A PREVIEW OF NATURAL VEGETATION AND PEDODIVERSITY IN INDIA ESTIMATED USING DATA MINING TECHNIQUES

Mrs.R.Vidhu[1] *Dr.S.Niraimathi*[2]

ABSTRACT

Our country is rich in Natural resources and Agriculture is the backbone of India. Due to the exploitation of mankind for development in various fields these resources get depleted. At the same time, the natural vegetation also called Virgin vegetation grows naturally without human aid. The vegetation might be classified based on the region, climate and soil where there is such growth. Soil is essential for the growth of any plant and its formation is called pedogenosis. Natural vegetation is strongly influenced by climatic changes such as atmospheric condition, ocean, sea and vegetation which form the Earth System Model (ESM). The soil is classified based on the different characteristics and soil samples are considered by index properties such as plasticity, consistency, toughness etc. Chemical substances like hydrogen pH,-sodium measured by Exchangeable Sodium Percentage (ESP) are also considered. This study reveals how the Data Mining (DM) techniques are incorporated to model the significance of vegetation. The agricultural products that affect the country's social, environmental and economic factors need to be evaluated. Data mining is a procedure of separating/comparing important information from an extensive database utilizing diverse tools and methods, and extracting only the vital data and outlining them into helpful data form. The DM algorithms and their specific approaches are used to assess vegetation and soil using vegetation maps and their related datasets.

KEYWORDS : Plasticity, Consistency,_ Earth System Models (ESM), _Data Mining (DM), Power of Hydrogen (pH), _Exchangeable Sodium Percentage (ESP).

[1]*Research Scholar*
NGM College
Pollachi.

[2]*Associate Professor*
PG Department of Computer Applications, NGM College
Pollachi.

1. INTRODUCTION

India being the seventh largest country in the world in terms of area is rich in its natural wealth. The natural resources and forests are depleted for the sustenance of growing population. The deterioration of nature will have a devastating effect on human health. The Natural Vegetation (NV), grows naturally is left undisturbed by mankind for a long time and can be preserved using Data Mining (DM) techniques. Vegetation is dependent on climate, humidity, rainfall and soil. Pedodiversity is a part of natural heritage and involves a study of the soil. Soil is important for plants to grow and the vegetation plays a crucial ecological role in supplying oxygen, reducing air pollution, controlling soil and water erosion, mitigating floods and conserving wildlife. DM is a method of selecting important knowledge from the extensive datasets by utilizing diverse tools and methods. DM is a way of assessing data from various standpoints, extracting only the vital data and outlining them into helpful data form. The DM algorithms and techniques use both statistical and numerical methods to predict the relationships among the given input variables. The prediction of data is deriving the relationship for a known object and the object to be known in future. It models the continuous valued function. The property of soil which has high impact on NV is discussed using DM techniques.

2. DATA MINING PURPOSES

DM include categorizing data, grouping them based on similar properties, finding relationships among the datasets, sequencing data based on time intervals and analyzing of outlier data.

(i) Classification is the process of finding a set of models or functions that describes and distinguishes data classes or concepts. It is done to label each kind of data with a distinct number of classes.

(ii)Cluster analysis for objects is to analyze the characteristics of each data in cluster to define subclasses for data that have similar characteristics.

(iii)Association analysis is the discovery of association rules displaying attribute-value conditions that frequently occur. It defines the probability of coincidence of data

(iv)Time series analysis comprises methods and techniques for analyzing time series data to extract meaningful statistics and helps in identifying the shapes of data from previous data gathered.

(v) Outlier analysis describes and models regularities or trends for objects which are significantly different from the remaining data.

2.1 Consequences of Data mining

Figure 1: Consequence of Data Mining

i. Competent issues of DM algorithms – The extraction of data found in larger database needs the DM algorithms to be more efficient and scalable to reduce cost.

ii. Partitioning of mining algorithms – Huge data are divided into partitions for easy references and the fragmented data is paralleled and the results from partition are merged.

iii. Dealing with similar and convoluted data types - the complex data can be of any type such as multimedia objects, temporal or spatial data objects. The variety of data found together is difficult to be separated and mined.

iv. Extraction of idea and information from heterogeneous databases and global information systems– the information of datasets available at various resources like internet and other networks are not grouped or structured. This increases the challenges in DM performance.

v. Machine learning- This method has been developing in a very large manner which deals with huge dataset and is used in processing spatial as well as spectral data. Big Data Analytics is dealt with it.

vi. Deep learning-it is an emerging discipline that has already been transformed the way the data are analyzed in many fields and it makes a great difference in the modeled way.

2.2. Classification of Vegetation

2.2.1. Tropical Rain Forests

· The tropical rain forests are also called tropical evergreen forests and found in the region where rainfall is more than 100 cm.

· Sunlight may not reach the ground due to frequent rainfall.

· This forest covers more than seven percent of total land area on earth.

· This forest has more than fifty percent of flora and fauna.

· This vegetation is rich with all kinds of plants, trees, herbs, shrubs and creepers.

2.2.2. Tropical monsoon Forests

· They are found all over our India.

· These forests are also called deciduous forests receiving several hundred centimeters of rainfall per year.

· These forests remain dry for a part of a year.

· They are categorized as dry deciduous and moist deciduous based on the amount of rainfall.

2.2.3. Tropical thorn forest

· The Thorn Forests are shrub-like vegetation found in regions with rainfall ranging from 250 to 500 millimeters.

· The Thorn forests and Scrubs are found in regions where rainfall is less than 70cm.

· The vegetation in this forest comprises thorny trees and bushes.

· Such forests have plants with long roots so that they can move far in search of nutrients in soil.

· Another characteristic feature is presence of sharp thorns to protect themselves from animals.

2.2.4. Mangrove Forests

· These forests are found near the coastlines and are flooded with sea water during high tide.

· These forests have trees to withstand high wind and salt tolerance.

· It is found between land and sea.

· Vivipara is usual.

· It is a less dense evergreen Forest.

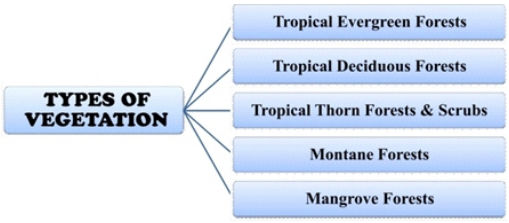

Figure 2 : Types of Vegetation

Table 1.Classification of Forests

Number	Shorthand name	Description
1	tropical evergreen forest	tropical broadleaved evergreen trees
2	temperate forest	temperate deciduous trees
3	mixed forest	temperate/boreal broadleaved and needleleaved trees
4	boreal evergreen forest	boreal needleleaved evergreen trees
5	boreal seasonal forest	boreal needleleaved deciduous trees
6	savanna	tropical broadleaved raingreen trees with grass groundcover
7	grasslands	temperate grass groundcover only
8	shrubland	tropical broaleaved shrubs with grass groundcover
9	semi-desert	tropical/temperate broadleaved shrubs with bare soil
10	tundra	boreal/alpine dwarf trees and shrubs with grass groundcover
11	desert	bare soil
13	tropical seasonal forest	tropical broadleaved raingreen trees
20	ice	permanent ice

3. Pedodiversity

Pedodiversity is the way of preserving soil or reconstructing the earth surface. It is closely related to the term biodiversity because the soil is rich in microorganisms which provide nutrients. Pedodiversity is a measure of soil variation, and various functions involved in the formation of soil. Soils information is a collected by samples and tested under various DM techniques. The texture of soil and its properties are analyzed. The Spectroscopic technology uses the visible and near infrared (vis-NIR) range and is very effective and rapid in measuring the soil properties. The soil Spectroscopy has quantitative information on Precision agriculture and vegetation.

3.1 Soil classification rules

There are various methods to classify soil and identical properties of spectral curves found are classified by Spectral Classification as a single group. Data Mining approaches such clustering, k-means and linear discriminant analysis are mostly used. The classification of soil is complex, and multivariate data analysis is used to establish their similarities. Data Mining algorithms like Multiple Linear Regression, Artificial Neural Networks, Naïve Bayesian Networks, K-nearest neighbor, Fuzzy C-means algorithms, GA Tree and Random Forests are used to classify soil texture and their properties.

3.2 Soil classification based on Indian standard method

Soil forms the surface of land and is called the Skin of the Earth. It is a complex mixture of organic and inorganic wastes, minerals, manures, water, numerous organisms and decay of living organisms. The flora is dependent on soil and plant life is essential part of the earth. The soil is classified based on the rainfall and its region. It is classified based on characteristics such as acidic properties like strongly acidic, highly acidic, moderately acidic and slightly acidic. The Indian Standard Method classifies soils based on dimension such as fine, coarse and medium or sand, slit or clay. Soil is a home of Wildlife and producer of plants. Apart from also the soil there are boulders, which are large-sized rock fragments. Bentonite is rich with highly clay minerals and

Cobbles are large-sized particles. Clay is the chemical decomposition of rocks and contains microscopic particles. Gravel is a coarse grained soil. Dunes are wind transported soils. Silt is fine-grained and not visible to the naked eye. The soil in India is classified based on the physical structure, geology, chemical composition, fertility and relief.

Table 2: Size of grain in different kinds of soil

Very coarse type soils	Boulder particle size		Greater than30 cm
	Cobble particle size		8 to 30 cm
Coarse type soils	Gravel particle size (G)	Coarse	2 to 8 cm
		Fine Soil	0. 475 to 2 cm
	Sand particle size (S)	Coarse Soil	0.2 to 0.475 cm
		Medium Soil	0.0425 to 0.2 cm
		Fine Soil	0.0075 - 0.0425 cm
Fine type soils	Silt particle size (M)		0.0002 to 0.0075 cm
	Clay particle size (C)		Less than 0.0002 cm

4. Spatial Data Mining

The Spatial features in the form of geo databases collected from different sources in various formats are stored in Geographical Information System (GIS).The new technology for spatial data analysis is Spatial Data Mining. The satellite images stored are used to monitor soil properties and the presence of Natural vegetation .The information from geodatabase is useful for predicting the future. Natural Vegetation using computational algorithms. The extraction of spatial data and use of computational applications through DM algorithms prove to be efficient for preserving Pedodiversity and Natural Vegetation.

4.1 Environmental Performance

Nature is destroyed for the survival of mankind and the result is the sudden occurrence of natural disasters. The climatic changes and rainfall are not periodic according to the seasons. Soil erosion during storms cannot be prevented and by using advanced DM techniques like SDM, GIS datasets, Remote Sensing and Internet of Things (IoT) the natural disasters can be well predicted. The geocomputational advancement in acquiring and analyzing data has paved way to meet the statistical as well as analytical approaches in a wide range of vegetation studies. The plants found in forests are rich in medicinal values. The medicinal plants if identified correctly would offer remedy for various kinds of diseases. A survey shows that more than 30% of drugs are obtained from plants. The multi-criteria decision analysis process called Analytical Hierarchy Process (AHP) covers both qualitative and quantitative analysis of complex data. The aspects selected to preserve vegetation and nature are:

i) GWP: The potency of greenhouse gases and that of carbon dioxide gases are compared to measure the greenhouse effect.

ii) AP: It measures the acidity of soil due to chemical emissions derived from gases like sulphur-methane and other nitrogenous compounds like ammonia.

5. CONCLUSION

Our Data Mining models is used to calculate the cover of earth i.e. the soil gradients depending on the datasets available and the approaches of DM helps to predict the natural disasters in advance. Data

Mining, through better management and data analysis, can assist Natural Vegetation to be preserved for future surveillance. DM offers useful tools for analyzing Pedodiversity and Natural Vegetation. The output received from various DM approaches would help the future work in modeling the texture of soil and predicting Natural

Vegetation with respect to natural disasters, climatic changes and the effect on human beings. The quality of soil depends on the composition of its chemical, biological, physical and various other nutrient compounds depending on the region. The soil health and environmental quality identified by various DM algorithms help technicians and pedologists in a better way. The emerging DM techniques like Deep Learning and Machine learning would provide improvement in predicting the future. New and improved information gathered from advanced approaches is needed for the welfare of managing the types of Vegetation and planning for a better welfare. By the use of data mining techniques, the acquired knowledge is used to make successful decisions which will preserve nature. Every individual has the responsibility to preserve natural resources for the human to live on earth.

.

References

1. Zahm F, Viaux P, Vilain L, Girardin P, Mouchet., C. Assessing farm sustainability with the IDEA method– from the concept of agriculture sustainability to case studies on farms. Sustain Dev 2008;16:271–81.

2. Meul M, Passel S, Nevens F, Dessein J, Rogge E, Mulier A, Hauwermeiren A. MOTIFS: a monitoring tool for integrated farm sustainability. Agron Sustain Dev 2008;28:321–32.

3. A. Mucherino, P. Papajorgji, P.M. Pardalos, Data Mining in Agriculture, Springer,2009.

4. Bezdek J C. 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press

5. Brown D J, Shepherd K D, Walsh M G, et al. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma, 132:273–290

6. A Novel Approach For Classification Of Soil And Crop Prediction Vrushal Milan Dolas, Prof. Uday Joshi, IJCSMC, Vol. 7, Issue. 3, March 2018, Pg.20 – 24

7. Monali Paul, Santosh K. Vishwakarma and Ashok Verma "Analysis of Soil Behaviour and Prediction ofCrop Yield using Data Mining Approach" 2015 International Conference on Computational Intelligenceand Communication Networks.

8. G. Bartzas, K. Komnitsas, An integrated multi-criteria analysis for assessing sustainability of agricultural production at regional level, Information Processing in Agriculture (2019),

9. Michele De Sanctis

10. ElizabethKearsley

11. Morellos, A., Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosystems Engineering 152, 104–116.

12. McCarty, G., Reeves, J., Reeves, V., Follett, R., Kimble, J., 2002. Mid-infrared and nearinfrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci Soc Am J 66 (2), 640–646.

13. Viscarra Rossel, R., Behrens, T., Ben-Dor, E., Brown, D., Demattê, J., Shepherd, K., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., et al., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews 155, 198–230.

14. Potapov, Laestadius & Minnemeyer (2011) generated a global map of potential forest cover at one km resolution

15. Levavasseur et al. (2012) and Tian et al. (2016) predict global PNV classes using environmental covariates such as climatic images and landform parameters. Griscom et al. (2017) recently produced a global reforestation map at one km resolution.

16. A common limitation of existing maps is their coarse spatial resolution (about 25 km)

limiting the use of these maps for operational planning (Marchant et al., 2009;

Levavasseur et al., 2012; Tian et al., 2016).

17. Bigelow NH, Brubaker LB, Edwards ME, Harrison SP, Prentice IC, Anderson PM, Andreev AA, Bartlein PJ, Christensen TR, Cramer W, Kaplan JO, Lozhkin AV,

Matveyeva NV, Murray DF, McGuire AD, Razzhivin VY, Ritchie JC, Smith B, Walker DA, Gajewski K, Wolf V, Holmqvist BH, Igarashi Y, Kremenetskii K, Paus A, Pisaric MFJ, Volkova VS. 2003. Climate change and Arctic ecosystems: 1. vegetation changes north of 55 N between the lastglacial maximum, mid-Holocene, and present. Journal of Geophysical Research: Atmospheres 108(D19):8170 DOI 10.1029/2002jd002558.

18. Extreme Learning Machines. http://www.extreme-learningmachines.org.

19.www.Tutorialspoint

20. Indian Government. Soil Health Card data.

21. Sehgal JL, Mandal DK, Mandal C, Vadivelu S. Agro-ecological regions of India. NBSS Publication; 1990. p. 24.

22. Soil Fertility Assessment and Information Management for Enhancing Crop Productivity. In: Rajasekharan P, Nair KM,Rajasree G, Sureshkumar P, Narayanankutty MC, editors.Agro-ecology of Kerala. Kerala State Planning Board; 2013. p.54–71.