

MODELLING A PREDICTIVE ANALYTICS METHODOLOGY FOR FORECASTING RICE VARIETY AND QUALITY ON YIELD ON FARM AND FARMING ATTRIBUTES USING BIGDATA

Senthil Vadivu . S¹, Dr. Prasanna Devi .S², Vinu Kiran .S³, Manivannan .S⁴

¹Research Scholar, MS University, Tirunelveli, India

¹senthumukund@gmail.com

²Head, Computer Science & Department, Apollo Engineering College, Chennai, India

³Data Analyst, Business Intelligence-Center of Excellence, HTC Global, Chennai, India

⁴Deputy Dean, Dr.MGR Educational & Research Institute University, Chennai, India

Abstract: In this paper, we have analyzed the attributes of farm land and farming methods and applied predictive analytics methodology over huge data set to forecast the outcome of the commodity. We have took “Rice” as a case study since we had done some geological survey work on Rice fields in and around Kanchipuram, India. The purpose of this study is to collect information on the current and foreseen land use practices with particular emphasis on the role of cropping/farming systems to predict the outcome of the crop quality (Rice).

Keywords: Agriculture, Big Data, clustering, Data mining, Machine learning, Predictive analytics.

1. Introduction

Like the rest of the delta region[1] in India, most of the land region is endowed with some of the most favourable conditions for agricultural production. Such climatic and soil quality conditions have given rise to relatively high yielding capacity for both human and livestock feeds. Approximately 40-45% of the population is engaged in a wide range of agricultural activities in full-fledge.

However, the natural resources of our continent is facing increasing threat due to population pressure, increase of agricultural and livestock activities and unsustainable farming practices. This has led to serious land degradation and decline in productivity of agricultural commodity.

In recent years, due to land degradation to a large extent by the inefficient farming practice, the quality of commodity were getting degraded. In addition to that due to soil quality and some of human activities along the rivers/lake is that soil erosion from the catchment areas increases the

sediment load of the rivers and lakes resulting in poor quality of the commodity.

we have analyzed the attributes of farm land and farming methods and applied predictive analytics methodology over huge data set to forecast the outcome of the commodity. We have took “Rice” as a case study since we had done some geological survey work on Rice fields in and around Kanchipuram, India. The purpose of this study is to collect information on the current and foreseen land use practices with particular emphasis on the role of cropping/farming systems to predict the outcome of the crop quality

2. Farm and farming conditions

Like the rest of the top agricultural areas of India, the Delta region and the region around Palar river basin is endowed with some of the most favourable conditions for agricultural production[9]. The temperatures over most of the region are likely to be favourable for Rice production all the year round. Most of the region receives between 400mm(min)-1500mm(max)[2] approx. of rain annually and this is utilized well by the farming activities. It is therefore possible to have more than two cropping seasons a year under rain fed conditions in much of the Kagera basin. The soils, though mainly weathered tropical soils liable to leaching, are able to support vigorous crop growth. The farmers or crop/farming systems experts are expected to show their expertise on quality and farming issues, adopting best practices to retain crop quality, identifying degradation indicators and assessment methods, to conduct survey/assessment and analyze the current growth conditions.

3. Bigdata analytics on farm data

Big data analytics helps us on turning information into insights by connecting the dots for evidence based decision making. The outcome of analytics helps us to integrate solutions for complex problems. Big data analytics on farming and agriculture [3] can be pretty much helpful for expanding capacity in plant biology, transformation of plants to improve crops and yields, facilitation of automation of large-scale plant analysis and for commercialization of commodity decision making based on trend analysis. It helps us on prediction of resources such as climate and weather for agricultural cultivation.

We can leverage the vast data sets such as short and long term weather, data sets and soil category from national soil database etc., BigData analytics [5] can helps us to implement and design a feasible network of integrated farm models, crop growth models, soil water balance model, algorithm for machinery selection based on crop type and algorithm for farm optimization. Data inputs can be drilled down to find out the pattern of crops and commodities for machinery and labor availability, crop rotations, yield and crop input maps to work out with marketing strategies and for picking and locating correct inventories to balance the demand and supply..

4. Methodology

4.1 Field Visit for Rice Assessment[6]

The entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special The participatory survey and agricultural field crop pattern assessment was conducted in six (6) locations manually and had conversation with members of the selected farms. Observations were made on current and dominant land use systems, range of crops grown, varieties of crops and their cropping systems, location of specific crop fields , crop plant population and their relative times and seasons of crop planting, the dominant weed species along with major pests and diseases. Visual observation attempts were made to assess the extent of soil erosion and land degradation with poor quality of yield.

4.2 Predictive analytics approach

4.2.1. Rice Variety and Quality Uniformity test[7]

An area 25 meters by 40 meters was planted with rice variety IR8. At harvest, a 1-meter border was removed around the field and discarded. Each square meter (1 meter x 1 meter) was harvested and weighed. We were able to observe the quality degradation at the middle and the corners due to insufficient water and sunlight.

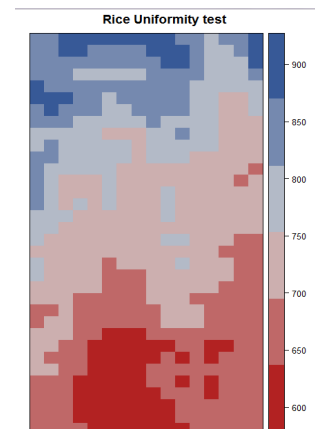


Figure.1 Rice uniformity test

4.2.2. Crop density test

We are observing on each segment, the same variety of Rice but density differs. In this observation we are able to see the crop outcome with 6 different densities.

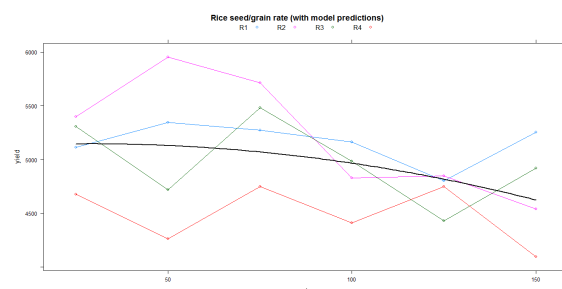


Figure.2 Rice seed/grain (with model predictions)

4.2.3. Split-plot experiment[8]

To determine the height of the rice. We made observations on split-plot experiment in three blocks. Whole-plot is 'management', sub-plot is 'time' of application, with two subsamples. The data are the heights, measured on two single-hill sampling units in each plot.

```

Error: rep
  Df Sum Sq Mean Sq
rep  2  2632    1316

Error: rep:manage
  Df Sum Sq Mean Sq F value Pr(>F)
manage  7  1482    211.77  2.239 0.0944 .
Residuals 14  1324    94.59
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: rep:manage:time
  Df Sum Sq Mean Sq F value Pr(>F)
time  3  820.8    273.61  7.945 0.000211 ***
manage:time 21  475.3    22.63  0.657 0.851793
Residuals 48 1653.1    34.44
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: within
  Df Sum Sq Mean Sq F value Pr(>F)
Residuals 96 167.4    1.744

```

Figure.3 Split-plot experiment

A strip-plot experiment with three reps, variety as the horizontal strip and nitrogen fertilizer as the vertical strip.

4.2.4. Grain yield was measured

Grain yield was measured at 3 locations with 3 reps per location. Within each rep, the main plot was 6 nitrogen fertilizer treatments and the sub plot was 2 rice varieties.

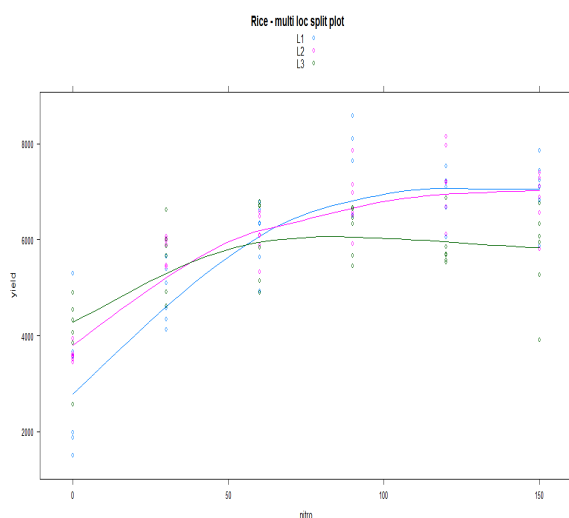


Figure.4 Rice-multi loc split plot

4.2.5. Soil nitrogen at three times for eight fertilizer treatment

Eight fertilizer treatments were tested. Soil nitrogen content was measured at three times. P1 = 15 days post transplanting. P2 = 40 days post transplanting. P3 = panicle initiation.

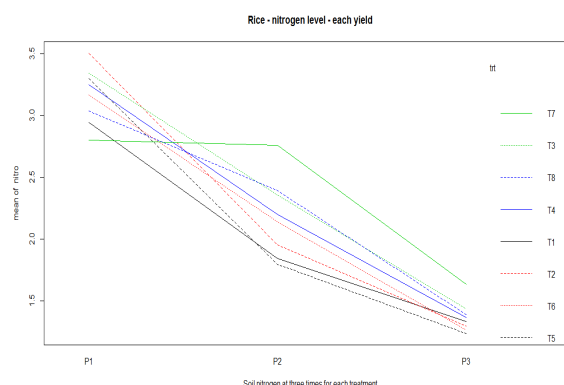


Figure.5 Soil nitrogen at three times for each treatment

5. Results and discussions

5.1.1. Long sowing period and poor timing

For “Rice” the sowing period may up to 12 weeks and the days interval between the last planting and the first harvest ranges around 30 days. This results in pest populations. The timing of the farming practices are proportional to the date of sowing that indicates that good crop management is a weak point in the current production system.

5.1.2. Intensive soil tillage

With the help of disc-plough and/or the offset-plough, dry soil tillage is carried out. The wet tillage is usually carried out with the chipper and the weedcutter attached to a tractor with cage wheels after which the fields are smoothed with a wooden beam. Four to five soil tillage could be distinguished, which is relative to their background in the soil humidity and machinery availability with respect to water and time. Fields tilled with the 3-disc plough had significantly more production with high quality commodity than other fields.

5.1.3. Late sown fields = low yields

Relatively late sown fields were resulted in low yield with moderate quality of commodity. The weight of the grain is comparatively lower indicating the stress conditions due to less irrigation water, unfavourable weather, issues due to insects and weed.

5.1.4. Soil fertility of the rice

In addition to nitrogen, phosphate appears to be low in most fields, and potash is expected only on a few

fields. Other nutrients are available plenty. The analyses show that the Rice variety with good weight and height shows, apart from nitrogen and phosphate, the irrigation methodologies have their impact and the results were homogeneous. Rice production yield differences among fields and parts of lands or soil types are therefore attributed to differences in crop management (that includes fertilization with nitrogen and phosphate).

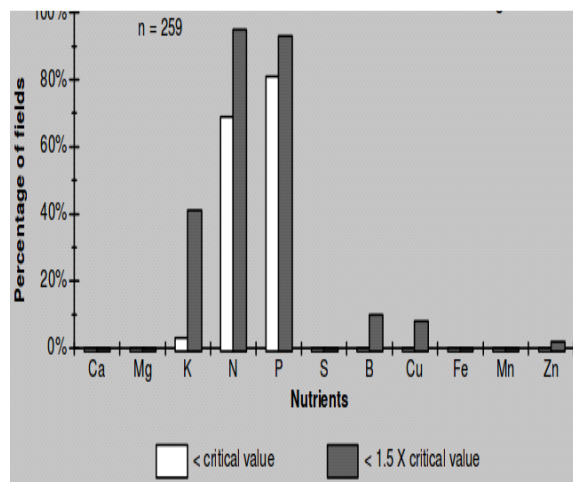


Figure.6 Soil fertility of crop nutrients

6. Conclusion

Our action-oriented methodology aims at adopting an integral ecosystem approach to achieve national, local and global benefits of predicting the outcome of the crop quality and finding out the ways for restoring degraded land and crop seeds with respect to climate change, native commodity conservation, and improved rural livelihoods.

References

- [1] Bowman, M., Debray, S. K., and Peterson, L. L. "Reasoning about naming systems", 1993
- Arulkumar C and Manimannan G (2014), Agriculture Productivity Performance of North-Western Zone in Tamilnadu Using Data Mining Techniques, International Journal of Statistika and Matematika, Volume 10, Issue 1, pp. 1-
- [1] Sjaak Wolfert, Lan Ge, Cor Verdouw, Marc-Jeroen Bogaardt, Big Data in Smart Farming – A review, Agricultural systems, Volume 153, May 2017, Pages 69–804.
- [2] Francis, C.C., and M.D. Clegg. 1990. Crop rotations in sustainable production systems. In C.A. Edwards, R. Lal, P. Madden, R.H. Miller, and G. House (Eds.) Sustainable Agricultural Systems (pp. 107-122). Soil and Water Conservation.
- [3] S Senthilvadivu, S Vinu Kiran, S Prasanna Devi, S Manivannan, Big Data Analysis on Geographical Segmentations and Resource Constrained Scheduling of Production of Agricultural Commodities for Better Yield, Vol. 87, pp. 80-85, Procedia Computer Science, 2016
- [4] <http://lib.icimod.org/record/28226/files/15.pdf>, Evaluation of the System of Rice Intensification (SRI) through participatory research and development.
- [5] M.Rajesh, K. Balasubramaniaswamy, "Open Issues in Routing Techniques in Wireless AdHoc Sensor Networks", vol. 1, no. 1, pp. 05-08, 2015.
- [6] Vijayakumar, <http://krishikosh.egranth.ac.in/bitstream/1/78633/1/1.%20INTRODUCTION.pdf>, Assessment Of Distinctiveness And Uniformity Of The Characteristics In Farmers' Rice (Oryza Sativa.) Varieties And Their Application For Pvp
- [7] http://dl.taq.ir/science/variations_on_splitplot_and_splitblock_experiment_designs_federer.pdf
- [8] S. Prasanna Devi, Y. Narahari, N. Viswanadham, SV Kiran, S. Manivannan E- mandi implementation based on gale-shapely algorithm for perishable goods supply chain, pp.1421-1426, Automation Science and Engineering (CASE), 2015, IEEE International conference.
- [9] https://en.wikipedia.org/wiki/Climate_of_Tamil_Nadu

