	Y2	T2	1. 773333
	Y1	T2	26. 16667
	Y1	T3	4. 536667
	Y2	T3	4. 803333
	Y1	T3	1. 453333
	Y2	T3	1. 856667
	Y1	T3	33. 76667
	Y1	T4	5. 076667
	Y2	T4	5. 063333
	Y1	T4	1. 46
	Y2	T4	1. 953333

Mean, Standard Deviation and Coefficient of Variation for Efficient Groups

Mean(μ) , Standard Deviation(σ) and Coefficient of Variation(CV)

Efficient Group	(μ+/- σ)	CV
Efficient Group 1	477.3167+/-7.123903	0.014925
Efficient	245.0167+/-4.05586	0.016553

Group 2		
Efficient Group 3	170.5+/-5.616039	0.032939
Efficient Group 4	42.725+/-3.389498	0.079333
Efficient Group 5	6.967895+/-3.026792	0.434391
Control	48.66433+/-52.43159	1.077413

Statistical Analysis through SAS

CASE I: Growth Performance of Bamboo seedlings

Table 4(a:) For Culm length 2008

ANOVA (Adjusted for Treatment Effects)

Source	DF	SS	MS	FCAL	PROB>F
Block (Unadj.)	2	0.00065333	0.00032667		
Treatment(Adj.)	4	2.10477333	0.52619333	543.40103270	0.00001000
Error	8	0.00774667	0.00096833		

R-Square	RMSE	General Mean	C.V.
0.996334	0.031118	1.695333	1.835512

Treatment Means and Their Standard Errors

Treatment	Mean	SD
1	1.236667	0.020817
2	1.406667	0.023094
3	1.640000	0.017321
4	1.893333	0.051316
5	2.300000	0.017321

All possible Paired Comparison of Treatments Prob>F

	1	2	3	4	5
1	.	0.00015	0.00001	0.00001	0.00001
2	0.00015	.	0.00002	0.00001	0.00001
3	0.00001	0.00001	.	0.00001	0.00001
4	0.00001	0.00001	0.00001	.	0.00001
5	0.00001	0.00001	0.00001	0.00001	.

Treatment Critical Difference

C.D. for Treatments (1%)	0.090325
C.D. for Treatments (5%)	0.058590

Interpretation: The analysis given by **OETPA** in various Efficient Group of Case I results revealed that the mean annual increment of plant height and collar diameter differed significantly between spacing and age of the seedlings. A steady growth in height and diameter was observed in during the first year after transplantation. However, during the 2nd year onwards a higher growth rate was observed. The slower growth rate during first year could be due to the fact that the seedlings were not able to establish properly through uptake of nutrient from soil required by them for their growth purposes. Maximum culm length was noticed in 5x5 m spacing followed by 4x4, 3x4 and 3x3 m spacing respectively. From this observation it was revealed that least number of new shoot developed in first year and its number gradually increased in successive year. Among these four spacing's, 5x5 m spacing have shown the better growth as well as shoot emergence was observed as compare to others spacing.

Conclusion

The advanced methodologies of optimality with ensemble techniques have been implemented in order to get the precision agriculture. The optimum variety of crop, dose of treatment, fertilizer, irrigation, soil etc. have been selected by optimality solution. The results of the methodologies have been tested also statistically at different level of significance. All possible levels of various treatments have been computed with inclusions of treatment means and standard errors. The advanced statistical computation of R- Square, RMSE, CV and treatment critical difference have also been computed and utilized for varying the different parameters. Rigorous testing, experimentation and critical evaluation and its implementation, results shows efficient improvement for various crops production and productivity and completely fulfulling the targets of achieving the Precision Agriculture.

References

1. Jiawei Han and Micheline Kamber ," Data Mining : Concepts and Techniques" Second Edition, Morgan Kaufmann Publishers, San Diego, USA, 2006.
2. Margaret H. Dunham, "Data Mining: Introductory and Advanced Topics", Southern Methodist University, Pearson Education Inc., Upper Saddle River, New Jearsey, 2003.
3. Kuncheva, L and others, "Evaluation of stability of K-means cluster ensembles with respect to random initialization)", journal of "IEEE Transcations on pattern analysis and machine intelligence, 1798-1808 pp, 2006
4. Kruglinski, David - Inside Visual C++, I Edition. Microsoft Press, Washington, 1996.
5.Richard A. Johnson and Dean W. Wichern, "Applied Multivariate Statistical Analysis", Prentice Hall, Upper Saddle River, New Jersey, 1979.
6. Jeff, Prosise. Programming Windows with MFC, II Edition. Microsoft Press, Washington, 1999.
7. Ahuja, S, et al., SPFE 1.0 Software, 2005.
8. SAS Software., 2015.