Prediction and Analysis of Soil properties in Guntur District, Andhra Pradesh

D.Venkata Revanth Naidu, Sridevi Sakhamuri, A.Venkat Vardhan

Abstract: A significant statistical analysis is done over Eleven Thousand and Two hundred soil data samples of thirteen villages in Guntur District of Andhra Pradesh. Useful relationship among major soil parameters was found. An analytical approach is used to assess the properties of the soil using pandas, numpy and matplotlib libraries in python. Hidden patterns among the soil properties are identified. By using Random Forest Regressor soil data of Guntur district is predicted with considerably higher accuracy.

Keywords : Soil Nutrients, Optimum fertilization rate, Yield quality, Random Forest Regressor.

I. INTRODUCTION
Assessment of crop cultivation is exceptionally fundamental for economical creation in agribusiness. The field supplements assume pivotal job in the yield just as illness reconnaissance. The consumption of soil ripeness is one of the key reasons for decline of production profitability. The major properties of soil like calcium, pH, magnesium and manganese are reducing substantially in the soil at an alarming rate. The main objective of soil management is to maintain the soil nutrients at a suitable level for cultivation[1]. After Green revolution, introduction of chemical fertilizers and pesticides have negatively impacted the soil. Running after higher yields resulted disadvantageous to soil and lower yield quality. It is accepted that the convergence of numerous supplements are emphatically or contrarily corresponded with major soil parameters. In the present setup we have done measurable examination to discover the shrouded examples of soil supplement relationship. Multivariate factual investigation is one of the adaptable devices in order to draw lines of similarity between various segments of soil[2]. Multivariate strategies are effectively implemented in a setup like this, for example, head segment investigation, sanctioned relationship analysis (PCA), factor analysis (FA), and authoritative correspondence examination (CCA) [3] accounted to be proficiently used to bring out the hidden patterns among the soil nutrients.

As Doran and Perkin stated, the nature of the soil solely depends on the environmental changes and the microorganisms breeding in the soil thus supporting the biological processes in the nature. Thus ensuring the ecological balance and topsoil quality [4].
The appraisal of soil condition is essential for feasible administration of farming. Soil condition assessment involves a holistic evaluation of data samples. In checking and safeguarding the soil nutrients SQI (Soil Quality Indices) are exceptionally helpful. Evaluating and analysing different soil parameters will be a great deal of advantage in assessing the soil.
One of the most important approaches to consider is multicollinearity of soil condition. The vast majority of the examination bunches tended to this issue utilizing the multivariate factual methods, for example, head part investigation. These multivariate examinations uncovered the bury connection of soil supplement with properties, for example, saltiness of soil, cut, mud and nearness of natural issue in soil and so on. These field qualities likewise impact dispersion of vegetation, the tree attributes, for example, size, vegetation spread, thickness and so on.[5]
For the present investigation we have chosen Tenali mandal of Andhra Pradesh, which is arranged nearly the focal point of the state. The district is placed between 15°18’0” - 16°50’0”N and 15°18’0” -16°50’0”E. The region covers a topographical region of 410km. In the present examination we chose information from thirteen grampanchayaths which are located in Guntur district. The yearly average downpour of the severe is 881 mm. The maximum and minimum temperatures recorded are 28 and 40 degrees respectively. Guntur district is considered as an administrative capital of Andhra Pradesh. River Krishna which serves as a geographical boundary separates it with Krishna district. According to the census (2011) agricultural sector constitutes to be a leading sector providing employment to majority of people in this district.[6]
According to agricultural contingency report out of 1139 hectares of geographical area only 682 hectares is used for cultivation. The net sown area in a cropping year is 597 hectares and area sown more than once is 206 hectares. When classified on the basis of irrigation modes irrigated area is 373 hectares apart from this, 223 hectares is rainfed area. The area under vegetables is 27 hectares. When classified on the basis of irrigation modes irrigated area is 373 hectares apart from this, 223 hectares is rainfed area. The area under vegetables is 27 hectares. When classified on the basis of irrigation modes irrigated area is 373 hectares apart from this, 223 hectares is rainfed area. The area under vegetables is 27 hectares.

Revised Manuscript Received on November 22, 2019

D.Venkata Revanth Naidu, Department of Electronics and Computer Engineering,Koneru Lakshmaiah Education Foundation,Vaddeswaram,AP,India. Email:revanth50045@gmail.com

Sridevi Sakhamuri, Department of Electronics and Computer Engineering,Koneru Lakshmaiah Education Foundation,Vaddeswaram,AP,India. Email:sridevisakhamuri@kluniversity.in

A.Venkat Vardhan, Department of Electronics and Computer Engineering,Koneru Lakshmaiah Education Foundation,Vaddeswaram,AP,India. Email:anumolvardhan@gmail.com
Different assortments of vegetables are likewise developed in
the region. On the other hand there is significant production of
spices, aromatic, medicinal and plantation crops in the
district. For the present experimental setup we have used
11,200 samples of soil data of Guntur district to work on
productive analytical approach. Implementing a testing model
to predict the soil data and figuring out the correlations among
different properties of the soil and statistically evaluating the
parameters of the soil around different villages using NumPy
and Pandas libraries in python.[6]

II. LITERATURE SURVEY

a) Analysing Soil Data using Data Mining Classification
Techniques - V. Rajeswari and K. Arunesh: This gives a
detail analysis on soil properties prediction and analysis
using Data mining classification techniques.
b) Analysis of Soil condition Based on pH value Using
Classification Techniques - Mrs. N. Hemageetha, Dr.
G.M. Nasira: This research gives analysis of soil data in
Salem district on the basis of pH value using Naive Bayes
concept
c) Improving the prediction accuracy of soil nutrient
classification by optimizing extreme learning machine
parameters. – M.S Suchitra , Maya L.Pai: This gives
information about prediction of village wise soil
parameter aids to reduce fertilizer waste depending on the
need of the crop cultivated.
d) Soil fertility evaluation of Chintayapalem village, Guntur
district, Andhra Pradesh – Sumathi.P: A thesis about how
soil behaves with respective to climate and soil type in
different areas of Guntur District and analysing its
properties.
e) A Statistical Analysis of Soil Fertility of Thrissur District,
Kerala – Indian Institute of Information Technology and
Management –Kerala by R.Ajith Kumar: Which is a
detailed analysis of soil properties of Thrissur district of
Kerala implementing ANOVA analysis.

III. SOIL FERTILITY ANALYSIS

The soil status of the region was assessed utilizing around
11,200 information focuses from various panchayaths. These
elements are illustrative of different yield generation
rehearses just as agro environmental zones. The present
investigation depends on the mean of the examples speaking
to various panchayaths exhibited in Table 1. We chose 14
distinctive panchayaths and the synopsis measurements are
displayed in Table 2. The relationship investigation is
appeared in Figure 1. The pH scale recommends the acidic
nature of the soil in the district, ranging from at least 6.94 to
a maximum of 7.68 with a mean of 7.45 and standard
deviation of 0.22. Considering these analytical results we can
conclude that the soil is emphatically acidic in nature. The
amount of rainfall can be accounted for such moderately
lower pH values which results in reducing the base nature of
the soil. Apart from these, Moisture content and temperature
are also responsible for change of soil pH.[7]

Measuring electrical conductivity ensures the health and
fitness of the soil with a substantial amount of salts present in
it. Optimum levels of electronic conductivity ensure effective
absorbance of water into soil. Sometimes it is suggested to
add organic waste to increase electronic conductivity[8]. The
results show that EC differs from 0.35 to 1.09 with an average
of 0.55 and standard deviation 0.18. The soil EC is beneath
ideal level for agribusiness in some villages of the district.

Electrical conductivity is important indicator of soil
health. Excess salts can be hinder plant growth and it affects
crop yield, sustainability &activity of soil microorganism.
Low rainfall regions account is higher EC than regions with
higher rainfall .Soil microorganism activity decreases with
increases in EC.[9] Residue decomposition, nitrification,
denitrification, respiration are effected by increases in EC.

The organic carbon in the district ranges from 0.29
to 1.07 with a mean of 0.57 and standard deviation of 0.21.
Where electronic carbon is an essential component for
the microorganisms in the soil to grow and decompose the
organic matter in the soil and provide required nutrients to
the soil.[10]

Nitrogen is a major component of the chemical
fertilizers used in the fields. The Nitrogen (N) value ranges
from 98.79 to 317 in Guntur District resulting a mean of
229.19 with a standard deviation of 55.48.[11] All other
parameters of soil are listed in Table 1 along with the
statistical analysis in Table 2 with respect to different villages.

Phosphorous is assessed by using three methods soil
pH , amount of organic matter and amount of phosphorous
nutrients. Phosphorous constituting to be vital nutrient of the
soil in Guntur district ranging from 18.13 to 59.32 with an
average of 26.58.[12] Phosphorous abundance can be
achieved by liming acid soils, using measures to increase
organic matter.
Therefore increasing the acid
Nitrogen concentration increases pH value decreases shows us how pH is correlated to Nitrogen. As the form of
would serve the purpose required. Thus Figure 1 clearly
In
and potassium are all primary components for plants growth.
also responsible. Monitoring Phosphates, nitrates, calcium
into details pH, moisture content and temperature of soil are
system in major driver for changes in soil pH value. Looking
restrict fertilizers usage.[13]
water by leaching .So there is a great need to regulate or
over usage excess of Nitrogen fertilizers effect the ground
usage makes it more acidic effecting the plant growth. Also
The pH value taken into account where excess of Nitrogen fertilizers effect the ground water by leaching .So there is a great need to regulate or restrict fertilizers usage.[13]

The form Nitrogen and type of Nitrogen in soil plant system in major driver for changes in soil pH value. Looking into details pH, moisture content and temperature of soil are also responsible. Monitoring Phosphates, nitrates, calcium and potassium are all primary components for plants growth. In order to check if the soil is fit or not checking EC of soil would serve the purpose required. Thus Figure 1 clearly shows us how pH is correlated to Nitrogen. As the form of Nitrogen concentration increases pH value decreases therefore increasing the acidity of soil

Table1: Mean of soil parameters for selected villages from Guntur District

<table>
<thead>
<tr>
<th>Panchayath</th>
<th>pH</th>
<th>EC</th>
<th>OC</th>
<th>N</th>
<th>P</th>
<th>K</th>
<th>S</th>
<th>Zn</th>
<th>Fe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angalkuduru</td>
<td>6.94</td>
<td>0.35</td>
<td>1.07</td>
<td>98.79</td>
<td>59.32</td>
<td>159.82</td>
<td>36.80</td>
<td>1.63</td>
<td>18.25</td>
</tr>
<tr>
<td>Burripalem</td>
<td>7.46</td>
<td>0.54</td>
<td>0.58</td>
<td>201.48</td>
<td>28.98</td>
<td>382.09</td>
<td>41</td>
<td>1.62</td>
<td>16.55</td>
</tr>
<tr>
<td>Chinavaru</td>
<td>7.68</td>
<td>0.52</td>
<td>0.45</td>
<td>268.04</td>
<td>18.13</td>
<td>569.02</td>
<td>36.92</td>
<td>1.06</td>
<td>13.21</td>
</tr>
<tr>
<td>Gudivada</td>
<td>7.65</td>
<td>0.47</td>
<td>0.38</td>
<td>260.53</td>
<td>18.46</td>
<td>382.45</td>
<td>382.45</td>
<td>0.81</td>
<td>14.70</td>
</tr>
<tr>
<td>Katevaram</td>
<td>7.67</td>
<td>0.60</td>
<td>0.43</td>
<td>256.10</td>
<td>18.37</td>
<td>434.52</td>
<td>59.40</td>
<td>1.71</td>
<td>18.07</td>
</tr>
<tr>
<td>Kokalapur</td>
<td>7.32</td>
<td>0.44</td>
<td>0.52</td>
<td>259.59</td>
<td>19.89</td>
<td>528.82</td>
<td>40.67</td>
<td>1.28</td>
<td>23.98</td>
</tr>
<tr>
<td>Nandivelugu</td>
<td>7.55</td>
<td>0.51</td>
<td>0.95</td>
<td>166.52</td>
<td>40.27</td>
<td>383.94</td>
<td>26.83</td>
<td>1.16</td>
<td>13.22</td>
</tr>
<tr>
<td>Nelapadu</td>
<td>7.52</td>
<td>0.51</td>
<td>0.62</td>
<td>212.31</td>
<td>23.53</td>
<td>446.63</td>
<td>52.53</td>
<td>1.75</td>
<td>20.02</td>
</tr>
<tr>
<td>Pedaravuru</td>
<td>7.40</td>
<td>0.40</td>
<td>0.59</td>
<td>211.95</td>
<td>34.27</td>
<td>660.13</td>
<td>35.46</td>
<td>1.68</td>
<td>10.03</td>
</tr>
<tr>
<td>Pinapadu</td>
<td>7.55</td>
<td>0.71</td>
<td>0.54</td>
<td>246.46</td>
<td>20.34</td>
<td>578.92</td>
<td>65.18</td>
<td>147.69</td>
<td>26.90</td>
</tr>
<tr>
<td>Sangam</td>
<td>7.16</td>
<td>0.38</td>
<td>0.29</td>
<td>317</td>
<td>23.73</td>
<td>501.17</td>
<td>20.08</td>
<td>1.32</td>
<td>28.53</td>
</tr>
<tr>
<td>Tenali</td>
<td>7.38</td>
<td>0.58</td>
<td>0.53</td>
<td>206.73</td>
<td>20.21</td>
<td>395.84</td>
<td>62.89</td>
<td>1.25</td>
<td>26.28</td>
</tr>
<tr>
<td>Tenali(U)</td>
<td>7.57</td>
<td>1.09</td>
<td>0.41</td>
<td>274.03</td>
<td>20.05</td>
<td>360.94</td>
<td>57.06</td>
<td>2.46</td>
<td>26.44</td>
</tr>
</tbody>
</table>

Table2: Statistical Analysis of soil Nutrients

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>7.4512</td>
<td>0.220534</td>
<td>7.525392</td>
<td>6.944776</td>
<td>7.684653</td>
<td>0.048635</td>
</tr>
<tr>
<td>EC</td>
<td>0.550496</td>
<td>0.189982</td>
<td>0.513618</td>
<td>0.352239</td>
<td>1.092152</td>
<td>0.036093</td>
</tr>
<tr>
<td>OC</td>
<td>0.570515</td>
<td>0.218801</td>
<td>0.531053</td>
<td>0.291778</td>
<td>1.075522</td>
<td>0.047874</td>
</tr>
<tr>
<td>N</td>
<td>229.1989</td>
<td>55.4872</td>
<td>246.4689</td>
<td>98.79104</td>
<td>317</td>
<td>3078.829</td>
</tr>
<tr>
<td>P</td>
<td>26.58429</td>
<td>11.95055</td>
<td>20.3445</td>
<td>18.13366</td>
<td>59.32836</td>
<td>142.8156</td>
</tr>
<tr>
<td>K</td>
<td>444.9512</td>
<td>126.5907</td>
<td>434.5227</td>
<td>159.8209</td>
<td>660.1366</td>
<td>16025.21</td>
</tr>
<tr>
<td>S</td>
<td>70.56357</td>
<td>94.74881</td>
<td>41</td>
<td>20.08899</td>
<td>382.4575</td>
<td>8977.337</td>
</tr>
<tr>
<td>Zn</td>
<td>12.72971</td>
<td>40.55466</td>
<td>1.626471</td>
<td>0.816644</td>
<td>147.6969</td>
<td>1644.681</td>
</tr>
</tbody>
</table>

The pH value taken into account where excess of fertilizers usage makes it more acidic effecting the plant growth. Also over usage excess of Nitrogen fertilizers effect the ground water by leaching .So there is a great need to regulate or restrict fertilizers usage.[13]

The form Nitrogen and type of Nitrogen in soil plant system in major driver for changes in soil pH value. Looking into details pH, moisture content and temperature of soil are also responsible. Monitoring Phosphates, nitrates, calcium and potassium are all primary components for plants growth. In order to check if the soil is fit or not checking EC of soil would serve the purpose required. Thus Figure 1 clearly shows us how pH is correlated to Nitrogen. As the form of Nitrogen concentration increases pH value decreases therefore increasing the acidity of soil

Figure1: Correlation (Relationship between pH and Nitrogen)

IV. PREDICTION REPORT

The prediction of soil properties is implemented using Random Forest Regressor in spyder IDE where we achieved 89% accuracy rate with a loss of 0.02 which is considered to be least. These prediction analysis findings are reported in the Table 3 below. As a future improvement we are planning to map the set crops with the results obtained after prediction and give an optimum result for relevant crop to be cultivated with respect to the soil requirements.

Table 3: PREDICTION ANALYSIS REPORT

<table>
<thead>
<tr>
<th>RMSE</th>
<th>ExpVar</th>
<th>MAE</th>
<th>MSE</th>
<th>MSLE</th>
<th>MedAE</th>
<th>RMS log error</th>
<th>R*R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.211</td>
<td>0.886</td>
<td>0.083</td>
<td>0.046</td>
<td>0.001</td>
<td>0.013</td>
<td>0.35</td>
<td>0.886</td>
</tr>
<tr>
<td>0.211</td>
<td>0.886</td>
<td>0.083</td>
<td>0.046</td>
<td>0.001</td>
<td>0.013</td>
<td>0.35</td>
<td>0.886</td>
</tr>
<tr>
<td>0.214</td>
<td>0.885</td>
<td>0.081</td>
<td>0.046</td>
<td>0.001</td>
<td>0.013</td>
<td>0.36</td>
<td>0.885</td>
</tr>
<tr>
<td>0.214</td>
<td>0.885</td>
<td>0.081</td>
<td>0.046</td>
<td>0.001</td>
<td>0.013</td>
<td>0.36</td>
<td>0.885</td>
</tr>
</tbody>
</table>

*RMSE =root mean square error  *MAE=Mean Absolute Error  *MSE=Mean Squared Error  *ExpVar=Exp Variance  *MSLE=Mean Squared Log Error  *MedAE=Median Absolute Error
V. CONCLUSION

Assessing the soil properties using data samples of Guntur District was performed using python in spyder IDE. The datasets are collected from 13 different panchayats in two different Cultivation years. The Analysis shows that there is a significant correlation between pH and Nitrogen parameters. A detailed Statistical Analysis is done on soil properties like pH, EC, OC, N, P, K, Fe, Zn. Apart from these an efficient predicting algorithm Random Forest Regressor is used to predict the soil properties.

REFERENCE

5. S. SriDevi*, M.Bindu Prathyusha, P.V.S.I.Krishna Teja , “ User behavior analysis on agriculture mining system”.
7. Gopi Krishna P., Ravi K. S., Nishanth C., Chowdary K.K.,reddyKU
9. DMAH P Vinay Kumar.”Soil nutrient measurement in Paddy farming Using IoT”
10. DMAH Moulan, “A frame work for efficient decision Tree using attribute Elimination and dual processing techniques”.
11. RB Badu , MA Hussain , RB babu , “Comparitive study of Algorithm on class imbalanced Datasets”
12. G.V.SaiPrassana,G.Vijay Kumar,”Controlling and Monitoring plant growth conditions using Embedded systems