Real Time CHISM Model for Efficient Sugarcane Plant Growth and Yield Estimation Model using Satellite Images

J.Saranya, N. Thenmozhi

Abstract: The research on plant growth estimation of sugarcane plants is a key factor ongoing now days. The problem of plant growth and yield estimation of sugarcane plants is well studied. There are number of solutions recommended by different researchers, still they suffer with poor accuracy. Existing methods measure the plant growth according to the rainfall and temperature which introduces poor performance. To improve the performance, an efficient Climate Hydro Image Soil Model (CHISM) is presented. The model considers various properties namely climate conditions like temperature, humidity and hydrologic features namely rainfall, water poured and soil conditions towards plant growth. The method uses the satellite images in obtaining the soil condition, by applying image processing technique, the soil condition are obtained. Remaining features are obtained through the regional data set provided by agriculture sector. Using all these features, the method estimates various influence measure on different features considered. The method computes rainfall influence measure (RIM), water influence measure (WIM), temperature influence measure (TIM), humidity influence measure (HIM), and soil influence measure (SIM). Using all these measure, the model computes the plant growth rate (PGR) and Yield Rate (YR) in different time window. According to the measures estimated, the model performs water regulation. The method improves the performance of plant growth estimation and crop yield.

Index Terms: Agriculture Industry, Satellite Images, Sugarcane Yield, Plant Growth, Water Regulation, CHISM.

1. INTRODUCTION

The agricultural industry has higher impact in the world economy as well as social mobility. The ratio of population getting increased every year and to meet the social requirement the agricultural industry has to work on that. The country like India has more population and struggle with higher scarcity in commodities like wheat, rice and so on. However, the agricultural society and industry performs cultivation of plants in the available agricultural area. In India the 68% of geographic area is covered by agricultural and forest. The growing population replaces the agricultural area by residential and industries. This reduces the yield and outcome of the entire agricultural sector. Also it is necessary to come up with some strategic solution in this problem. The growth of agricultural sector is highly depending on various factors. For example, the yield of any plant is depending on temperature, humidity, rainfall, water poured, soil type, evaporation and so on.

By considering all these features, the possible growth and the yield would be produced can be measured. By measuring the yield possible, and the growth of the plant, the necessary actions can be performed. There are number of approaches available to perform plant growth estimation and yield estimation, but they consider only limited features. For example, there are a method which uses area and rainfall in predicting the yield and plant growth. Similarly, there are number of methods available to perform plant growth estimation and yield estimation.

By considering all these issues, the data mining approach can be adapted to the problem. Similarly, the soil type and the plant growth estimation can be performed based on satellite images. The satellite images are used for several conditions and problems. The agricultural growth can also be used depend on satellite images. The satellite images are used in predicting the rainfall, predicting climate changes. So, by using them, the plant growth estimation and yield prediction can be performed. This paper discusses such approach towards yield estimation of sugarcane plants. The proposed CHISM model has various stages of plant growth estimation and yield prediction. The method maintains the rate of yield produced in different time window of any region. Such records have been used in estimating the yield and growth of plants. Similarly, the satellite images are used for two different purposes. First, from the satellite image, the soil type can be identified, and the plant growth can be performed. Second, from the satellite images of various time windows, the change in climate and the rain fall can be predicted. This would support the water regulation and can perform water management efficiently. Such works can be performed by applying image processing technique.

This CHISM model measures different influence measures towards plant growth estimation. The method considers the influence of different factors in achieving higher growth and yield of sugarcane. According to this the method estimates computes rainfall influence measure (RIM), water influence measure (WIM), temperature influence measure (TIM), humidity influence measure (HIM), and soil influence measure (SIM). Using all these measure, the model computes the plant growth rate (PGR) and Yield Rate (YR) in different time window. Among these measures, the temperature, humidity and rainfall features are measured according to the image features obtained from satellite images. The image processing techniques are used in obtaining the features. The detailed approach is presented in the next section.

Revised Manuscript Received on March 16, 2020.

J. Saranya, Department of Computer Science, Government Arts and Science College, Coimbatore, Tamil nadu, India.

Dr. N. Thenmozhi, Assistant Professor, PG and Research Department of Information Technology, Government of Computer Science, Coimbatore, Tamil nadu, India.
II. RELATED WORKS

There are number of solution have been recommended from various researchers and such methods are discussed in this section. In [1], the author present crop yield estimation framework which consider the satellite images. The method uses satellite images and obtains various factors from the image using different image processing techniques and obtained result has been used to perform crop yield estimation. In [2], the author presents a mathematical model for crop yield estimation which uses remote sensing techniques. The images captured have been used to extract various features and the mathematical model uses energy balanced equation to hike the performance of crop yield estimation. In [3], the author presents a image processing framework through android phones to estimate the yield of Kiwifruit. The method considers the area of cultivation and the number of Kiwi fruit. Using these two, the method estimates the yield of fruit. In [4], the author presents a crop yield estimation approach by applying background subtraction towards wheat. Similarly, a vision oriented approach on infection identification on plants is presented in [5], which uses color features. The image is segmented by applying k-means approach and GLCM is used of classification of disease affected by. In [6], an IoT based plant growth estimation algorithm is presented, which uses color, texture and shape features of leaves to perform pattern matching towards plant growth estimation. In [7], the author present yield estimation algorithm which uses image processing technique in Vineyard. The method cluster the area under different condition and for each cluster a set of weight is measure. The segmentation of the image is used to perform yield estimation. In [8], the author present a detailed review on various image processing methods would support the problem of crop yield estimation. The author surveys the adoption of communication facilities in the growth of predicting the crop yield. Plant Diseases Recognition Based on Image Processing Technology [9], present a combined approach of image processing and region growing methods for the yield estimation. The system uses different linear regression methods and image processing feature extraction methods. The method produces noticeable results on crop yield estimation. In [10], the crop yield estimation is approached with improved deep learning pipeline. The method optimizes the parameters of crop yield according to threshold, output size and so on. The network is trained with number of images and evaluated for its performance in various parameters. In [11], the author present a yield estimation algorithm for red macroalga Kappaphycus alvarezi which uses satellite images in Indonesia. It notices that the carrageenan yield is higher when the temperature is moderate and the salinity range is higher. The biomass value should be less for the better growth of yield. In [12], the author presents an evapotranspiration value estimation approach using remote sensing. The method captures the remote sensing data and consider the evaporation occur on water source. The method considers the ratio of water being used at green season which is obtained from the satellite maps. Based on these values, the method estimates the crop water consumption. In [13], the author presents a crop yield estimation algorithm using AI and SI. The prediction of crop yield is performed from the satellite images. From the images, the method extracts various features of temporal nature. From the images, the method extracts the features like temperature, humidity, area of cultivation, water source and so on. Using these data, the CNN algorithm is used to predict the crop yield. In [14], the problem of crop yield estimation is handled with the combination of random forest and decision tree algorithms. The method is validated for its performance using the data set obtained from Terra. Decision tree is used to generate the trained model where the random forest is used to predict the result. The method archives efficient decision which is able to produce efficient values on plant growth. In [15], a satellite image based yield estimation and classification of crop has been presented. The method extracts the color and texture features to perform image classification towards yield estimation. In [16], the author presents a crop yield estimation algorithm which uses satellite images which consider both contextual information. It works towards the criteria of fixing compensation for the farmers at huge loss. The method identifies various factors and performs classification according to the crop properties. The remote sensing images are used to extract various features to perform classification. In [17], the author performs yield estimation on Maize using satellite images of Zimbabwe. The method performs inference on yield according to the data obtained from small-scale commercial farming sector (SSCF) is used for analysis. The method has been validated with different national yield model with different data sets. In [18], the author presents a corn height estimation scheme towards yield prediction. The UAV dataset has been used for evaluation which contains information like RGB data. The same data has been used for evaluation and estimate the height of corn height estimation. In [19], the author presents a satellite image based crop yield estimation model to support fertilizer application on corn field. Similarly, in [20], the author present a satellite image based chlorophyll content estimation scheme SMLR-PSO model. The method uses spectral images and extracts various features using random field segmentation to estimate the dependencies. PSO algorithm is used to perform prediction to produce efficient results. All the methods discussed has been suffer with achieving higher yield and introduces poor accuracy.

A. CHIS Sugarcane Plant Growth and Yield Estimation Model:

The proposed CHIS (Climate Hydrology Image Soil) model uses the agricultural data obtained from different region of the country. Also, the satellite images obtained from weather research center are used to extract different features. From the data set available, the method split the data into number of time domain. Also the data has been split into number of regions space and for each space; the method estimates different influence measures. Similarly, the method preprocessed the input images and performs segmentation to extract features.
The architecture of proposed CHIS model is presented in Figure 1 which shows various functional components present. Each functional part is discussed in detail in this section.

**B. Agriculture Feature Extraction:**

The agricultural data set ADs given has been taken here for feature extraction. The data set would have trace for number of years according to different regions of the country. The data present in the data set would contain noisy or incomplete values. Such values are identified and the records with incomplete features have been eliminated from the data set to produce preprocessed data set. From the preprocessed set, the method split the records into number of time space. The records may be available for different time space like year, month and so on. Such records are split into number of groups. Further the records split into region space and for each region space, the method extract the features to be used to measure various values.

Consider the agriculture data set Ads, which contains 30 year data which represent 30 time space. First the list of features and facts available in the data set is identified as follows:

$$F_l = \sum (\text{Features} \in F_l) \cup \sum ADs(i). \text{Feature} \in F_l$$

(1)

Now the list of features present in the data set is available in fact list Fl. By using this list, the noise values are identified and such records are eliminated from the data set. Noise removed set Nrs = \sum (\text{records} \in Nrs) \cup i=1\text{size}(ADs)ADs(i)\text{Features}(Fl)$

(2)

In the above equation (2), the records covers all the features of feature list is added to the noise removed data set.

Now the list of time space present in the data set should be identified. It is identified as follows:

$$T_l = \sum (\text{time} \in T_l) \cup (\text{Nrs}(i). \text{Time} \in T_l)$$

(3)

The equation (3) identifies the unique time window present in the logs of noise removed set.

Now the records of the data set have been grouped under number of clusters according to the time space. If the size of Tl is N, then N number of clusters are generated. Now according to the time space, the method groups the records in to different clusters. Now the clustering is performed as follows:

$$\sum_{i=1}^{\text{size}(\text{Nrs})} C_s(Ts) \leftarrow (\text{Nrs}(i). Ts = Cs(Ts))$$

(4)

The equation (4) indexes the logs available in Nrs to different time stamp clusters according to the time value present in each log. Generated clusters are used to estimate various measures towards plant growth estimation.

**C. Satellite Image Feature Extraction:**

The feature from satellite image is extracted in this stage. The satellite image would have noise values due to the capturing device. It is necessary to remove the noisy pixels. It is performed by applying Gabor filter. Further, the image has been applied with Histogram equalization technique to improve the quality of image. Further, the method generates gray scale image and performs segmentation by applying gray threshold segmentation. The gray threshold segmentation is the process of grouping the pixels of image according to their gray value. From the grouped image, the method identifies the region of image which represents water particles and the soil type is detected. The soil type is detected according to the gray scale value present in the image. Extracted features are used to perform plant growth estimation and crop yield estimation.
The Figure 2 A and 2B shows different information obtained from different countries, where Figure 2.A display the crop health information of United States of America published by NASA. Similarly, Figure 2(B) shows the area being left unplanted is displayed in the country Australia. So, from these images, the features like soil type, area of cultivation, area of vacant land, area of water available can be extracted to support plant growth estimation and yield estimation.

Algorithm:

Input: Satellite Image SI
Output: Feature Vector Fv

Start
Read satellite image SI
Initialize Gabor Filter GF = \( \sum_{i=1}^{\text{No of Level}} \text{Initialize}(i, \text{coefficient}) \) (5)

The equation (5) applies the gabor filter on the input image.

For each level 1

\( \text{SI} = \text{Gabor}() \)

End

SI = Histogram Equalization (SI)
GI = Grayscale(SI)

Compute minimum gray scale \( M_{\text{gs}} = \sum_{i=1}^{\text{size(SI)}} \text{Min}(SI(i), \text{value}) \) (6)

//equation (6) identifies the minimum gray scale value available in the image.

Compute maximum gray scale \( M_{\text{xgs}} = \sum_{i=1}^{\text{size(SI)}} \text{Max}(SI(i), \text{value}) \) (7)

//equation (7) identifies the maximum gray scale value available in the image.

Compute histogram values of GI as GIHist = Histogram (GI)

Choose least 2 gray values as two.

Choose maximum gray scale value as WT.

Segmented Image Simg = Perform segmentation with WT.

Simg = Perform segmentation with Two values.

Compute Area of cultivation CA.

\( CA = \text{size(Simg)} - (\sum_{i=1}^{\text{size(Simg)}} \text{Simg}(i). \text{value} < \text{LeastSecondThreshold}) \) (8)

//equation (8) identifies the cultivation area by counting the number of pixels with the color value which is less than the least second threshold which represent the green layer and plant feature.

Soil Type S = if LeastFirstThreshold<100?1:2

Compute volume of water \( W_v \).

\( W_v = \text{size(Simg)} - (\sum_{i=1}^{\text{size(Simg)}} \text{Simg}(i). \text{value} < \text{LeastSecondThreshold}) \)

//the water features are identified with the color value greater than the least second threshold

Compute plant growth \( P_g = \frac{\text{size(Simg)}}{\text{Simg}(i). \text{Green} > \text{Threshold}} / \text{size(Simg)} \)

//the plant growth is measured according to the green pixels of the image and threshold.

Feature vector Fv = \{CA, size(Simg), S, Wv, Pg\}

Stop

The above discussed algorithm represents how the features from the satellite image have been extracted. Extracted values and features are used to measure different influence measure to perform sugarcane plant growth and yield estimation.

D. CHIS Plant Growth Estimation:

The growth of any plant being cultivated can be measured according to different factors. The proposed approach considers different factors in estimating the growth of the plant. First, the features from the satellite image are used and second, the features from the previous records are used. Also, the current state of art is considered in measuring the growth of the plant. The current satellite image is taken and the features from the image are extracted to obtain the feature vector. From the data set, the method extracts different features from the entire satellite image to obtain feature set. Similarly, the features from the cluster generated in the Agricultural feature extraction stage are extracted in terms of clusters. We merge both the values of feature extraction to measure the plant growth.

Algorithm:

Input: Agricultural Data Set Ads, Satellite Image Set SIS, Sat Image SI, Current Data CD
Output: Plant Growth Value

Start
Read Ads, SIS, SI, CD.
Cluster Set Cs = Agricultural-Feature-Extraction (Ads)
Feature Set Fs = Satellite Feature Extraction (SIS)
Feature F = Satellite Feature Extraction (SI)

Compute Time Domains \( T_d = \text{size(Cs)} //\text{Number of Clusters} \)

For each time domain Tdi

Compute Rain Fall Ratio \( R_{fr} = \)
The rain fall is measured according to the average rainfall in the time window logs.

\[
W_{fr} = \frac{\sum_{i=1}^{\text{size}(Cs(Tdi))} cs(Tdi). Rainfall}{\text{size}(Cs(Tdi))}
\]

The Wfr is measured according to the average water poured in the time stamp logs.

\[
CR = \frac{\sum_{i=1}^{\text{size}(Fs(Tdi))} Fs(Tdi). CA}{\text{size}(FS(Tdi))}
\]

The CR is measured according to the average area of cultivation in the time stamp logs.

\[
P_{gr} = \frac{\sum_{i=1}^{\text{size}(Fs(Tdi))} Fs(Tdi). P_g}{\text{size}(FS(Tdi))}
\]

P_{gr} is measured according to the average plant growth obtained in the time stamp log.

The above discussed algorithm represents how the plant growth on sugarcane cultivation and its plant has been measured. The method estimates various ratio of factors based on which the method performs plant growth estimation.

E. CHIS Crop Yield Estimation:

The crop yield of sugarcane has been measured according to different factors. The crop yield is measured based on water poured, rain fall, soil type, temperature, humidity, area of cultivation and so on. For each factor, the method computes influence measures. The rainfall influence is measured according to the ratio of rainfall and rainfall obtained in the current time window. Similarly, the temperature influence is measured according to values of temperature currently recorded and the mean value of temperature in the region. The humidity influence measure is computed accordingly.

Algorithm:

Input: Agricultural Data Set Ads, Satellite Image Set SIS, Sat Image Si, Current Data CD

Output: Yield Value

Start

Read Ads, SIS, SI, CD.

Cluster Set Cs = Agricultural-Feature-Extraction (Ads)

Feature Set Fs = Satellite Feature Extraction (SIS)

Feature F = Satellite Feature Extraction (SI)

Compute Time Domains Td = size (Cs) //Number of Clusters

For each time domain Tdi

Compute Rain Fall Ratio Rfr = \[
\frac{\sum_{i=1}^{\text{size}(Cs(Tdi))} cs(Tdi). Rainfall}{\text{size}(Cs(Tdi))}
\]

Compute Water Pour Ratio wpr = \[
\frac{\sum_{i=1}^{\text{size}(Fs(Tdi))} Fs(Tdi). Rainfall}{\text{size}(FS(Tdi))}
\]

Compute cultivation ratio CR = \[
\frac{\sum_{i=1}^{\text{size}(Fs(Tdi))} Fs(Tdi). CA}{\text{size}(FS(Tdi))}
\]

Compute plant growth ratio Pgr = \[
\frac{\sum_{i=1}^{\text{size}(Fs(Tdi))} Fs(Tdi). P_g}{\text{size}(FS(Tdi))}
\]

Stop
Compute Rain fall influence measure RIM = 
\[ \text{Dist}\left( \left( \sum_{i=1}^{\text{size}(C_t)} \frac{R_f i}{T_d} \right), CD.Rainfall \right) \]
\[(19)\]
//distance between the average rainfall and the rainfall obtained in the input feature given

Compute Water Pour influence Measure WIM.

\[ \text{WIM} = \text{Dist}\left( \left( \sum_{i=1}^{\text{size}(C_t)} \frac{W p r i}{T_d} \right), CD.waterpoured \right) \]
\[(20)\]
//distance between the volume of water poured in the current input sample, average water poured in the logs.

Compute Temperature Influence Measure TIM = \text{Dist}\left( \left( \sum_{i=1}^{\text{size}(C_t)} \frac{\text{Temp}_i}{T_d} \right), CD.Temp \right)
\[(21)\]
//distance between the temperature in current time and the average temperature.

Compute Humidity Influence Measure HIM = \text{Dist}\left( \left( \sum_{i=1}^{\text{size}(C_t)} \frac{\text{Hum}_i}{T_d} \right), CD.Humidity \right)
\[(22)\]
//distance between the average humidity and the humidity monitored currently.

Compute yield value \( Yv = \frac{\text{RIM}}{\text{HIM}} \times \frac{\text{TIM}}{\text{WIM}} \times Yr \)
\[(23)\]

Stop

The above discussed algorithm represents how the yield has been estimated based on the various influence measure computed.

**F. CHIR Water Regulation:**

The water regulation becomes more important because of the shortfall of rain which is being increased every year. To achieve higher plant growth and yield, it is necessary to regulate the required volume of water to the plants. It is measured according to the rainfall, temperature and humidity of specific area of cultivation. According to the value of different factors, the specific volume of water has been regulated to the plant area.

Input: Agricultural Data Set Ads, Satellite Image Set SIS, Sat Image SI, Current Data CD

Output: Yield Value

Start

Read Ads, SIS, SI, CD.

Cluster Set Cs = Agricultural-Feature-Extraction (Ads)

Feature Set Fs = Satellite Feature Extraction (SIS)

Feature F = Satellite Feature Extraction (SI)

Compute Time Domains Td = size(Cs) //Number of Clusters

For each time domain Tdi

Compute Rain Fall Ratio Rfr = \[ \int_{i=1}^{\text{size}(Cs(Tdi))} \frac{\text{Rf}(i)}{\text{size}(Cs(Tdi))} \]

Compute Water Poured Ratio wpr = \[ \int_{i=1}^{\text{size}(Fs(Tdi))} \frac{\text{Wp}(i)}{\text{size}(Fs(Tdi))} \]

Compute cultivation ratio CR = \[ \int_{i=1}^{\text{size}(Fs(Tdi))} \frac{\text{C}(i)}{\text{size}(Fs(Tdi))} \]

Compute Humidity Ratio Hr = \[ \int_{i=1}^{\text{size}(Fs(Tdi))} \frac{\text{H}(i)}{\text{size}(Fs(Tdi))} \]

Compute Temperature Ratio Tr = \[ \int_{i=1}^{\text{size}(Fs(Tdi))} \frac{\text{T}(i)}{\text{size}(Fs(Tdi))} \]

Compute Yield Ratio Yr = \[ \int_{i=1}^{\text{size}(Fs(Tdi))} \frac{\text{Y}(i)}{\text{size}(Fs(Tdi))} \]

End

Water Regulation value = \[ \frac{\text{Rfr}+\text{wpr} \times \text{Cr} \times \text{Tr} \times \text{Hr}}{\text{CD}(\text{R})} \]

Stop

The above discussed algorithm represents the volume of water to be regulated to the cultivation area, which is measured according to different factors.

**III. RESULTS AND DISCUSSION**

The proposed CHIS model has been implemented using Matlab and evaluated for its performance under various factors. The method has been evaluated for its performance using different data set. The agricultural data has been collected from Coimbatore region maintained by agricultural department of India. Similarly, the satellite images are obtained from the same. Using the both data set, the method evaluates the performance of proposed algorithm.
The details of evaluation being used for the performance measure have been presented in Table 1. According to the values of Table 1, the evaluation is performed under different parameters and the result obtained has been compared with the result of other methods.

### Table 1: Details of evaluation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool used</td>
<td>Matlab</td>
</tr>
<tr>
<td>Total data</td>
<td>5 years</td>
</tr>
<tr>
<td>Data source</td>
<td>ARI India</td>
</tr>
<tr>
<td>Data types</td>
<td>Numeric and Images</td>
</tr>
</tbody>
</table>

The performance on plant growth estimation, crop yield estimation and water regulation has been measured for various methods and presented in Table 2. The proposed CHIS model has produced higher performance in all the parameters considered.

### Table 2: Performance on various parameters

<table>
<thead>
<tr>
<th></th>
<th>Plant Growth Estimation</th>
<th>Crop Yield Estimation</th>
<th>Water Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepLearning</td>
<td>76</td>
<td>68</td>
<td>43</td>
</tr>
<tr>
<td>CNN</td>
<td>79</td>
<td>74</td>
<td>56</td>
</tr>
<tr>
<td>SMLR-PSO</td>
<td>84</td>
<td>79</td>
<td>67</td>
</tr>
<tr>
<td>CHIS</td>
<td>95</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>

The performance on plant growth estimation is measured between different methods and plotted in Figure 3. The proposed CHIS model has achieved higher performance than other methods.

The performance on crop yield has been measured and compared with the result of other methods. The proposed CHIS model has produced higher crop yield than other methods.

The performance on water regulation has been measured between different methods. Obtained results are compared with each method in Figure 5. The proposed CHIS model has produced higher performance than other methods.

### IV. CONCLUSION

In this paper, an efficient CHIS model is presented to perform plant growth and yield estimation. The model maintains the traces of previous agriculture growth and yield obtained from different geographic area of India. In particular, the records from Coimbatore region has been collected and grouped under different time stamp. In the prediction and estimation, the clustered data has been used. Using these data, the method computes different influence measures like Temperature influence measure (TIM), Humidity Influence Measure (HIM), Water Influence Measure (WIM), Rainfall Influence Measure (RIM) and etc. Using these measures, the method estimates the possible yield and estimate the growth of the sugarcane. Similarly, using the details, the method estimates the volume of water to be regulated to the plants. The adaption of CHIS model has support the performance development in estimating the crop yield, plant growth and water regulation. The proposed method improves the performance in growth estimation and yield estimation than other methods.
REFERENCES

10. Weilu Li, Automatic Localization and Count of Agricultural Crop Pests Based on an Improved Deep Learning Pipeline, (SCIENTIFIC REPORTS), Number 7024, 2019.

AUTHORS PROFILE

J. SARANYA, Research Scholar, Department Of Computer Science, Government Arts College, Coimbatore, Tamil Nadu, India. saranyaresearch2016@gmail.com

Dr. N. Thenmozhi, Assistant Professor, PG and Research Department of Information Technology, Government Arts College, Coimbatore, Tamil Nadu, India. nthenmozhi300@gmail.com