

Desertification characterization using predictive soil modelling and pattern recognition

by

VIRAL A. DAVE
201721015

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

to

DHIRUBHAI AMBANI INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGY



August, 2023

Dedicated to my Father

Ashokkumar R. Dave

Declaration

I hereby declare that

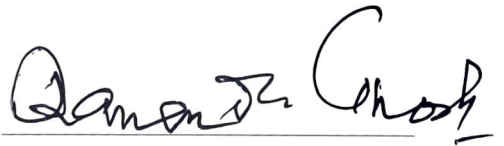
- i) the thesis comprises of my original work towards the degree of Doctor of Philosophy at Dhirubhai Ambani Institute of Information and Communication Technology and has not been submitted elsewhere for a degree,
- ii) due acknowledgment has been made in the text to all the reference material used.



Viral A. Dave

Certificate

This is to certify that the thesis work entitled DESERTIFICATION CHARACTERIZATION USING PREDICTIVE SOIL MODELLING AND PATTERN RECOGNITION has been carried out by VIRAL A. DAVE for the degree of Doctor of Philosophy at *Dhirubhai Ambani Institute of Information and Communication Technology* under my supervision.



Prof. Ranendu Ghosh
Thesis Supervisor

Acknowledgments

First and foremost, praises and thanks to God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I would like to express my sincere gratitude to Dr. Ranendu Ghosh, my research advisor, for his invaluable contributions to this work. His unwavering encouragement, guidance, and motivation have been instrumental in the completion of this research. Dr. Ghosh has always been accessible and approachable, offering expert scientific assistance and valuable insights throughout my Ph.D. journey. I am truly grateful for his technical expertise and his skills in effective problem-solving, presentation, and documentation. Working under his guidance has been a privilege, and I have learned a great deal from his wealth of knowledge in the fields of remote sensing and GIS. His positive attitude and unwavering support have been a constant source of inspiration, propelling me toward the successful completion of this thesis. I acknowledge that any mistakes or shortcomings in this work are entirely my own. Once again, I extend my heartfelt thanks to Dr. Ranendu Ghosh for his extraordinary patience and motivational guidance, which has played a pivotal role in bringing me across the finish line.

I am extremely thankful to my RPS committee members, Dr. A.S. Rajawat, Space Applications Centre (SAC), ISRO, Ahmedabad and Dr. Suman Mitra, Prof. Dhirubhai Ambani Institute of Information and communication technology, Gandhinagar for being constructive scientific evaluators and providing meaningful suggestions for the improvement of my research during the RPC meetings.

I am extremely grateful to Director, Dr. K. S. Dasgupta, for allowing me to register as a Ph.D. student and for all the support to carry out my Ph.D. study.

I am deeply thankful to Swati Priya, Megha Pandya, Abhijeet Ghodgaonkar

and Srikumar Sastry, coresearchers at Geospatial Lab, DA-IICT, Gandhinagar for their constant encouragement and support during this research work.

Last but not least, I would like to express my deepest gratitude and heartfelt appreciation to my beloved family for their support and encouragement throughout my research journey.

Contents

| | |
|--|-------------|
| Abstract | viii |
| List of Acronyms | xi |
| List of Tables | xii |
| List of Figures | xiv |
| 1 Introduction | 1 |
| 1.1 Causes of desertification & land degradation | 4 |
| 1.2 Indicators of desertification and land degradation | 5 |
| 1.3 Related Work | 6 |
| 1.3.1 Concept of desertification & land degradation | 6 |
| 1.3.2 Land degradation assessment on global scale | 8 |
| 1.3.3 Land degradation mapping in the context of India | 14 |
| 1.4 Importance of Remote Sensing and Machine Learning | 17 |
| 1.5 The motivation of the work | 19 |
| 1.6 Objectives and Accomplishments of the Thesis | 20 |
| 1.7 Organization of Thesis Chapters | 22 |
| 2 Pattern recognition of Desertification process | 26 |
| 2.1 Study Area | 29 |
| 2.2 Dataset | 31 |
| 2.3 Methodology | 32 |
| 2.4 Results & Discussion | 47 |

| | | |
|----------|---|------------|
| 3 | Predictive Soil modeling | 53 |
| 3.1 | Study Area | 56 |
| 3.2 | Soil Sampling | 56 |
| 3.3 | Covariates selection & preparation | 57 |
| 3.4 | Methodology | 59 |
| 3.5 | Result & Discussion | 66 |
| 4 | Desertification Vulnerability Assessment | 76 |
| 4.1 | Study Area | 78 |
| 4.2 | Datasets | 79 |
| 4.3 | Methodology | 80 |
| 4.3.1 | Medalus Model | 81 |
| 4.3.2 | Fuzzy logic approach | 86 |
| 4.4 | Results & Discussion | 88 |
| 4.4.1 | Climate Index (CI) | 88 |
| 4.4.2 | Soil Index (SI) | 89 |
| 4.4.3 | Elevation Index(EI) Slope Index | 90 |
| 4.4.4 | Vegetation Index (VI) | 91 |
| 4.4.5 | Socio-Economic Index (SEI) | 92 |
| 4.4.6 | Land Utilization Index (LUI) | 93 |
| 4.4.7 | Desertification Vulnerability Index (DVI) | 93 |
| 5 | Desertification hot-spot using Aridity Index | 100 |
| 5.1 | Datasets | 102 |
| 5.2 | Methodology | 103 |
| 5.3 | Results & Discussion | 104 |
| 6 | Conclusion | 108 |
| 6.1 | Key Findings and Implications of the thesis | 111 |
| 6.2 | Future Work | 112 |
| | References | 115 |
| | List of Publications | 137 |

| | |
|--|------------|
| Appendix A Desertification Status classification using Gabor filters | 139 |
| A.1 Abstract | 139 |
| A.2 Introduction | 139 |
| A.3 Methodology | 140 |
| A.4 Result & Discussion | 141 |
| A.5 Conclusion | 141 |
| | |
| Appendix B Action plan generation for combating land degradation at micro-watershed level | 143 |
| B.1 Methodology | 144 |

Abstract

This thesis presents a hierarchical methodology for land degradation mapping, land use land cover classification, degradation process identification and mapping using multispectral LISS-3 images. The study aims to demonstrate the importance of remote-sensing images for various applications, both social and environmental. The study compares the results of different algorithms for different terrains, demonstrating that Simple Linear Iterative Clustering (SLIC) segmentation with the random forest(RF) method outperforms CNN and pixel-based Support Vector Machine (SVM) with an accuracy of 85% for level 1 land cover classification. Vegetation degradation in forest areas is assessed in central parts of Gujarat, India, and land degradation in agricultural areas due to soil salinity is studied, particularly in southeastern parts of Gujarat, India. ML algorithms like support vector machine(SVM) and RF was applied to different features to identify the degradation process. Temporal data were used to find the severity of desertification using the change in degraded areas.

Further, it discusses soil degradation causing desertification and severely reducing potential soil productivity. The study uses machine learning algorithms and an ANN-based model to predict soil properties like EC, pH, and OC, which are important indicators of soil degradation. Environmental parameters are taken as covariates in prediction models, including vegetation indices, terrain indices, soil parameters, spatial attributes, and meteorological parameters of the study region. Field soil sampling data of the study region obtained from Soil Health Card (SHC) for the year 2014 is incorporated in training the model. The SHC data is divided into different ratios for training and testing the model. The SCORPAN model is considered the base approach for the development of the ANN-based

prediction model. Moreover, the thesis also discusses the mapping of vulnerable areas to desertification. The study combines remote sensing and geographic information system (GIS) to map sensitive areas. Two different approaches were used for vulnerability assessment: Mediterranean Desertification and Land Use (MEDALUS) approach and the fuzzy logic (FL) method. Soil, climate, land utilization, geography, and vegetation contribute to the land degradation of any area. However, man's intervention leads to significant changes in the environment, making socio-economic factors a considerable input to assess desertification vulnerability. Indices related to these factors are generated, and both methods are used to find the severity level of the desertification vulnerability in the Panchmahal district.

Lastly, the role of climate in the process of desertification is discussed. The study uses the aridity index (AI), which incorporates most of the weather data like temperature, rainfall, humidity, wind, and solar radiation, to identify the desertification hot-spot using AI over the Gujarat state. The study uses weather data from more than 18 locations all over Gujarat for the past 20 years to calculate AI, and the FAO Penman-Monteith method was used to calculate PET. The study generates an annual AI map for the whole of Gujarat using these values and compares it with a globally published AI map. It also compares the change in climate with the change in vegetation over the years using the vegetation index for Gujarat. In summary, this thesis provides a comprehensive approach to land degradation mapping using degradation process identification, soil prediction, and climate variable using geospatial technology and machine learning. The study demonstrates the importance of remote sensing images in various applications, including social and environmental. The study employs different machine learning algorithms and approaches to achieve high accuracy and identify vulnerable areas to desertification. The study also highlights the importance of soil properties and climate in the process of desertification.

List of Acronyms

| | |
|-----------|--|
| ASSOD | The Assessment of Human-Induced Soil Degradation in South and Southeast Asia |
| CGIAR-CSI | Consultative Group on International Agricultural Research - Consortium for Spatial Information |
| CNN | Convolutional Neural Network |
| DSM | Desertification status mapping |
| DVI | Desertification Vulnerability Index |
| FAO | Food and Agriculture Organization |
| FL | Fuzzy Logic |
| GIS | Geographic Information System |
| GLADA | Global Land Degradation Assessment in Dry Lands |
| GLADIS | Global Land Degradation Information System |
| GLASOD | Global Assessment of Human Induced Soil Degradation |
| IPBES | Intergovernmental Panel on Biodiversity and Ecosystem Services |
| ISRIC | International Soil Reference and Information Centre |
| ISRO | Indian Space Research Organization |
| LADA | Land Degradation Assessment in Dry Lands |

| | |
|-----------|--|
| LD | Land Degradation |
| LISS | Linear Imaging and Self Scanning Sensor |
| LULC | Land Use Land Cover |
| MEDALUS | The Mediterranean Desertification and Land Use |
| ML | Machine Learning |
| MoEF & CC | Ministry of Environment, Forestry & Climate Change |
| NAP | National Action Programmes |
| NBSS-LUP | National Bureau of Soil Survey and Land Use Planning |
| NDVI | Normalized difference vegetation index |
| NPP | Net Primary Productivity |
| PAGE | Pilot Assessment of Global Ecosystems |
| PSM | Predictive Soil Modeling |
| RAP | Regional Action Programme |
| RF | Random Forest |
| RFE | Recursive feature elimination |
| RS | Remote Sensing |
| SAC | Space Applications Centre |
| SCEP | Stockholm Conference on Environmental Problems |
| SLIC | Simple Linear Iterative Clustering |
| SVM | Support Vector Machine |
| TGA | Total Geographic Area |
| TPN | Thematic Program Network |

| | |
|--------|---|
| UNCCD | United Nations Convention to Combat Desertification |
| UNDP | United Nations Development Programme |
| UNEP | United Nations Environment Programme |
| UNESCO | United Nations Educational, Scientific, and Cultural Organization |
| WOCAT | World Overview of Conservation Approaches and Technologies |

List of Tables

| | | |
|-----|---|----|
| 1.1 | Information on the extent of human-induced soil degradation by type, degree, and causative factor for the world and major continents or regions published by GLASOD in 1999, expressed in millions of hectares[25]. | 10 |
| 1.2 | Land Degradation estimation by various national agencies of India | 16 |
| 1.3 | The development of remote sensing data and methods for assessing land degradation.[55] | 18 |
| 2.1 | Three-level hierarchical desertification/land degradation classification system | 28 |
| 2.2 | Satellite Data Description | 32 |
| 2.3 | Vegetation degradation in forest area classification scheme | 45 |
| 2.4 | Percentage area under different severity conditions for desertification | 47 |
| 2.5 | Accuracy comparison of different segmentation methods for Level 1 Land cover object creation with Random forest classification (values in %) | 48 |
| 2.6 | Level 1 Land Cover classification accuracy using three different algorithms (values in %) | 48 |
| 2.7 | Object based classification result for two different study area(values in %) | 51 |
| 2.8 | The classification results for forest vegetation degradation for two model performance (values in %) | 51 |
| 2.9 | The classification results for soil salinity for two model performance (values in %) | 52 |

| | | |
|------|---|-----|
| 2.10 | The degraded areas change for two different year (Area in %) | 52 |
| 3.1 | Ranges for soil sample test values used in soil rating for SOC, pH, and SEC[33] | 56 |
| 3.2 | Descriptive statistics of soil properties in the topsoil of the north Gujarat Agro-climatic zone | 57 |
| 3.3 | Soil covariates used for study | 59 |
| 3.4 | The performance of different Soil pH prediction models using three indicators of model evaluation with descriptive statistics | 66 |
| 3.5 | The performance of different Soil EC prediction models using three indicators of model evaluation with descriptive statistics | 67 |
| 3.6 | The performance of different Soil OC prediction models using three indicators of model evaluation with descriptive statistics | 67 |
| 4.1 | Dataset Specifications | 80 |
| 4.2 | Vulnerability index definition in MEDALUS model | 81 |
| 4.3 | Quantization of Climate Index | 82 |
| 4.4 | Quantization of Soil parameter | 83 |
| 4.5 | Quantization of Soil Index | 83 |
| 4.6 | Quantization of Slope Index | 84 |
| 4.7 | Quantization of Vegetation Index | 85 |
| 4.8 | Land Utilization Index | 87 |
| 4.9 | Classes definition in Fuzzy logic | 88 |
| 4.10 | Vulnerability index definition in Fuzzy logic | 88 |
| 4.11 | Percentage area under different severity condition for desertification vulnerability | 94 |
| 5.1 | The aridity zones based on Aridity index values | 101 |
| A.1 | Dataset Specification | 140 |
| B.1 | Strategic Framework: Micro-Watershed Level Action Plan for Combating Land Degradation | 145 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | Organization of the thesis chapters. | 22 |
| 2.1 | Study Area shown in as red polygon in Bhavnagar district and Panchmahal district | 30 |
| 2.2 | Land Use Land Cover map of Study area of Panchmahal (above) and Bhavnagar (below) Districts with the legends showing 16 classes and their color code | 33 |
| 2.3 | Desertification Status Map of Study area for two-cycle over 10 years apart | 34 |
| 2.4 | Methodology flow chart | 35 |
| 2.5 | The rotation-based augmentation process generates the patch that would be fed to the CNN model | 38 |
| 2.6 | Patch Based CNN Architecture | 38 |
| 2.7 | Object Based classification model flowchart | 40 |
| 2.8 | Level 1 classified Map using (a) SLIC segmentation comparison with (b) original GT class map for part of study area | 50 |
| 3.1 | Study Area showing soil sample points in North Gujarat Agro-climatic area | 70 |
| 3.2 | Predictive Soil modeling methodology flowchart | 71 |
| 3.3 | The above image shows a 3-layer artificial neural network: One input layer, two hidden layers, and one output layer. | 72 |
| 3.4 | Performance of Predicted Soil pH property for six different prediction models | 73 |

| | | |
|------|---|-----|
| 3.5 | Performance of Predicted Soil EC property for six different prediction models | 74 |
| 3.6 | Performance of Predicted Soil OC property for six different prediction models | 75 |
| 4.1 | Panchmahal District of Gujarat | 79 |
| 4.2 | Severity index for vegetation using fuzzy logic | 89 |
| 4.3 | Aridity Index Map | 90 |
| 4.4 | Soil Index Map | 91 |
| 4.5 | Elevation Index Map | 92 |
| 4.6 | Slope Index Map | 93 |
| 4.7 | Vegetation Index Map | 94 |
| 4.8 | Cumulative Amenities Index Map | 95 |
| 4.9 | Economic Development Index Map | 95 |
| 4.10 | Land Utilization Index Map | 96 |
| 4.11 | Desertification Vulnerability Index Map | 97 |
| 5.1 | District level climate classification of aridity zones of India.[129] . . | 101 |
| 5.2 | Meteorological observatories in Gujarat | 102 |
| 5.3 | Average (a) PET, (b) rainfall and (c) aridity index map of Gujarat (d) CGIAR-CSI global aridity index | 105 |
| 5.4 | Comparison of (a) normalized difference vegetation index(NDVI)during rabi season of 2008 and (b) aridity index map of 2008 | 106 |
| 5.5 | Comparison of (a) normalized difference vegetation index(NDVI) during rabi season of 2012 and (b) aridity index map of 2012 | 106 |
| 5.6 | Comparison of Histogram of rabi NDVI and AI map for year (a) 2008 and (b) 2012. | 106 |
| 5.7 | Relation between NDVI and AI value of random points for the year 2008 and 2012 | 107 |
| A.1 | (a) Original Classification Training Image, (b) Classification using SVM, (c) Classification using Gabor + SVM | 142 |

CHAPTER 1

Introduction

"Land" signifies a complicated bio-productive system that contains soil, vegetation, ecological and hydrological processes, and other biota that work inside the system[170].

"Land degradation" is the degradation or loss of the productive capacity of the land, caused by human activities, such as deforestation, overgrazing, soil erosion, and agricultural mismanagement. It results in the decline of soil fertility, reduced biodiversity, and other negative impacts on the environment, as well as social and economic consequences. Land degradation can have serious consequences for food security, water availability, and human well-being, and can be a major problem, particularly in arid and semi-arid areas[170].

The term "desertification" was first used by a French ecologist named Louis Lavauden in his 1927 publication titled "La desertification de la France méridionale" (The desertification of southern France)[95]. However, Lavauden's use of the term referred to the expansion of desert-like conditions in areas that were previously fertile, rather than solely the low productivity of rangelands.

Later, in the 1940s and 1950s, researchers began to use the term "desertification" specifically to refer to the degradation of drylands and rangelands, which were becoming increasingly arid and unproductive due to human activities such as overgrazing, deforestation, and poor land management practices. Later the term was utilized in 1949 by the French forester Aubreville, who used the term to allude to the removal of tropical rain-forest by secondary savannah and scrub in those areas of Africa where forestland was being burned and cleared to give land for cultivation [12]. Glantz et.al.[64] in 1983 examined the meaning of deser-

tification on a survey of more than 100 unique meanings of desertification taken from the writing. Various definitions center around changes in soil (e.g. salinization), vegetation (e.g. decreased biomass density), water (e.g. water-logging), or air (e.g. expanded albedo). Regardless of the main focus, the majority of them also describe changes in biological production while making observations about the type of vegetation. According to the United Nations Environment Programme (UNEP), desertification is "land degradation in arid, semi-arid, and dry sub-humid environments arising from adverse human influence." The United Nations Convention to Combat Desertification (UNCCD) backed the UNEP's 1992 revision of the definition of desertification as "Land degradation in arid, semi-arid and dry sub-humid areas arising from different sources, including climate fluctuations and human activities." in 1994 [170]. Later this term is increasingly widely used to characterize the effects of desertification. (Aridity and its characterization are given in detail in Chapter 5)

The process of turning fertile land into desert is known as desertification. Desertification is not just the growth of the desert. It is a form of land degradation when moderately dry land becomes more and more arid as a result of the loss of water, vegetation, and life due to a combination of natural and anthropogenic factors in arid, semi-arid, and dry-subhumid environments. Desertification can be considered to be a subsection of land degradation, occurring in arid, semi-arid, and dry sub-humid areas [52]. Thus desertification and Land Degradation (LD) can be summed up as a process that affects land due to the intermingling effect of various physical phenomena and anthropogenic activities in an intricate way. This millennium is constantly facing challenges from desertification all over the globe because 38% of the total global population inhabits drylands, which is about 41% of the total Earth's land surface[172]. Constant ecological changes due to rapid growth in population pressure, industrial growth and urbanization have encountered serious effects on the entire dynamics of dry land regions. Understanding these changes and the acting forces of desertification is thus the need of the hour. The most efficient and cost-effective way of monitoring these changes in spatial and temporal dimensions is by using remote sensing techniques through

various satellite imageries provided by different satellite sensors. Indian administrative boundary occupies 2.4% of the total geographical area of the world but supports 17% and 15% of the world's total human population and livestock [30]. The increasing pressure of human and animal populations is leading to excessive demand for natural resources, thereby leading to desertification and land degradation [176]. In India, 25.14% of the total geographic area of the country had been undergoing the process of desertification (arid, semi-arid and dry sub-humid regions) during the year 2011-2013 [136]. Gujarat is one of the fastest-growing states in India. Gujarat is located in the western part of India, with a 1,96,244 sq km area, which is 6.2 percent of the total geographical area of India. The state has a population of 6,04,39,692 (4.99% of country population); with 308 population density, 919 sex ratio and 78.03% literacy [31]. Gandhinagar is the capital of Gujarat. The major geographic features of Gujarat are the Rann of Kutch, the alluvial plain, the Saurashtra peninsula, the Girnar hills, the Vindhyan ranges and the coastal plain. The main rivers of the state include Sabarmati, Mahi, Narmada, Tapi, Bhadar and Shetrunji. Gujarat soils are of various types like sandy, saline, Clay, loamy and black cotton soil. The state comprises characteristics of an arid region in the western and northern parts and a semi-arid region in the southern and eastern parts. Gujarat experiences an extreme climate of very hot and dry summer and very cold winter with an average annual rainfall of 625 mm. The 19th Livestock Census (2012) [30] of India has placed the total livestock population at 512.05 million and the total of poultry birds at 729.2 million, out of which, there are 271.28 lakhs livestock (5.29%) and 150.03 lakhs poultry (2.06%) in the state of Gujarat. Thus, looking at the population pressure of human beings and livestock and their activities in addition to climatic wind and water effects, an attempt was made to measure the extent of land degradation and sustainable agriculture in Gujarat. The statistical summary and analysis of the Land Degradation of Gujarat state reveal that 52.22% (10.24 million ha) of the total geographical area is undergoing Desertification/Land Degradation [136]. The most significant process of desertification/land degradation in the state is Water Erosion (19.53%) followed by Salinity (13.24%), Vegetation Degradation (11.84%), and Wind Erosion (5.99%).

1.1 Causes of desertification & land degradation

Two different types of factors affect land degradation[35, 181].

1. Natural Factors

Natural factors that lead to land degradation and desertification are related to the natural conditions of a particular area[184]. These factors are typically stable over time, or at least change slowly enough to be observed at a human time scale. Examples of natural factors include topography, historical climate, vegetation, and soil conditions. However, even these stable conditions can be disrupted by extraordinary climatic events such as floods, earthquakes, and thunderstorms. These events can cause significant and rapid changes to the natural environment, resulting in soil erosion, loss of vegetation cover, and other forms of land degradation[182].

2. Human factor

Land degradation and desertification are often caused by man-made actions such as inadequate agricultural practices, deforestation, fires, tourism, urbanization, industrial activities, improper waste disposal, and mining. For example, excessive use of fertilizers and pesticides can damage the soil and reduce its fertility, leading to land degradation. Deforestation removes trees that hold the soil together, leading to erosion and soil degradation. Similarly, fires, whether natural or man-made, can destroy vegetation and contribute to soil erosion. Urbanization and industrialization can lead to the destruction of natural habitats and ecosystems, as well as the pollution of soil and water resources. Improper waste disposal and mining can also contribute to land degradation and desertification. All of these human activities can have a negative impact on the environment and the sustainability of land resources, leading to long-term ecological and economic consequences[5].

1.2 Indicators of desertification and land degradation

Indicators of desertification and land degradation are used to measure and assess the health of ecosystems[86].

1. Physical indicators

The physical indicators of desertification include changes in rainfall patterns and drought conditions. These changes can lead to a decline in the quality and quantity of ground and surface water, which can negatively impact vegetation growth and soil fertility. Additionally, desertification can lead to decreases in soil depth and organic matter, which can result in a decrease in the ability of the soil to support plant life. In some cases, desertification can lead to the formation of dunes and sandstorms, which can cause further damage to the ecosystem.

2. Biological indicators

Biological indicators of desertification include changes in vegetation cover, above-ground biomass, and yields of agricultural products. These changes are often the result of soil erosion and other forms of land degradation. When vegetation cover decreases, it can lead to increased soil erosion and a decline in soil fertility. This can cause a decrease in the availability of water and nutrients to plants, which can reduce their growth rates and overall health. Additionally, changes in above-ground biomass and agricultural yields can indicate that the ecosystem is becoming less productive and less able to support human and animal populations.

3. Animal/Livestock indicators

Animal and livestock indicators of desertification are also important in assessing the health of ecosystems. Changes in the distribution and frequency of key species can indicate that the ecosystem is becoming less hospitable to these species. Additionally, declines in livestock production and yield can be an early warning sign of ecosystem stress, as livestock are often a key component of the social and economic fabric of many communities.

4. Social/economic indicators

Social and economic indicators can provide valuable insights into the impacts of desertification and land degradation on communities. Changes in land use and water use can indicate that communities are struggling to adapt to changing environmental conditions. Additionally, changes in settlement patterns and population parameters such as migration statistics and public health information can provide important information on the social and economic impacts of land degradation. Changes in social processes such as increased migration, decrease in incomes and assets, and changes in relative dependence on cash crops versus subsistence crops can all be important indicators of the social and economic impacts of desertification and land degradation.

1.3 Related Work

1.3.1 Concept of desertification & land degradation

Changes in land use and land cover (LULC) is being caused by rising demands placed on the environment as a result of economic growth, expanding cities, and rising rural populations. Land degradation is then a result of unsustainable land usage and land cover changes. Therefore, land degradation is defined as a temporary or permanent reduction in the productive capacity of land [127]. It addresses many types of soil degradation brought on by anthropogenic and natural sources. Clearly, land degradation is a result of a variety of processes that both directly and indirectly decrease the usefulness of the land.

Desertification is the word for land degradation caused by a variety of reasons, such as climate changes and human activity, in arid, semi-arid, and dry sub-humid environments[171]. The combination of several land degradation processes acting over a landscape results in desertification, which increases in harsh circumstances. It is a complicated phenomenon that necessitates the knowledge of academics from a variety of fields, including geography, political science, eco-

nomics, climatology, soil science, meteorology, hydrology, agronomy, and veterinary medicine[63].

Historical Perspective

Since the dawn of human history, environmental damage in arid and semi-arid areas has been a major issue for human civilizations. According to the initial report of the South African Drought Investigation Commission, both northern and southern Africa experienced severe droughts in the first two decades of the 19th century [62]. Numerous other tragic occurrences caught people's attention, including the Dust Bowl in the United States in the early 1930s and the Sahara's seeming extension[157, 98]. These are a few of the occasions that might be considered benchmarks or turning points in the emergence of the idea of desertification.

The Arid Zones Research program was started by the United Nations Educational, Scientific, and Cultural Organization (UNESCO) in the early 1950s. This initiative produced a perspective framework regarding dry zones and led to the release of the first encyclopedia on LD[177]. Later, as a result of the negative effects of an extended drought in the West African Sahel in the early 1970s, an increase in interest in environmental conservation was visible throughout the developed world.

Finally, as a result of this awareness, the Stockholm Conference on Environmental Problems was held in 1972, and the United Nations Environment Programme was established(UNEP). Following this, the UN General Assembly mandated the establishment of the UN Sudano-Sahelian Office and directed UNEP to plan an international convention on desertification. The meeting was held in Nairobi in 1977 and brought together representatives from several nations whose landscapes had been either directly or indirectly impacted by the phenomena of desertification. This immediately drew the attention of national and international policy-makers, as well as academics on dry regions, who set out to develop plans and tactics to tackle the global desertification crisis.

The United Nations (UN) General Assembly established the 17th of June as "World Day to Combat Desertification and Drought" in 1994 with the goal of rais-

ing public awareness of the issue and encouraging the implementation of United Nations Convention to Combat Desertification (UNCCD) plans and initiatives in nations with severe desertification issues. The convention has now matured and is transitioning from developing National Action Programmes (NAP) to putting them into action at the local and national levels, which is resolving significant problems.

1.3.2 Land degradation assessment on global scale

Scientists all throughout the globe have created many ways for assessing and monitoring LD and desertification, concentrating on the scale and severity of the damage so that conservation efforts may be planned. Over the last 30 years, several attempts have been made by international organizations, as well as individual researchers, to conduct LD and Desertification studies on a worldwide scale using various methodologies. International conferences such as the United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro in 1992, the Convention to Combat Desertification (CCD) in Paris in 1994, and the International Conference on Population Development in Cairo in 1994 helped to clarify the issues surrounding LD and Desertification.

It was observed that soil directly affects the land's ability to produce, hence it is important to investigate efficient soil degradation control measures. The International Soil Reference and Information Centre (ISRIC) was given a job by UNEP in 1987 to establish a system for a global evaluation of human-induced soil degradation. A Global Assessment of Human-Induced Soil Degradation map (GLASOD) with a scale of 1 : 10,000,000 (Mercator projection) was created by ISRIC in partnership with numerous environmental experts and soil scientists from around the world. The map shows the degree and type of degradation brought on by human stresses on the land and soils [25]. The GLASOD map's biggest flaw is that no specific form of soil degradation can be identified on the map. Four main stages of soil degradation were discovered by GLASOD map. The first two categories deal with soil material movement brought on by wind or water activities. The third category addresses in-situ soil degradation caused by chemical damage, while

the fourth deals with physical degradation. However, it should be noted that just a little amount of remote sensing was used in this map, which entirely relied on the opinions of experts [79]. From GLASOD data, a number of themed maps have been created and published in UNEP's World Atlas of Desertification [109]. The GLASOD classification scheme and the statistical analysis based on the continental area are displayed in Table 1.1. The types of degradation listed in the table include water, wind, nutrient decline, salinization, pollution, acidification, compaction, waterlogging, and subsidence of organic soil. The degree of degradation is classified as light, moderate, strong, and extreme. The causative factors include deforestation, overgrazing, agriculture mismanagement, overexploitation, and industrial activity.

According to the table, the total extent of human-induced soil degradation in the world is 1965 million hectares, with the largest areas affected found in Asia (747 million hectares) and Africa (494 million hectares). Water degradation is the most common type of soil degradation, with 1094 million hectares affected worldwide. Overgrazing is the leading causative factor of soil degradation, affecting 678 million hectares globally. The table highlights the significant impact that human activities have on soil degradation, and the need for sustainable land management practices to mitigate further degradation and maintain soil health[25].

The Assessment of Human-Induced Soil Degradation in South and Southeast Asia (ASSOD), a follow-up initiative to GLASOD, was created for the Asian area. In October 1993, regional office of the UN's Food and Agriculture Organization (FAO) for Asia and the Pacific-Bangkok decided to complete ASSOD maps using the same classification system as GLASOD but with more detail, so Asia, as well as the the region of Asia-Pacific, were mapped at a greater level of 1:5 million instead of 1:10 million scale[175].

Based on the georeferenced datasets of GLASOD, the SOVEUR project from 1997 to 2000 mapped the vulnerability of the soil and topography in Central and Eastern Europe, paying particular attention to soil contamination issues[17]. The Pilot Assessment of Global Ecosystems (PAGE) contrasted a newly determined global region of agriculture with GLASOD data as its fundamental foundation.

Table 1.1: Information on the extent of human-induced soil degradation by type, degree, and causative factor for the world and major continents or regions published by GLASOD in 1999, expressed in millions of hectares[25].

| | World | Asia | North Africa & West Asia | Africa | South Amer- ica | Central Amer- ica | North Amer- ica | Europe | Oceania |
|----------------------------|-------|------|--------------------------------------|--------|-----------------------|-------------------------|-----------------------|--------|---------|
| Degradation Type | | | | | | | | | |
| Water | 1094 | 440 | 84.1 | 227 | 123 | 46 | 60 | 114 | 83 |
| Wind | 548 | 222 | 145.2 | 187 | 42 | 5 | 35 | 42 | 16 |
| Nutrient Decline | 135 | 14 | 6.3 | 45 | 68 | 4 | - | 3 | - |
| Salinization | 76 | 53 | 46.9 | 15 | 2 | 2 | - | 4 | 1 |
| Pollution | 22 | 2 | 0.3 | - | - | - | - | 19 | - |
| Acidification | 6 | 4 | - | 2 | - | - | 1 | - | - |
| Compaction | 68 | 10 | 3.6 | 18 | 4 | - | - | 33 | 2 |
| Waterlogging | 11 | - | 0.1 | - | 4 | 5 | - | 1 | - |
| Subsidence of Organic Soil | 5 | 2 | - | - | - | - | - | 2 | - |
| Total | 1965 | 747 | 286.5 | 494 | 243 | 62 | 96 | 216 | 102 |
| Degradation Degree | | | | | | | | | |
| Light | 749 | 295 | 142.8 | 173 | 105 | 2 | 17 | 60 | 96 |
| Moderate | 910 | 344 | 113.7 | 192 | 113 | 35 | 78 | 144 | 4 |
| Strong | 296 | 108 | 29.6 | 124 | 25 | 26 | 1 | 10 | 2 |
| Extreme | 9 | - | 0.4 | 5 | - | - | - | 4 | - |
| Causative Factors | | | | | | | | | |
| Deforestation | 579 | 298 | 52.7 | 67 | 100 | 14 | 4 | 84 | 12 |
| Overgrazing | 678 | 198 | 152 | 243 | 68 | 9 | 29 | 48 | 83 |
| Agriculture Mismanagement | 552 | 204 | 49.1 | 121 | 64 | 28 | 63 | 64 | 8 |
| Overexploitation | 133 | 46 | 32.3 | 63 | 12 | 11 | - | 1 | - |
| Industrial Activity | 23 | 1 | 0.4 | - | - | - | - | 21 | - |

The PAGE findings indicated that degradation caused by humans had been more severe than anticipated by the GLASOD since the mid-1900s[186].

In 1991, UNEP merged GLASOD data with maps from the International Centre for Arid and Semiarid Land Studies of Texas Technical University, which showed the primary land use classifications, to conclude that about 2,600 Mha, mostly the rangelands, were affected by vegetation degradation that was not clearly reflected in GLASOD. This study's publication received harsh criticism for GLASOD [137]. After 1992, it became clear that a consistent approach was needed for both the detailed assessment of soil degradation and its conservation. At this point, the approach for the World Overview of Conservation Approaches and Technologies (WOCAT) developed. The Center for Development of the Environment in Bern,

Switzerland supports the worldwide collaboration that organizes the world network known as WOCAT.

In addition to more than 35 national partner institutions, this forum included FAO, UNEP, Regional Land Management Unit representatives, International Centre for Integrated Mountain Development, Swiss Agency for Development Co-operation, International Atomic Energy Agency, Danish International Development Agency and Swiss Agency for Development Co-operation. The main goal of WOCAT was to make it possible for individuals from around the world to share knowledge and experiences in order to recognize and address soil degradation. In addition to this, a database of trustworthy sources of information covering numerous geographic regions was also intended to be created.

The data acquired provided suggestions for how to improve current procedures and helped uncover research problems. To gather data on technology, field-level techniques, and mapping relevant to soil and water conservation, WOCAT created a set of three extensive questionnaires. Data were analyzed in a systematic way using a standardized tool. In addition, WOCAT started offering training sessions on how to complete the questionnaires and ultimately update the database. As a result, WOCAT is the first significant effort to record soil conservation and watershed management actions in a consistent manner at all scales, from global to regional [174].

Although WOCAT has produced the intended outcome efficiently and cost-effectively, there is a need to shift the focus from developing tools and techniques using new-age technology to their implementation in nations and regions that are experiencing land degradation.

The Millennium Ecosystem Assessment (MA) was another method for figuring out the implications of ecosystem change. Kofi Annan, the Secretary-General of the UN, called MA in 2000. It evaluated the demands of the Ramsar Convention on Wetlands, the Convention on Migratory Species, the Convention to Combat Desertification, and the Convention on Biological Diversity, as well as the needs of other users in civil society. The MA makes use of data from models, datasets, practitioners, and local communities as well as research articles. In four different

expert working groups, more than 1,300 writers from 95 countries contributed to the creation of the global assessment.

MA offers management and planning tools as well as insight into how decisions affecting ecosystems may affect them in the future. Since the MA has not been focused on global land degradation, only one of the systems, the "Dryland Systems," has thoroughly analyzed land degradation. The MA examined land degradation in dry regions and revealed that by 1990, 14 major terrestrial biomes and more than half of the area of four additional biomes had been changed, primarily to agricultural and livestock production systems. Lepers et al.[93] used remote sensing and regional data sets that partially overlapped to assess desertification for the MA in 2003. The soil's condition was regarded as being represented by the vegetative cover and its Net Primary Productivity (NPP). Based on this methodology, it was determined that 10% of the world's drylands, including hyper-arid regions, were degrading between 1981 and 2000[100]. However, the MA assessment's main flaw is that its authors limited their conclusions to the generalization that while the entire costs of ecosystem service degradation are not easy to quantify, the information now available shows that they are significant and rising[99].

The Land Degradation Assessment in Dry Lands (LADA) was created as a result of additional research on this. This project aimed to produce national, regional, and global assessments for planning and implementing interventions to ameliorate land degradation and promote sustainable land use and management practices. The objectives included quantifying and analyzing the type, extent, severity, and impacts of land degradation on ecosystems as well as carbon storage in dry lands at various spatial and temporal scales. The Global Land Degradation Assessment in Dry Lands (GLADA), which came after LADA, was based on NPP and how it changed over time utilizing the NDVI and rain-use efficiency data.

In accordance with GLADA findings, 3,510 Mha of the terrestrial land surface, or 24% of the world's land area, experienced degradation between 1981 and 2003. Tropical Africa and south of the equator, Southeast Asia, South China, North-central Australia, drylands and sloping lands of Central America and the

Caribbean, Southeast Brazil, the Pampas, and the boreal woods were the areas most severely affected [14].

A Global Land Degradation Information System (GLADIS) was created by FAO and partners as a follow-up to LADA/GLADA [115]. As GLADIS set out to ascertain the status of ecosystems' capacity to deliver goods and services and the change in this capacity, GLADA concentrated on the production function. The six quantifiable categories of biomass, soil health, water quantity, biodiversity, economic services, and social services were used to categorize goods and services. Global data sets on land use and management, climatic conditions, socioeconomic conditions, etc. were evaluated using models to provide a baseline condition for the assessment of the state [57]. GLADIS offers a number of global maps that can be downloaded and queried that display the condition and trends of ecosystem services. A growing number of maps and databases that list the input information used to calculate each individual parameter are available to assist GLADIS. Ancillary maps, such as global land use and land cover map with attributes, are also given to accompany the later datasets. GLADIS relies mostly on socioeconomic information and a few bio-physical characteristics for its indicators, which are very few. In an effort to get a sense of the subnational condition, an effort was made to concentrate on the most important criteria and, whenever possible, include geo-referenced indicators. This has only been achievable for biophysical variables, while country statistics are often provided for socio-economic factors. Therefore, spatial distribution over the globe needs to be completed.

The United Nations Environment Programme (UNEP) published the World Atlas of Desertification (WAD) in 1992 as a comprehensive global assessment of the state of desertification and land degradation. The goal of the atlas was to provide a comprehensive picture of the extent and severity of desertification, as well as the underlying causes and drivers, as well as the impacts on human and natural systems.

The WAD identified the global distribution of desertification, highlighting that approximately one-third of the earth's land surface is arid or semi-arid, with desertification affecting approximately one-third of this area. Climate change, land

use change, and unsustainable land management practices such as overgrazing and deforestation were also identified as major drivers of desertification in the report[108]. The WAD contributed significantly to the global understanding of desertification and land degradation by providing a comprehensive and easily accessible source of information on the subject. Recent assessments, such as the Intergovernmental Panel on Climate Change's Special Report on Climate Change and Land (2019) and the Intergovernmental Panel on Biodiversity and Ecosystem Services(IPBES)' Global Assessment Report on Biodiversity and Ecosystem Services (2019), have highlighted the need for updated data and methods to better understand the extent and impact of land degradation and desertification at the global level.

The United Nations Convention to Combat Desertification (UNCCD) conducts the Global Land Outlook (GLO), a comprehensive global assessment of the status, trends, and challenges of land and land-based ecosystems, as well as the drivers of land degradation. The first edition of the GLO was released in 2017, and the second edition was released in 2021. The second edition of the GLO, 2021, provides updated information and analysis on the status and trends of land degradation, as well as the drivers of change, and emphasizes the critical need for action to address land degradation and strengthen the resilience of land-based ecosystems in the face of multiple global challenges, such as climate change, biodiversity loss, and the COVID-19 pandemic[173].

1.3.3 Land degradation mapping in the context of India

India, which primarily relies on agriculture for its economy, is particularly more concerned about land degradation. Agriculture provides a means of subsistence for two-thirds of the people. In the past 20 years, numerous national and regional policies have been launched to address this issue, but the outcomes have been insufficient. For the purpose of creating effective policies to address the LD issue, analysis of causes and extents is crucial. The Convention to Combat Desertification (CCD), which went into effect in 1996, has 176 signatories, including India. To stop additional dry land degradation, the Convention outlined a framework

for national, regional, and sub-regional programs. The CCD is represented by the Ministry of Environment, Forestry & Climate Change (MoEF & CC) in the case of India. Additionally, it is in charge of creating and carrying out the National Action Plan (NAP). The Regional Action Programme (RAP) was developed under UNCCD [170] to fight desertification. In order for the member nations of the Asian region covered by RAP's programme to effectively combat desertification, those nations must have stronger capacities. Six theme programme areas were chosen for the RAP. The first Thematic Program Network (TPN-1) focuses on monitoring and evaluating desertification.

To develop the national network for desertification monitoring and assessment in India, the Space Applications Centre (SAC), Indian Space Research Organization (ISRO), Ahmedabad, was chosen as the national focal institution. Using IRS LISS-III data, a pilot project was launched to standardize and produce a detailed classification system and methodology for desertification status mapping at a scale of 1 : 50,000 under varied dry-land conditions in both hot and cold deserts in India. The entire nation's desertification status was mapped out at a scale of 1 : 5,00,000 using IRS P6 AWIFS data. Using satellite data, it has been possible to identify and map the main causes of land degradation, including water erosion, vegetative degradation, wind erosion, salinization/alkalization, water-logging, frost heaving, frost shattering, mass movement, etc.

The study finds that 105.48 mha (million hectares) areas or 32.07% of the nation's overall geographic area are experiencing land degradation. The desertification zone is 81.4 mha [7]. Additionally, the estimates have been improved by using digital satellite photos rather than hardcopy satellite images, the statistical summary and analysis indicates that there has been an increase in the area undergoing land degradation in India over the years. The most recent data for the timeframe 2018-19 shows that 97.85 million ha, which is equivalent to 29.77% of the country's Total Geographic Area (TGA), is undergoing land degradation. This is higher than the areas undergoing degradation in the timeframes of 2011-13 and 2003-05, which were 96.40 mha (29.32% of TGA) and 94.53 mha (28.76% of TGA), respectively. Moreover, the data shows that there has been a cumulative increase

of 1.45 mha (0.44% of TGA) undergoing degradation from 2011-13 to 2018-19, indicating a worsening trend. In contrast, there was a cumulative increase of 1.87 mha (0.57% of TGA) from 2003-05 to 2011-13, which suggests that the rate of increase has slowed down in recent years[136, 154]. Other national organizations have evaluated LD over India as well. These assessments produce a wide range of findings, primarily because diverse methodologies for designating degraded areas and distinct evaluation criteria were applied [135]. The LD estimation by several national agencies in India is shown in Table 1.2. Land degradation has been cited in numerous studies as one of the main environmental issues facing our nation [119, 132, 147, 189].

Table 1.2: Land Degradation estimation by various national agencies of India

| Agency | Estimated Area (Mha) | Criteria for Delineation |
|---|----------------------|--|
| National Commission of Agriculture (1976) | 148.09 | Based on secondary data collection |
| Ministry of Agriculture (1978) | 175 | Based on Net Cultivated Area estimates |
| Society for Promotion of Wasteland Development (1984) | 129.58 | Based on secondary data collection |
| National Remote Sensing Agency (1985) | 53.28 | Mapping on 1:1 million scale based (1980-82) on remote-sensing techniques |
| Ministry of Agriculture (1985) | 173.64 | Land degradation statistics for the states based on Net Cultivated Area estimates |
| Ministry of Agriculture (1994) | 107.43 | Elimination of duplication of area above |
| NBSS&LUP (1994) | 187.7 | Mapping 1:4.4 million scale at country level and then deducting at state level based on Global Assessment of Soil Degradation (GLASOD) Mapping 1:4.4 million scale at country level and then deducting at state level based on GLASOD guidelines |
| NRSC(2000) | 63.85 | Based on satellite data (1986-1996) |
| NRSC (2005) | 55.27 | Based on satellite data (LISS-III sensor data of 2003) |
| NBSS&LUP (2005) | 146.82 | Mapping of all the states at 1:250,000 scale. Global Assessment of Soil Degradation (GLASOD) guidelines |
| SAC (2016) | 96.40 | Based on satellite data (2011-2013)1:5,00,000 |
| SAC (2021) | 96.85 | Based on satellite data (2019-2021)1:5,00,000 |

1.4 Importance of Remote Sensing and Machine Learning

Machine Learning (ML) and Remote Sensing (RS) have become increasingly useful in land degradation and desertification assessment and mapping. The combination of these two technologies allows for a more accurate and efficient analysis of the changes in land use and land cover over time, which are indicators of land degradation and desertification.

Remote sensing is a powerful tool that enables scientists to gather information about the Earth's surface and atmosphere using data collected from satellites, aircraft, and other platforms. These sensors capture data in the form of electromagnetic radiation, which can be used to detect changes in vegetation cover, soil moisture, and other environmental parameters that are important indicators of land degradation and desertification[104, 178]. Over the years, the development of satellite remote sensing technology has significantly improved, providing us with more detailed and accurate information about the globe.

Table 1.3 summarizes some of the key developments in satellite remote sensing technology over the years, starting with aerial cameras, to the present satellite imagery. Some of the notable developments include the launch of the Landsat (Land Remote Sensing Satellite) series of satellites in the 1970s, which provided high-resolution imagery of the Earth's surface, and the development of the Advanced Very High Resolution Radiometer (AVHRR), which has been widely used for vegetation mapping and monitoring. The evolution of remote sensing-based methods for land degradation (LD) mapping, monitoring, and assessment are also summarized in Table 1.3. Among these methods, the analysis of vegetation cover dynamics and vegetation decline analysis is the most commonly applied ones[55].

Machine learning algorithms can be trained to analyze large datasets generated by remote sensing and to classify land cover and land use changes over time. This allows for the detection of patterns and trends in land degradation and desertification that may be difficult to identify using traditional manual methods.

ML and RS can also be used to identify areas that are at high risk of land

Table 1.3: The development of remote sensing data and methods for assessing land degradation.[55]

| | 1970-1980 | 1980-1990 | 1990-2000 | 2000-2010+ |
|--------------------------------|---|---|---|---|
| Input data | Multi-spectral images, aerial photos | Multi-spectral images, aerial photos and derivatives, vegetation indices | Multi-spectral images, aerial photos and derivatives, vegetation indices, derivatives from spectral transformation (Tasseledcap, PCA, SMA), vegetation biophysical parameters | UAV images and aerial photos (and derivatives), multi- and hyperspectral image (and derivatives) time series, vegetation indices and vegetation productivity estimates, yield estimates, spectral transformation, vegetation biophysical parameters |
| Methods (examples) | Visual interpretation of aerial photos, photogrammetric methods, manual mapping | Image classification, map digitalization, expert mapping, photogrammetric methods, manual mapping | Image classification, spectral transformation, change detection, semi-quantitative image analysis, expert mapping | Time series analysis, data fusion, LD modeling, image classification, spectral transformation, change detection, participatory mapping methods |
| Sensors/platforms (examples) | Aerial cameras, CORONA, Landsat MSS | Landsat TM, SPOT, AVHRR | Landsat ETM/ETM+, SPOT, ASTER, AVHRR, LISS | LISS, Cartosat, Landsat, SPOT, AVHRR, Aster, MODIS, MERIS, Sentinel, RapidEye, IKONOS, Quickbird, GeoEye, Hyperion, UAV |
| Resolutions (exemplary ranges) | 180 m, panchromatic and few multispectral bands | 1 m to 8 km, panchromatic, multispectral and thermal bands | 1 m to 8 km, increasing number of bands | 0.01 m to 8 km, increasing number of bands, hyperspectral sensors |

degradation and desertification and to develop predictive models that can help to forecast future changes in land use and land cover. This information can then be used to inform policy decisions related to land management and conservation.

The use of machine learning and remote sensing techniques for land degradation and desertification assessment and mapping has become increasingly important in recent years. There have been several studies that have focused on the application of machine learning algorithms and remote sensing data for accurate and efficient assessment and mapping of land degradation and desertification.

Many studies demonstrate the effectiveness of machine learning algorithms

and remote sensing data for mapping and monitoring land degradation, soil erosion, and desertification in various regions of the world. Several studies used machine learning algorithms to map land degradation and soil erosion in different regions of the world[77, 193, 191].

In India many research studies have utilized various remote sensing and GIS techniques to map and monitor land degradation and soil erosion in different parts of the country[40, 41, 138, 83]. These studies used machine learning algorithms such as Random Forest and Support Vector Machines for the classification of satellite data.

Overall, the use of Machine Learning and Remote Sensing in land degradation and desertification assessment and mapping has significant potential and can be a powerful tool to improve our understanding of these phenomena and to guide effective land management practices.

1.5 The motivation of the work

Land degradation is one of the most pressing environmental issues that we face today. It refers to the loss of productivity of land due to natural or human-induced factors, including deforestation, erosion, and overgrazing, among others. Land degradation can result in a wide range of negative impacts on the environment, including the loss of biodiversity and ecosystem services, reduced food security, and increased vulnerability to natural disasters. Desertification, in turn, exacerbates the impact of climate change and reduces the ability of people to sustain themselves in affected regions. Therefore, developing effective strategies to monitor and mitigate the effects of land degradation and desertification is crucial for the sustainable use of natural resources and the livelihoods of millions of people living in affected regions. The Land Degradation Neutrality (LDN) Target Setting Programme was launched by the UNCCD in 2014 to help countries achieve LDN, which is defined as a state whereby the amount and quality of land resources necessary to support ecosystem functions and services and enhance food security remains stable or increases. India, being a party to the UNCCD, has been working

towards the goal of achieving Land Degradation Neutrality. India faces significant land degradation challenges due to factors like deforestation, unsustainable agricultural practices, urbanization, and climate change.

To address this issue, it is essential to monitor land degradation and its status on a timely basis to identify areas that require immediate attention. Avoiding, reducing, and reversing land degradation and restoring degraded land is an immediate need to secure the biodiversity and biological system benefits that are imperative to life on Earth. To pause and reverse the latest trends in land degradation, there is an urgent need to monitor these process and its status on timely manner. Remote sensing can play a significant role for generating this indicator. Distinguishing land degradation and perceiving its different sorts is a need to go to the practical lengths for combating it just as preserving and keeping the soil healthy. Remote sensing techniques, which allow for the systematic monitoring and mapping of land surface changes, offer a promising tool for achieving this goal. Therefore, there is a need to develop strategies to gain from medium to high-resolution satellite EO data.

In response to this issue, the presented thesis proposes an innovative approach to monitoring land degradation and desertification. Our objectives target for presenting the application of remote sensing to monitor land degradation and desertification in arid areas by incorporating satellite data with meteorological information and in situ information to detect land degradation and desertification processes. This thesis proposes using remote sensing as a tool to monitor land degradation and desertification in different areas. The proposed approach will not only provide effective information to monitor changes in land surface but also help decision-makers in producing relevant relief measures for sustainable resource exploitation.

1.6 Objectives and Accomplishments of the Thesis

The main objective of this thesis is to identify and understand the land degradation and desertification process by utilizing remote sensing data and machine

learning models. To accomplish this, four primary objective objectives were set for this thesis.

- The first objective is to pattern recognition of desertification process. For this, semi-automatic algorithms have been developed that utilize satellite data and machine learning models to identify desertification status. A new three phase methodology has been developed to map the different stages of desertification.
- The second objective is to focus on soil degradation, which occurs when soil health or productivity is declining. Here, developed an algorithm for predictive soil modeling using satellite data and in-situ ground observations. Artificial neural network algorithms have been developed for this soil mapping.
- The third objective is to identify the vulnerability of areas to desertification. Environmental parameters were used in combination with anthropological data to find areas vulnerable to desertification. Two different approaches were explored for vulnerability assessment. This objective helps in identifying the areas that are most vulnerable to desertification.
- The fourth objective of this thesis is to study and identify desertification hotspots using Aridity index. Climate change is a driving force for desertification and land degradation, and this objective helps in identifying the areas that are most vulnerable to the effects of climate change. Desertification hotspots are vulnerable areas within defined aridity zones. This objective helps in identifying the areas that are most affected by desertification.

Overall, the objectives of this thesis aim to provide effective information for identifying and monitoring land degradation and desertification. The utilization of remote sensing data and machine learning models can aid in identifying the areas that are most affected by these issues, and appropriate measures can be planned and implemented to improve the situation. Identifying the desertification hotspots and vulnerable areas can help in formulating and implementing targeted mitigation measures.

We discuss in brief each of our above-mentioned objectives in the following thesis chapters.

1.7 Organization of Thesis Chapters

The thesis is being planned to be structured in six chapters as shown in Fig. (1.1).

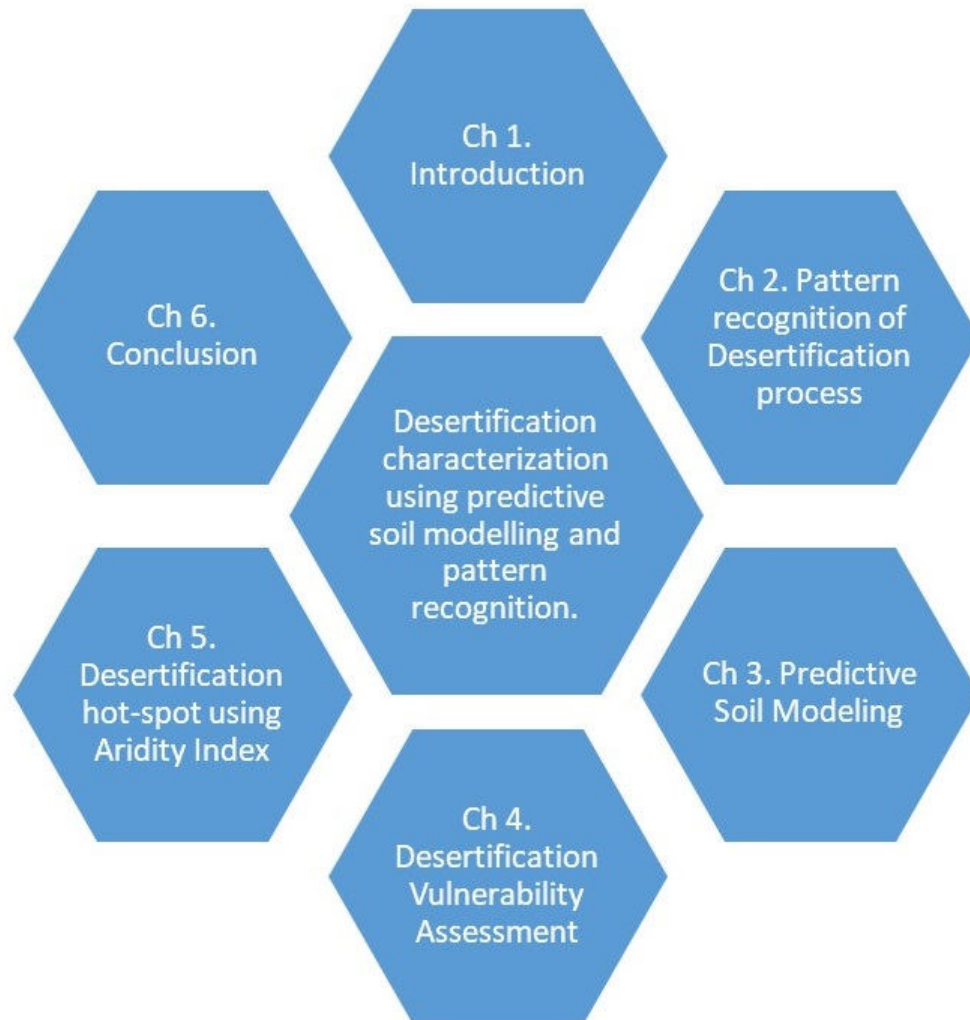


Figure 1.1: Organization of the thesis chapters.

Remote sensing images are essential for various applications be it social or environmental. This study aims at providing a hierarchical methodology for land degradation mapping, land use land cover classification, degradation process identification and mapping using multispectral LISS-III images.

Chapter 2: Pattern recognition of the desertification process was performed in

this chapter by proposing patch-based CNN architecture, a deep learning architecture and Object-based image segmentation to achieve high accuracy on level 1 land cover classification. The results of the different algorithms are compared for different terrains. The study demonstrates that segmentation with the random forest method outperforms CNN and SVM with an accuracy of 85% for level 1 land cover classification. Further, vegetation degradation in forest areas is assessed in central parts of Gujarat, India, having deciduous tropical forest cover. And with this land degradation in agricultural areas due to soil salinity is studied, particularly in southeastern parts of Gujarat, India. Degradation due to soil salinization is very common in arid and semi-arid regions. ML algorithms like SVM and RF were applied to different features to identify the degradation process.

Chapter 3: Soil degradation causes desertification and severely reduces potential soil productivity. Due to this, the degradation of the ecosystem and its associated ecosystem services are severely affected. There are several soil properties through which one can get an idea about the soil condition. Soil properties information can indicate the status of the soil desertification of the particular region. Different soil properties play different roles in the soil system. The machine learning algorithms used and ANN-based model has been developed to predict the soil properties like EC, pH and OC. These three soil properties are important indicators of soil degradation. Environmental parameters were taken as covariates in prediction models. Various covariates used in the study are, vegetation indices, terrain indices, soil parameters, spatial attributes and meteorological parameters of the study region. Field soil sampling data of the study region was obtained from Soil Health Card (SHC) for the year 2014 and was incorporated in training the model. The SHC data was divided into different ratios like 60:40, 70:30, and 80:20 for training and testing the model respectively. The SCORPAN model [106] has been considered as the base approach for the development of the ANN-based prediction model.

Chapter 4: Discusses mapping of vulnerable areas to desertification. Land degradation is a complex set of processes including climate change and human

activities, which interact over space to reduce or lose land productivity and lead to desertification. In the present study, a desertification vulnerability assessment was carried out combined with remote sensing and geographic information system (GIS) to map sensitive areas. Assessment through remote sensing offers a series of advantages such as data consistency, fairly near real-time data acquisition and a source for having spatially explicit data. Two different approaches were used for vulnerability assessment. Mediterranean Desertification and Land Use (MEDALUS)[22, 40] approach identifies such sensitive areas based on an index in which environmental quality and anthropogenic factors were included as layers for the Panchmahal district of Gujarat state in India. Second approach followed was fuzzy logic (FL) method in which fuzzy membership function[38, 41] was used to identify the risk area prone to desertification. Soil, climate, land utilization, geography and vegetation contribute to the land degradation of any area. However, man's intervention leads to significant changes in the environment, making socio-economic factors a considerable input to assess desertification vulnerability. Indices related to these factors have been generated. As a result, both methods have used natural and socio-economic factors to find the severity level of the desertification vulnerability in the Panchmahal district.

Chapter 5: Discuss the role of climate in the process of desertification. For the climate parameters, I have used the aridity index which incorporates most of the weather data like temperature, rainfall, humidity, wind and solar radiation etc. The aridity index (AI) is a useful parameter to study desertification conditions and its pattern. The AI formulation adopted by United Nations Environment Program (UNEP), Food and Agriculture Organization (FAO), and United Nations Convention to Combat Desertification (UNCCD), represents a simple but effective scientific investigation tool. AI is calculated by dividing the total annual precipitation by the annual potential evapotranspiration (PET).[16] The objective is to study and identify the desertification hot-spot using AI in the Gujarat state. Desertification hot spots are vulnerable areas within defined aridity zones. The weather data e.g. minimum temperature, maximum temperature, solar radiation, wind speed, humidity and rainfall for more than 18 locations all over Gujarat for

the past 20 years has been used in this study. FAO Penman-Monteith method[8] was used to calculate PET. Which along with rainfall were used to calculate AI for different locations. Annual AI map for the whole of Gujarat has been generated using these values and compared with CGIAR(Consultative Group on International Agricultural Research) based aridity map. MODIS-Terra NDVI product for the past 20-year period of *rabi* season has been used to get a correlation of AI with NDVI. In addition to comparing annual AI and NDVI data, thirty years average AI map has been generated for the State.

In the upcoming chapters of the thesis, a detailed discussion of the objectives outlined earlier is given. The following chapters will provide a exploration of each objective, along with an in-depth analysis of their work. This will involve a meticulous examination of the methodology employed, the datasets utilized, the study area selected, and a thorough review of existing literature relevant to each objective. This comprehensive approach ensures that readers are equipped with the necessary information to comprehend and evaluate the subsequent findings and conclusions of the thesis.

CHAPTER 2

Pattern recognition of Desertification process

Remote sensing techniques are an efficient tool for monitoring the Earth in a short time and at low cost.[102] One of the primary advantages of remote sensing in land cover classification is its ability to capture data over large and inaccessible areas. Remote sensing can cover vast areas in a short period of time, making it a valuable tool for monitoring and assessing changes in land cover. Remote sensing can also capture data in areas that are difficult to access, such as rugged terrain or dense forests, making it a useful tool for mapping areas that are not easily accessible by ground-based methods[183, 144]. Remote sensing data, along with advancements in image processing techniques, machine learning and artificial intelligence, geographical information systems, etc. have been used to monitor activities like land use mapping and change detection, deforestation rates etc [178]. Remote sensing images are essential for various applications be it social or environmental[183, 68, 166, 158]. One of the most significant advancements in remote sensing and machine learning-based land cover classification is the availability of high-resolution satellite imagery. The use of this imagery has enabled researchers to develop more accurate and detailed land cover maps. The integration of machine learning algorithms, such as support vector machines, random forests, and artificial neural networks, has also enhanced the accuracy of land cover classification by automating the process and reducing human error. Several studies have compared the performance of different machine learning algorithms for land cover classification[162, 164, 194, 148].

Another significant development in land cover classification is the use of multi-spectral and multi-temporal remote sensing data. Multi-spectral data refers to the

use of different spectral bands to capture information about different features on the earth's surface. Multi-temporal data refers to the use of satellite imagery taken at different times to capture changes in land cover over time. The combination of multi-spectral and multi-temporal data has been shown to improve the accuracy of land cover classification[185].

This study aims at providing methodological options for land cover classification, degradation process identification, and mapping using multispectral LISS-III imagery. A different stage hierarchical approach is proposed in this study. Table 2.1 shows the classification system in the hierarchical desertification mapping process[7, 136], where different levels are related to different findings. Level 1 study involves land cover classification, whereas the study of land degradation processes is conducted at level 2 of the hierarchy. Level 3 shows the severity level of degradation. Generally, as we move further in hierarchical levels, attention to detail increases. For example, one of the classes in level 1 may be agriculture, which represents a broad category. It is a general classification that encompasses various land covers. Moving up to hierarchy level 2, the focus narrows down to the degradation process itself. This level examines the specific mechanisms and factors contributing to land degradation for the level 1 class. It may involve studying aspects such as soil erosion, deforestation, wind/water erosion, and any activities or processes that lead to land degradation as shown in the table.

Up-to-date and accurate land cover information is crucial to many resource monitoring, planning and management programs. Several ecological models make use of such information. Conventional methods of land use classification rely on inefficient, uneconomical and time-consuming methods like field surveys. These methods are very impractical when the data is required for instantaneous global mapping or extending over large areas of land. Satellite images are a reliable source for obtaining land cover information due to their concise view and periodical coverage. Here, the pixels of the remote sensing image are grouped into meaningful classes, each class representing a land cover type, also known as digital image classification. Different approaches for level 1 land cover classification were performed including Pixel-based classification, Object-based classification,

and Convolution Neural network.

Table 2.1: Three-level hierarchical desertification/land degradation classification system

| Level 1 | Level 2 | Level 3 |
|---|--|--|
| Land Cover Detection | Process of Degradation | Severity of process |
| Agriculture Forest / Plantation Grassland/ Grazing land | Water Erosion Wind Erosion Water Logging | Based on temporal thresholding and change severity will be decided |
| Scrub land | Salinization/ Alkalinization | |
| Barren / Rocky Area Water body / Drainage Urban | Man made Vegetation Degradation | |

The identification of land degradation processes like vegetation degradation, salinity, water erosion, wind erosion, and water logging is a challenging task. This is because of the varying information required to identify different degradation processes, which are not easily available. Vegetation degradation is prevalent in forest areas that are accessible and occur in areas where human population density is high. Assessment and monitoring of such regions are crucial to help create conservation strategies and save the region's biodiversity. Romero-Sanchez et. al. demonstrates the use of above-ground biomass calculated from multispectral data to obtain thresholds for areas to be considered as degraded. [134] According to the study, above-ground biomass shows a positive correlation with most vegetation indices like NDVI(Normalized Difference Vegetation Index). In another study, Connette et. al. have shown a comparison of the classification-based machine learning models to map degradation in different forest types[37]. This thesis used two different vegetation indices and pixel-based classification models like SVM and random forest to identify the regions that have undergone vegetation degradation. Comparison in the performance of the two models and validation of the results are shown in subsequent sections. Another degradation phenomenon that

is affecting the environment, especially agricultural lands is salinity. Poorly vegetated areas usually characterize areas affected by saline soils. Thus, knowledge about the extent of degradation due to soil salinization is crucial for the planning and implementation of effective soil improvement strategies.[11] In most studies salinity is assessed and mapped based on the band ratios in visual, near-infrared, and short-wave infrared spectral ranges.[11, 117] In [37], different indices were obtained from multispectral images, showing an encouraging statistical correlation between soil electrical conductivity (EC) measurements conducted in the field and spectral indices calculated from remote sensing data. The study also showed that the salinity index had the highest correlation with soil electrical conductivity. Another study conducted by Asfaw et. al. demonstrates that there was a strong correlation between the sodium and chlorine contents of certain plants and the soil's spectral reflectance. Thus showing that reflectance spectroscopy methods can be used to monitor and delineate soil salinity in growing vegetation, especially agricultural lands[11].

2.1 Study Area

Two distinct areas were chosen to serve as explanations or proofs of concept for the methodology, one with vegetation degradation at Panchmahal and other soil salinity in the Bhavnagar district, located in the central and southeastern parts of Gujarat, India. These are hot semi-arid regions with dry climates. Satellite images were used for land use classification at level 1 followed by the classification of land degradation processes at level 2.

Bhavnagar district lies in the southeast corner of the Gujarat state. It is bounded in the north by Surendranagar and Ahmadabad districts, in the west by Rajkot and Amreli districts and in the south by the Arabian Sea and part of Amreli district, and in the east by the Gulf of Khambhat. It occupies an area of 8334 sq. km. Bhavnagar district has a population of 2,880,365 with 574 population density, 933 sex ratio and a literacy rate of 75.5%[30]. Bhavnagar district forms a part of the Kathiawar Peninsula and is subdivided into four sub-micro regions, namely,

Bhavnagar Coastal Plain, Palitana Savarkundla upland, Songadh Forested Plain and Keri, Kalubhar and Ghelo Plain on the basis of topography, climate, geology, soils and natural vegetation. The main rivers which drain this district are the Kalubhar river, Bagad river, and Gomati river. Bhavnagar is observed with 35.64% of total geographical area under land degradation/ desertification for the period 2011-13. The area under land degradation/ desertification in the district has decreased about 3.07% since 2003-05. The most significant process of land degradation/ desertification in the district is Vegetation Degradation (11.46% during 2011-13 and 11.49% during 2003-05) followed by Water Erosion (10.86% during 2011-13 and 10.79% during 2003-05)[136].

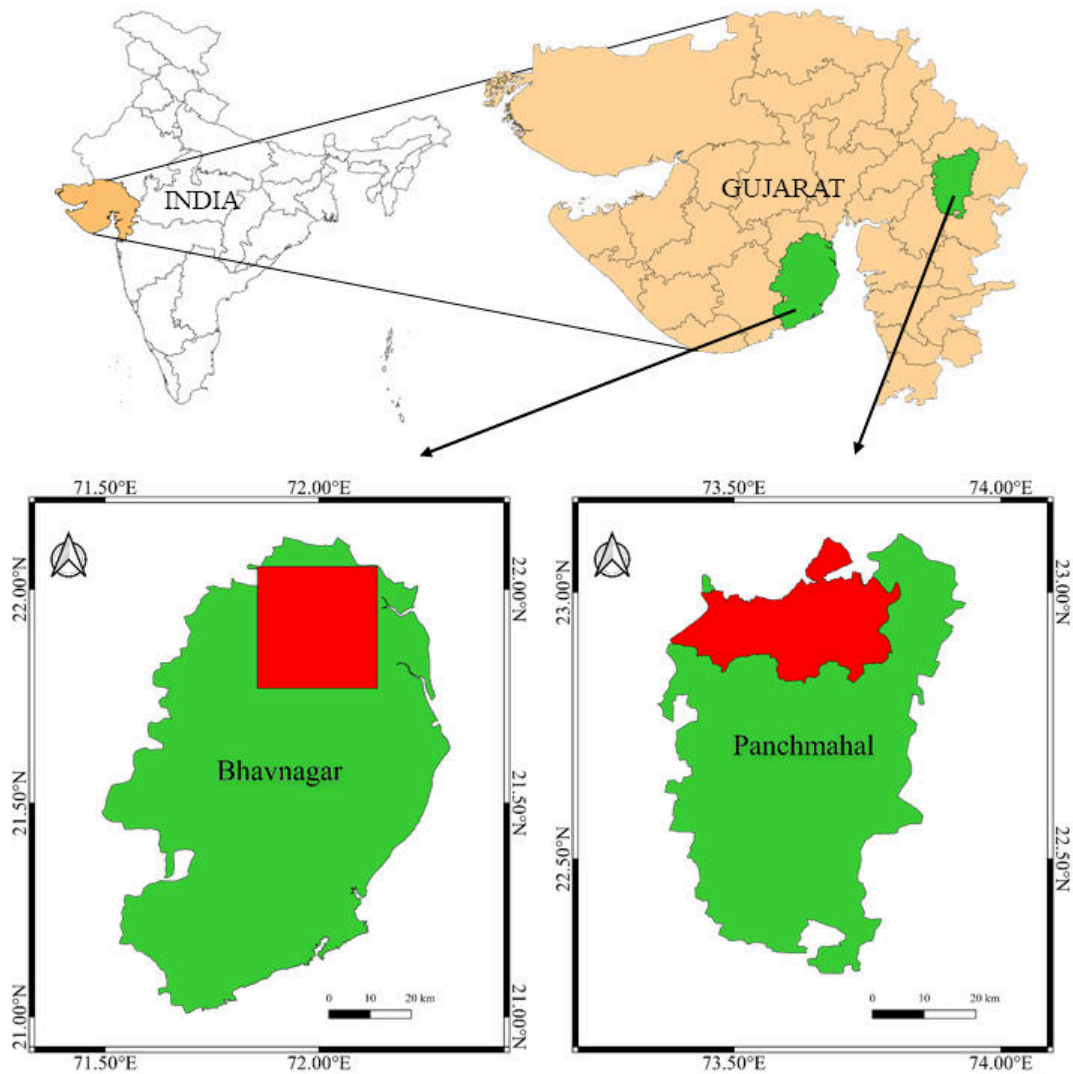


Figure 2.1: Study Area shown in as red polygon in Bhavnagar district and Panchmahal district

Panch Mahals district falls in the eastern portion of Gujarat State. It is bounded on the north by the Sabarkantha district and Rajasthan State, on the east by Dohad district, Vadodara district to the south and on the west by Kheda and Anand districts. It covers an area of 5231 sq. km area. The district has a population of 2,390,776 with 457 population density, 949 sex ratio and a literacy rate of 70.99%[30]. Topographically, the district is divided into two major zones hills and plain. The northern and eastern parts of the district are hilly whereas the southern and western areas of the district are plain. The general elevation of the district is between 100 meters and 400 meters from mean sea level. The district is rich in water resources. It is drained by several rivers like Mahi, Goma, Kun, Panam, Karad, Kali and Meshri. Mahi is the longest river.

Panch Mahals is observed with 52.07% of total geographical area under land degradation and desertification for the period 2011-13. The area under land degradation and desertification in the district has decreased about 0.22% since 2003-05. The most significant process of land degradation/ desertification in the district is Vegetation Degradation (40.42% during 2011-13 and 41.28% during 2003-05) followed by Water Erosion (11.03% during 2011-13 and 10.54% during 2003-05)[136].

2.2 Dataset

Satellite Data

Multi-temporal digital IRS LISS-III data, ancillary information, collateral data and forest cover layer of Forest Survey of India (FSI) were used. IRS LISS-III is 10 bits data with 23.5 meters spatial resolution, 24 day repeativity, swath of 141 km in four spectral channels, i.e. 520-590 nm (Green), 620-680 nm (Red), 770-860 nm (NIR) and 1550-1700 nm (SWIR)[32]. Individual bands in Geo-TIFF format for each single date were stacked into a single file False Color Composite (FCC) prepared using first three bands. All the spectral bands of three season for year 2011 and 2019, LISS data were used for the land cover classification(Table:2.2).

Table 2.2: Satellite Data Description

| | Bhavnagar | | Panchmahal | |
|----------|------------|------------|------------|-------------|
| Row_Path | 93_57 | | 94_56 | |
| | 2011 cycle | 2019 cycle | 2011 cycle | 2019 cycle |
| Rabi | 26-Feb-12 | 20-Jan-19 | 07-Feb-12 | 1-Jan-19 |
| Summer | 26-Apr-12 | 26-May-19 | 20-May-12 | 6-May-19 |
| Kharif | 11-Oct-12 | 9-Nov-2018 | 04-Oct-12 | 14-Nov-2018 |

Ground Truth Data

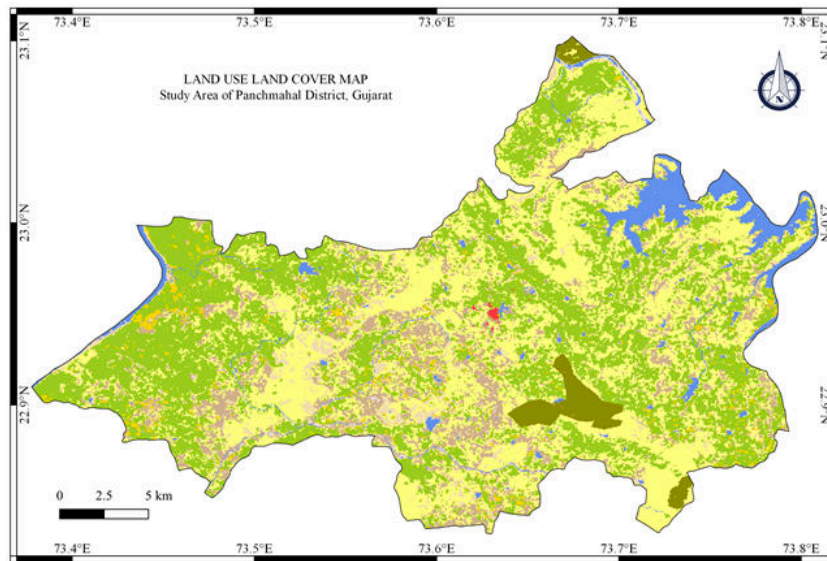
Ground truth data was used to train machine learning algorithms efficiently and further, validating the results obtained along with multi-spectral satellite images. Land use land cover map having 16 classes as seen in (Fig.2.2) published by Bhuvan-ISRO was used as level 1 ground truth dataset[118].

The Desertification Status map (DSM) cycle 2011 and 2019 published map were used as ground truth dataset for level 2 class(Fig. 2.3)[136]. In order to evaluate the classes of Desertification and land degradation, geo-coded LISS III digital data were analysed utilising on-screen visual interpretation techniques combined with auxiliary data. In a geographic information system (GIS) context, preliminary DSM with a scale of 1:50,000 were created for each district. Based on the National Spatial Framework on 1:50K with LCC projection and WGS 84 datum, a geodatabase was produced in GIS[154].

2.3 Methodology

The methodology is discussed into three components, namely - land cover classification, the process of degradation, and severity of degradation. Level 1 involves land cover classification using multi-spectral remote sensing images. The study was based on reflectance values using different classification techniques. The aim is to classify each pixel into one of the 6 land cover classes (Waterbody, Settlement area, Scrubland, Agricultural land, and Forest area, Barren land).

In Level 2, the aim is to detect degradation processes like salinity, erosion, water logging, etc that affect these land types. For the purpose of this thesis, I have



| Value | Description | Color | Red,Green,Blue |
|-------|-----------------------|--------------|----------------|
| 1 | Built-up | Red | 255,0,0 |
| 2 | Kharif Crop | Yellow | 255,209,0 |
| 3 | Rabi Crop | Orange | 255,158,0 |
| 4 | Zaid Crop | Brown | 158,81,43 |
| 5 | Double/Triple Crop | Light Green | 158,207,31 |
| 6 | Current Fallow | Light Yellow | 245,245,219 |
| 7 | Plantation | Green | 0,204,0 |
| 8 | Evergreen Forest | Dark Green | 0,94,0 |
| 9 | Deciduous Forest | Olive Green | 107,120,31 |
| 10 | Degraded/Scrub Forest | Light Green | 115,184,43 |
| 11 | Littoral Swamp | Dark Green | 5,130,94 |
| 12 | Grassland | Light Green | 184,234,120 |
| 13 | Shifting Cultivation | Purple | 158,31,235 |
| 14 | Wasteland | Light Brown | 209,181,133 |
| 15 | Rann | Grey | 191,191,191 |
| 16 | Waterbodies max | Light Blue | 94,209,242 |
| 17 | Waterbodies min | Dark Blue | 0,158,222 |
| 18 | Snow Cover | Pink | 255,191,196 |

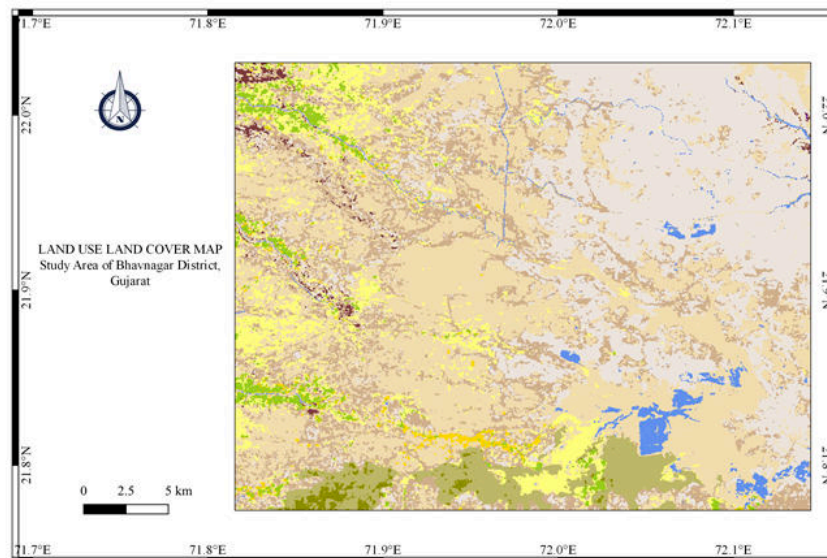
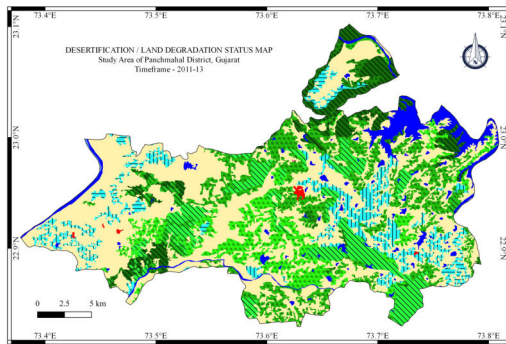
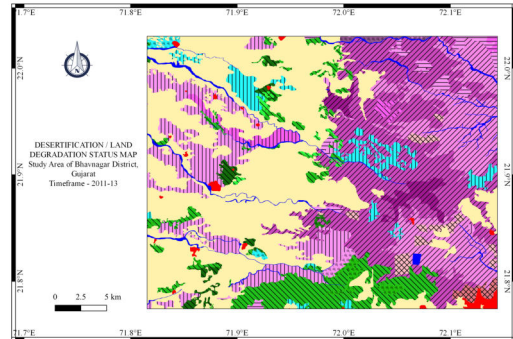


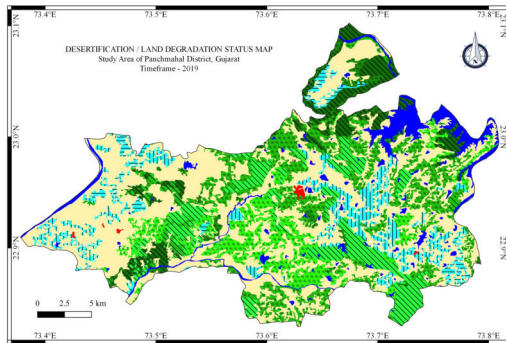
Figure 2.2: Land Use Land Cover map of Study area of Panchmahal (above) and Bhavnagar (below) Districts with the legends showing 16 classes and their color code



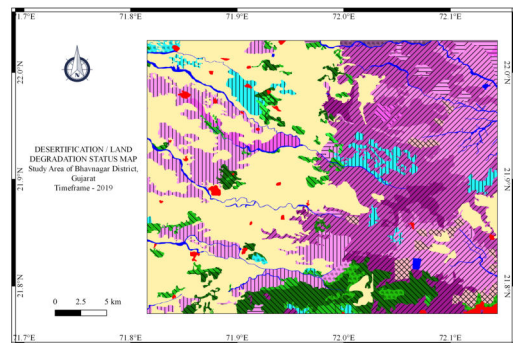
(a)



(b)



(c)



(d)

Legend

Symbol Code Description

| | | |
|--|-----|--|
| | B | Barren |
| | Bs1 | Barren, salinity / alkalinity, Slight |
| | Bs2 | Barren, salinity / alkalinity, Moderate |
| | Bs3 | Barren, salinity / alkalinity, Severe |
| | Ds1 | Agriculture unirrigated, salinity / alkalinity, Slight |
| | Ds2 | Agriculture unirrigated, salinity / alkalinity, Moderate |
| | Dw1 | Agriculture unirrigated, water erosion, Slight |
| | Fs1 | Forest, salinity / alkalinity, Slight |
| | Fs2 | Forest, salinity / alkalinity, Moderate |
| | Fv1 | Forest, vegetation degradation, Slight |
| | Fv2 | Forest, vegetation degradation, Moderate |
| | Fv3 | Forest, vegetation degradation, Severe |
| | Fw1 | Forest, water erosion, Slight |
| | Fw2 | Forest, water erosion, Moderate |
| | Is1 | Agriculture irrigated, salinity / alkalinity, S |
| | Is2 | Agriculture irrigated, salinity / alkalinity, M |

Symbol Code Description

| | | |
|--|-----|---|
| | Iw1 | Agriculture irrigated, water erosion, Slight |
| | Iw2 | Agriculture irrigated, water erosion, Moderate |
| | Iw3 | Agriculture irrigated, water erosion, Severe |
| | NAD | No Apparent Degradation |
| | S | Settlement |
| | Ss1 | Land with scrub, salinity / alkalinity, Slight |
| | Ss2 | Land with scrub, salinity / alkalinity, Moderate |
| | Ss3 | Land with scrub, salinity / alkalinity, Severe |
| | Sv1 | Land with scrub, vegetation degradation, Slight |
| | Sv2 | Land with scrub, vegetation degradation, Moderate |
| | Sv3 | Land with scrub, vegetation degradation, Severe |
| | Sw1 | Land with scrub, water erosion, Slight |
| | Sw2 | Land with scrub, water erosion, Moderate |
| | Tm1 | Others, man made, Slight |
| | Tm2 | Others, man made, Moderate |
| | Tm3 | Others, man made, Severe |
| | W | Water body/ Drainage |

Figure 2.3: Desertification Status Map of Study area for two-cycle over 10 years apart

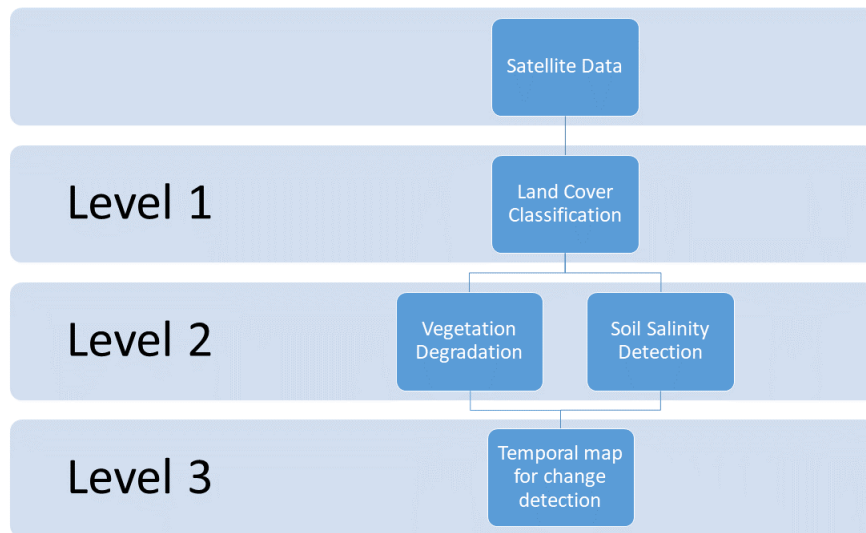


Figure 2.4: Methodology flow chart

worked on two processes, namely soil salinity in agricultural land and vegetation degradation in forest areas. The result is a land degradation process map distributed in each of the level 1 classes. Level 2 study was based on extraction of features from satellite images and using them to further assess degradation processes. In level 3, satellite dataset of year 2019 was used for level 1 and level 2 classification system. And the total area was compared with 2011 map for severity classification.

Level 1 : Land Cover Classification

Land use land cover mapping involves assigning each pixel of the satellite data to a particular class. Each class represents a land cover type. In this thesis the pixels were classified into the following six land cover classes - Agricultural land, Forest area, Scrubland, Barren land, Settlement, Waterbody. Three Classification algorithms were attempted,

1. Pixel based classification - Support Vector Machine
2. Convolution Neural Network
3. Object Based Classification

1. Support Vector Machine

SVM is one of the widely used algorithms for land cover classification using remote sensing data. SVM model has the ability to perform well even with limited training data, making it a viable option in remote sensing use cases.[113]

SVM works by locating the ideal hyper-plane to classify data points into distinct groups. This division is based on the idea of structural risk minimization (SRM), which maximizes and divides the hyper-plane and data points closest to the hyperplane's spectral angle mapper (SAM). Both continuous and categorical variables, as well as linear and non-linear samples with various levels of class membership, can be handled by SVM. The vectors used in the SVM process ensure that the margin between the various classes is maximized. Support vectors are the training samples that define the SVM's hyper-plane or margin. The hyper-plane is defined by the equation,

$$wx + b = 0 \quad (2.1)$$

where, w is a vector of weights, x is a data point, b is a bias term.

The kernel function, a crucial part of SVM, is used to precisely establish the hyper-planes and reduce classification errors. The most popular remote sensing kernel is the polynomial and radial basis function (RBF) kernel, with the RBF technique being the most widely used for classifying land use and land cover (LULC) because it is more accurate than other conventional techniques[101, 103]. The Radial Basis Function (RBF) kernel is defined as:

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (2.2)$$

Here, x and x' are feature vectors, and γ is the kernel parameter that controls the shape of the kernel. The RBF kernel measures the similarity between data points in the feature space.

The decision function for making predictions using the trained SVM with the

RBF kernel is given by:

$$f(x) = \sum_i \alpha_i y_i K(x_i, x) + b \quad (2.3)$$

Here, b is the bias term.

To do with the problem of skewness and overfitting, a data augmentation technique called the smote algorithm (Synthetic Minority Oversampling technique) is applied to the minority class in the training data [34]. This helps in balancing the ratio of training data belonging to each class. The SMOTE algorithm works by selecting the data points that are close to the feature space.

2. Patch Based Convolution Neural Networks

The pixel values are scaled between zero and one. This is achieved using a min-max scalar. For the purpose of the patch-based model, instances are generated which are equal to the number of pixels in the original data. This involves the extraction of small three-dimensional patches of fixed size, centered at each pixel. This helps in taking into account the information of the spatial neighborhood and not just a single pixel. [29]

Here, the aim is to classify the central pixel by using the local spatial features from neighboring pixels. This model works well because it is observed that pixels with the same neighborhood show similar underlying objects. The size of the patches depends on the extent and resolution of the original image. It was found that a patch size of 5 works best for the current use case.

Rotation-based data augmentation techniques were carried out to allow the model to learn varying spatial distributions. Also, since the data is skewed, this technique helps in increasing the number of data points, especially for classes with the least data points, thus reducing the problem of overfitting. CNNs learn representations of original data with varying levels of understanding, thus acting as feature extractors. [29] Several layers like the convolutional layer, non-linearity layer, max-pooling layer, and fully connected layers are included in the model. Figure 2.6 shows the CNN architecture and the corresponding parameters used to train the classification model.

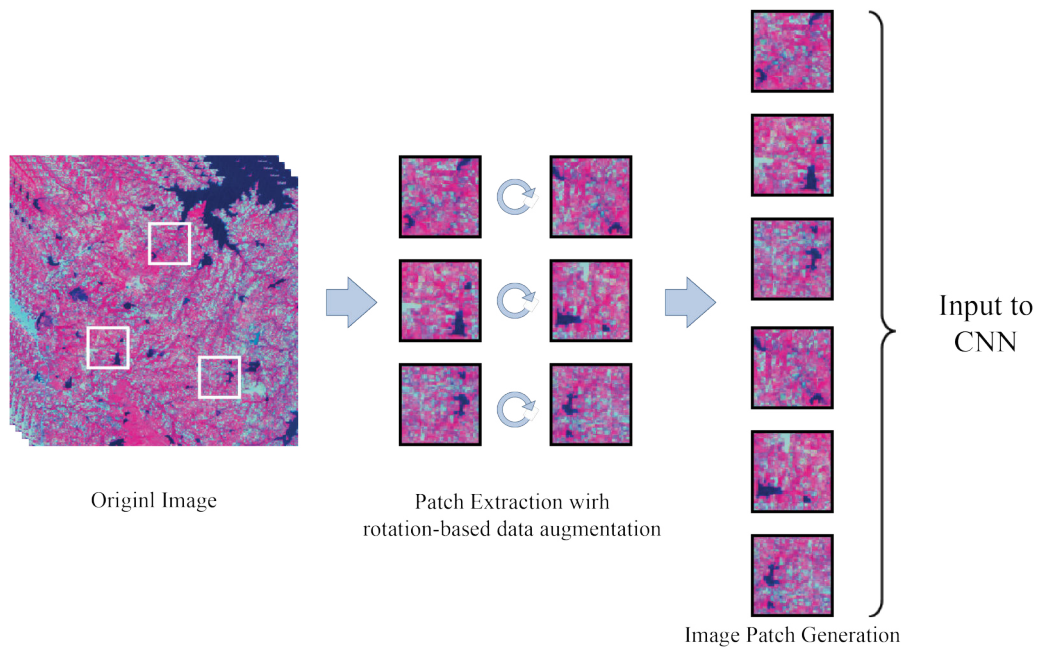


Figure 2.5: The rotation-based augmentation process generates the patch that would be fed to the CNN model

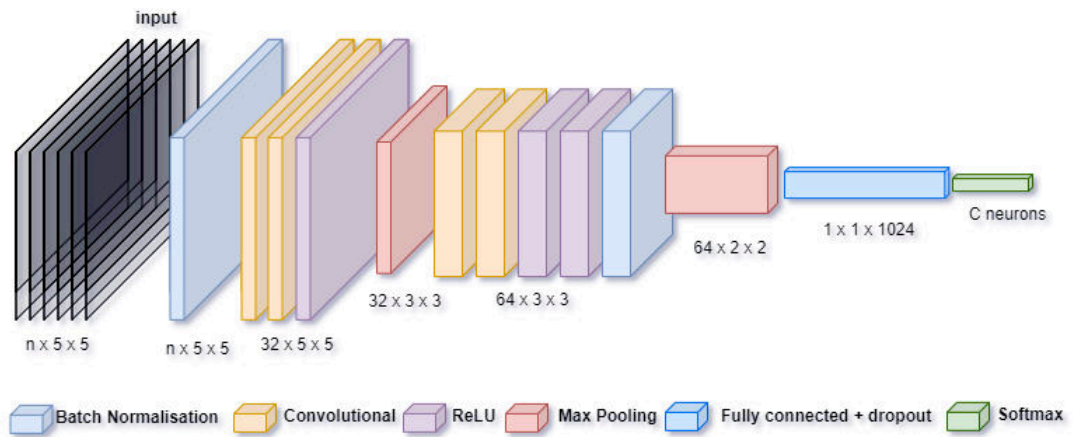


Figure 2.6: Patch Based CNN Architecture

- The input data in the convolutional layers is convolved with m 2D kernels, where m is the desired number of convolutional filters. The sum of m convolutional responses produces a feature map. Next, this feature map is passed through the rectified linear unit (ReLU), a nonlinear activation function.
- The max-pooling layers reduce the number of parameters, computations, and help in controlling overfitting by selecting superior invariant features.
- The fully connected layers have connections to all the activations in the previous layers. The results from convolution are converted into a 1D feature vector. The last layer of the network performs classification using the softmax activation function. The output thus obtained gives a probability distribution that represents the chances of belonging to a particular class.

3. Object-based classification

Fig.2.7 shows the flowchart of the model for object-based classification. It is divided in two phases. In the first phase called as object definition phase, three different segmentation methods eg. Region Growing segmentation, Mean-shift segmentation, and Simple Linear Iterative Clustering (SLIC) algorithms were used to create objects, and a random forest classifier is used to assign classes to these objects in second phase as object classification phase.

Phase 1: Object Definition Phase Phase one is for the object definition phase in which different segmentation algorithms were used for defining objects based on the reflectance value of single date LISS-IV bands. It is a process of grouping pixels in which the intent is to simplify the image into meaningful pixel groupings called segments or objects. Segments are relatively homogeneous regarding one or more characteristics like shape, size, color, texture etc.[19] Three different segmentation methods eg. Region Growing segmentation[19], Mean-shift segmentation[122], and Simple Linear Iterative Clustering (SLIC) Segmentation[4] used to create objects.

The use of segmentation for LULC classification has gained attention in recent years[180]. Several studies have proposed object-based and improved region-

growing algorithms combined with machine learning techniques, such as random forest and deep learning, for LULC classification[114, 53]. Approaches have shown promising results and can potentially improve the accuracy and efficiency of LULC classification in various applications, including environmental monitoring, natural resource management, and urban planning[56].

These studies provide valuable insights into the application of segmentation techniques for LULC classification using satellite imagery and highlight the importance of selecting appropriate techniques for accurate classification.

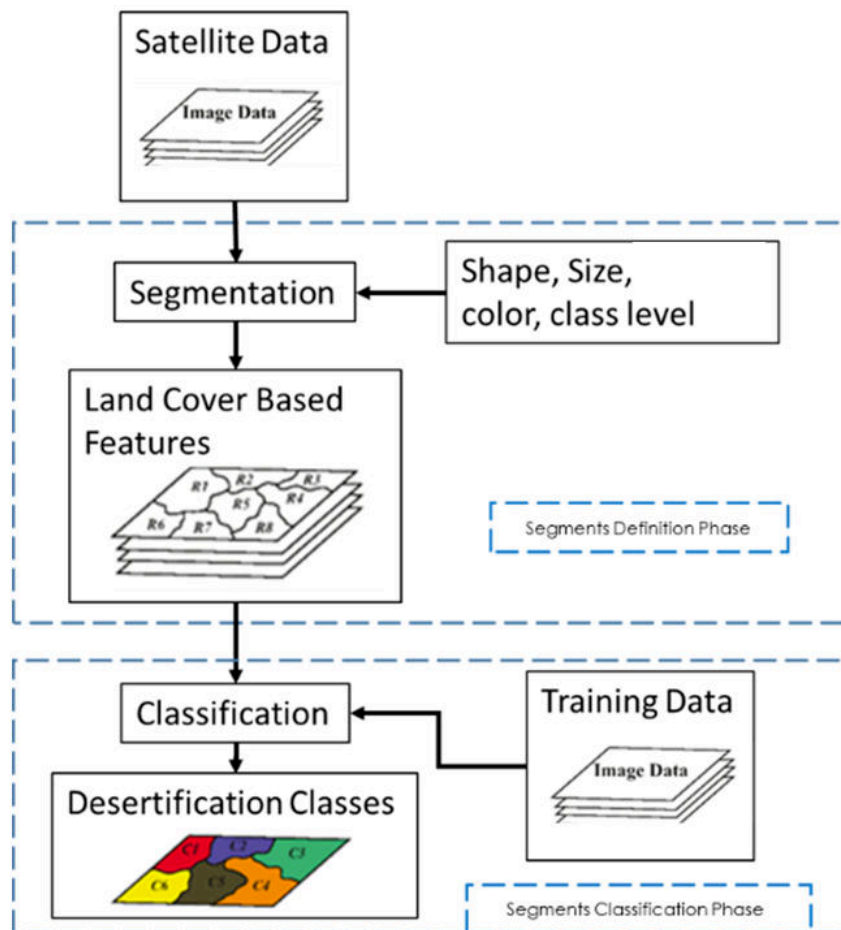


Figure 2.7: Object Based classification model flowchart

1. Region Growing Segmentation

A region-growing algorithm-based segmentation was performed. Two parameters, "similarity" and "area", were used to control the segmentation procedure. To clarify, "similarity" is a threshold value that determines whether

two neighboring objects are merged, while the area threshold is used to filter out the objects smaller than the established value [19]

In this method, the set of regions in the image is denoted by the symbol \mathbb{R} , and an individual region is represented by $R \in \mathbb{R}$. The threshold value below which two regions are considered similar at a particular instant t is denoted by $T(t)$. The mean value vector of a region is represented by M_i , and the Euclidean distance between the spectral mean values of two regions R_i and R_k is represented by $D(R_i, R_k) = \|M_i - M_k\|$. The set of neighboring regions of a region R is denoted by $N(R)$. The segmentation process involves the following stages:

- Initially, a list of regions $R_i, i = 1, \dots, n$ ($n = \text{number of pixels in image}$) is created, with each region composed of a seed pixel. For each region R_i , its mean value vector and neighboring regions are stored.
- For each region R_i , its neighboring regions $N(R_i)$ are examined, and the most similar neighboring region $R_k \in N(R_i)$ is chosen. If the similarity between the two regions is below the threshold value $D(R_i, R_k) < T(t)$, the neighboring region is considered the best neighbor of the current region.
- If the best neighbor of a region exists, the two regions are merged, and the best neighbor is removed from the list.
- The mean value vector of the resulting region is updated every time two regions are merged.
- The process is repeated until no joinable regions remain.
- In the final step, small regions are merged with larger adjacent regions, according to an area threshold value defined by the user.

2. Mean-Shift Segmentation

Mean-shift segmentation Mean-shift algorithm was introduced by Fukunaga et.al.[60]. It is a versatile, non-parametric density gradient estimation for mode finding or clustering procedure.

For each data point, the mean shift defines a window around it and computes the mean of the data point. Then it shifts the center of the window to the mean and repeats the algorithm till it converges. After each iteration, we can consider that the window shifts to a denser region of the dataset. At the high level, we can specify Mean Shift as follows:

- Initialize a random seed point and fix a window.
- Compute the mean of data within the window.
- Shift the window to the mean and repeat till convergence

3. Simple linear iterative clustering (SLIC)

Simple Linear Iterative Clustering, SLIC is similar to the k-means and mean shift algorithms regarding being both gradient-ascent-based algorithms and unsupervised data clustering algorithms. SLIC creates dense and homogeneous superpixels with a minimal effort and complexity. It also provides control over the number and compactness of superpixels, features that are highly preferable. Thus, it is one of the most commonly used algorithms for image segmentation. We can consider SLIC as a special version of the k-means with two crucial differences, which provide major enhancements in the performance. First, in the k-means algorithm, the search region is the whole feature space meaning every centroid is compared to every point/pixel in the space resulting in relatively higher complexity. In contrast, SLIC performs local clustering in which a limited region is defined as the search space contributing to lowering the computational cost. Second, SLIC introduces a new distance measure that takes into account not only the spatial but also spectral distance offering at the same time the feature of controlling the number and compactness of the superpixels[190, 195]

- The distance measure is constructed by a v -dimensional vector including the spectral feature vector $(band_1, band_2, \dots, band_{v-2})$ and the pixel position coordinates (x, y) .
- Suppose the total number of pixels of the image is N , and it is planned to be divided into the same number of segments K , thus the size of each

segment is N/K and the distance between every two adjacent segments $S \approx \sqrt{\frac{N}{K}}$.

- Choose K cluster centers $C_k = [band_{1_k}, band_{2_k}, \dots, band_{(v-2)_k}, x_k, y_k]^T$ where $k = [1, K]$ at regular grid intervals S . The computational formula of distance measure is

$$d_{spatial} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$$

$$d_{spectral} = \sqrt{(band_{1_i} - band_{1_k})^2 + \dots + (band_{(v-2)_i} - band_{(v-2)_k})^2}$$

$$D_s = d_{spectral} + \frac{m}{S} d_{spatial}$$

where, D_s denotes the normalized distance measure; $d_{spectral}$ denotes the Euclidean distance of the spectral feature vector $(band_1, band_2, \dots, band_{v-2})$, $d_{spatial}$ denotes the Euclidean distance of pixel (x_k, y_k) and pixel (x_i, y_i) in the image plane; m is a variable to control the compactness of a segments.

The greater the value of m , the more important spatial proximity and the more compact the cluster. Experience indicates that the optimal range of values for m is $[1,20]$, which can get a good tradeoff between spatial similarity and spatial proximity[4].

Phase 2: Object Classification Phase Phase 2 is for the classification of objects and was performed using RF classifier with a train-test ratio of 0.7:0.3. Random Forest is an ensemble machine learning algorithm that is used for both regression and classification problems. The key idea behind Random Forest is to build multiple decision trees and combine their predictions to get a more accurate and stable result. The Random Forest algorithm is known for its ability to handle large datasets with high dimensionality and for its robustness to overfitting. For classifying object/segments multivariate temporal four bands LISS data were used. In classification, majority class values were selected as final object class. Classification is performed on objects instead of pixels. Algorithms operate on many object-related features than typically available with pixel-based approaches. Reduces salt and pepper effect of classifications and speeds up processing.

Level 2: Degradation Process classification

Feature-based classification is performed using pixel-based Random Forest and Support vector Machine classifier to identify different degradation processes.

Vegetation Degradation in Forest Area

As a result of deforestation and/or overgrazing, vegetation degradation is defined as a decrease in biomass and/or a decline in the vegetative ground cover. With relation to soil erosion and the loss of soil organic matter, such degradation is a significant contributor to land degradation. The protection of land and the fertility of the soil are significantly influenced by vegetation. Vegetation degradation speeds up soil degradation, which then speeds up land degradation. When organic material is lost, the soil's ability to store water and its nutrient content are decreased, which puts further stress on vegetation viability. Agriculture practiced on forestland is likewise categorized as contributing to vegetation degradation.

Southeastern and central parts of Gujarat have dry deciduous forest cover. Forest cover in Shehra, taluka of Panchmahal district is classified into 3 categories, namely- Scrub forest, Open forest and Moderate dense forest. The Forest Survey of India map was used as a training dataset. The classification of these forest types is shown in table 2.3. Forest area in Shehra Taluka is selected for vegetation degradation.

Level 2 classification refers to a more specific categorization or classification system used for identifying vegetation degradation in a forest. It is described as a feature-based classification.

Feature-based classification means that certain features or characteristics of the vegetation are used as indicators or criteria for categorizing and identifying the degradation. In this case, two specific features or indices are used: NDVI (Normalized Difference Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index)[134].

Normalized Difference Vegetation Index (NDVI)[169]:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2.4)$$

Green Normalized Difference Vegetation Index (GNDVI)[61]:

$$GNDVI = \frac{NIR - Green}{NIR + Green} \quad (2.5)$$

Here, NIR represents the reflectance value in the near-infrared region, Red and Green represents the reflectance value in the red and green region respectively.

Table 2.3: Vegetation degradation in forest area classification scheme

| Forest Type | Classification scheme |
|-----------------------|--|
| Scrub Forest | Forest areas with poor tree growth having canopy cover less than 10% . |
| Open Forest | Forest areas having canopy cover between 10% and 40%. |
| Moderate dense Forest | Forest areas with tree cover having canopy cover between 40% and 70%. |

NDVI is a widely used index in remote sensing and vegetation studies. It measures the difference between the reflectance of near-infrared (NIR) and red light, providing an estimate of the vegetation greenness and health. NDVI values range from -1 to 1, with higher values indicating healthier and more abundant vegetation[76, 44].

GNDVI is similar to NDVI but focuses specifically on the green band of the electromagnetic spectrum. It measures the difference between the reflectance of green light and NIR, providing a more targeted assessment of the vegetation's greenness and health[133, 42].

By utilizing these two indices (NDVI and GNDVI), it becomes possible to analyze and classify the vegetation in a forest area.

Salinity in Agriculture Area

Alkalinity or salinity is the primary chemical characteristic of soils. The salt that can dissolve in water that is present in the soil is known as salinity. Salinity can arise naturally or as a result of human activity. It occurs primarily in culti-

vated regions, particularly in the areas that are irrigated. While alkalization is not always visible and is typically assumed based on ground truth, soil sample analysis, and information or published maps, there are instances where salinity is readily visible on satellite pictures. Excessive evapotranspiration, drought, excessive irrigation, and an increase in toxicity are the main causes of salinity. By means of capillary action, salts from the groundwater are brought to the soil's surface. As time passes, water evaporation leaves the salt on the soil's surface.

Three soil salinity indices, Normalized Differential Salinity Index(NDSI), Salinity Index(SI), Vegetation Soil Salinity Index (VSSI) [84, 45] were identified and calculated as,

Normalized Differential Salinity Index (NDSI):

$$NDSI = \frac{RED - NIR}{RED + NIR} \quad (2.6)$$

Salinity Index (SI):

$$SI = \sqrt{NIR * Red} \quad (2.7)$$

Vegetation Soil Salinity Index (VSSI):

$$VSSI = 2 * GREEN - 5 * (RED + NIR) \quad (2.8)$$

Here, NIR represents the reflectance value in the near-infrared region, RED and GREEN represents the reflectance value in the red and green region respectively.

VSSI takes into account both the soil salinity characteristics captured by NDSI and the vegetation-related information represented by SI, providing a comprehensive index for assessing soil salinity while considering the vegetation's influence.

These equations utilize reflectance values in specific spectral bands, such as NIR, red, and green, to calculate the respective indices. By applying these indices to multi-temporal LISS-III data, the study aimed to map and analyze soil salinity patterns in Bhavnagar district, Gujarat.

Level 3: Severity status

The severity of a degradation process can be determined by comparing two different time frames of data, here the degradation process' severity was determined using data from 2011 and 2019. To evaluate the severity of the degradation process, the total area affected over both time frames was calculated and compared. The threshold for percent area change to classify its severity level is shown in Table 2.4.

Table 2.4: Percentage area under different severity conditions for desertification

| Change in Area | Severity Level |
|----------------|----------------|
| < 10% | Low |
| 10% – 20% | Medium |
| >20% | High |

For example, the severity level is considered low if the percentage change is less than 10%. Any percentage change between 10% and 20% is considered significant. The severity level is high if the percentage change exceeds 20%. It is simpler to categorize the severity level of the degradation process based on the percentage change in the affected area by using these thresholds. Using this knowledge, appropriate mitigation strategies for the effects of the degradation process can then be developed.

2.4 Results & Discussion

Desertification status and severity were calculated using the proposed three-stage algorithm. In level 1 classification of land cover were performed using three different methodology. Object-based image classification is one of the method, in which three segmentation algorithms were used: Region Growing, Mean-shift, and SLIC. Random forest is then used to label the objects created in the segmentation part. Table 2.5 show the accuracy comparison of different segmentation methods for Level 1 land cover classification. Based on the results, SLIC appears to be the most accurate segmentation method for Level 1 land cover classification,

with an accuracy of 80%. The mean shift comes in second with an accuracy of 78%, and Region Growing is the least accurate with an accuracy of 71%.

Table 2.5: Accuracy comparison of different segmentation methods for Level 1 Land cover object creation with Random forest classification (values in %)

| % | Region Growing | Mean-shift | SLIC |
|-----------|----------------|------------|------|
| Accuracy | 71 | 78 | 80 |
| Kappa | 45 | 63 | 68 |
| Precision | 70 | 78 | 78 |
| Recall | 71 | 79 | 80 |
| F1 score | 69 | 78 | 78 |

Table 2.6: Level 1 Land Cover classification accuracy using three different algorithms (values in %)

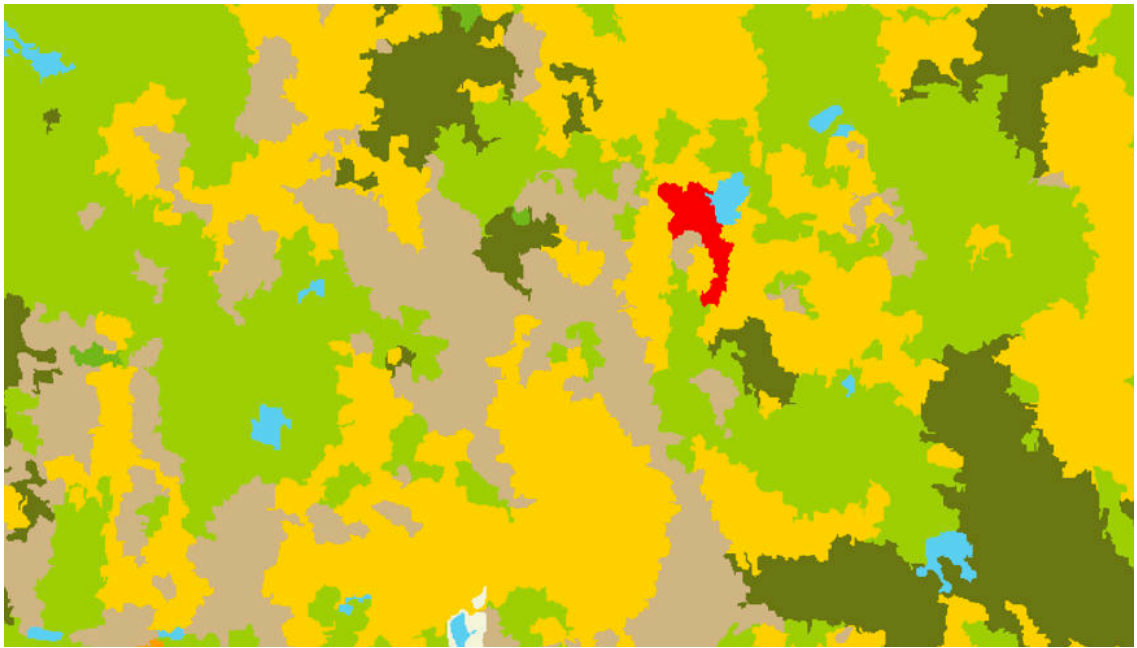
| Model | SVM | Patch based CNN | SLIC + RF |
|-----------|-----|-----------------|-----------|
| Accuracy | 72 | 79 | 80 |
| Precision | 71 | 70 | 78 |
| Recall | 73 | 79 | 80 |
| F1 Score | 71 | 74 | 78 |

Table 2.6 presents the results of three different algorithms for land cover classification: Support Vector Machine (SVM), Patch-based Convolutional Neural Network (CNN), and SLIC (Simple Linear Iterative Clustering) combined with Random Forest (RF). First, it is important to note that all three algorithms achieve relatively high accuracy scores, with the SLIC+RF approach performing the best at 80%. The precision and recall scores for all three models are also quite high, with the SLIC+RF approach again outperforming the other two in terms of precision and recall. Looking at the F1 score, which is a measure that combines precision and recall, we can see that the patch-based CNN approach performs better than the SVM approach, but still falls behind the SLIC+RF approach. It's also worth noting that while the differences in accuracy between the three models are relatively small, the SLIC+RF approach consistently outperforms the other two in terms of precision, recall, and F1 score. SVM is a straightforward model with moderate accuracy, low computational complexity. Patch-based CNN is a more intricate, more accurate model, but it uses more computing power. SLIC+RF

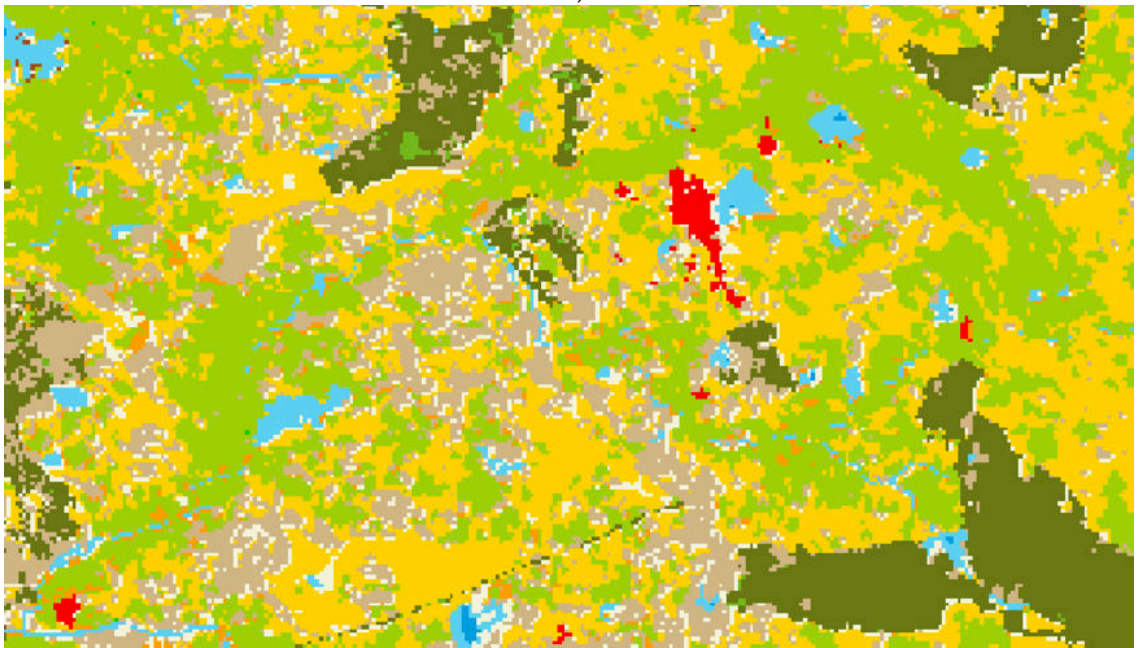
is a hybrid model that first extracts features with a segmentation algorithm before classifying data with a random forest classifier. In comparison to CNN, it achieves high accuracy and is computationally simpler. These results suggest that the SLIC+RF approach is a promising method for land cover classification, particularly in terms of its ability to achieve high precision and recall scores. Object-based classification typically results in more cohesive and smooth polygons representing different land cover classes, as compared to pixel-based classification where each pixel is classified independently and can result in fragmented and irregular patches of land cover. Object-based classification is based on grouping similar pixels together to form objects and classifying these objects based on their attributes, which can result in more accurate and visually appealing land cover maps. Fig.2.8 shows the visual comparison between SLIC with RF classified results with GT class map for level 1 land cover classification.

SLIC segmentation with RF classifier was used as the final model for level 1 classification of land cover for both study areas. The object-based classification results for two different terrains show that the overall accuracy for both areas is good, with Bhavnagar having a slightly higher accuracy (85%) than Panchmahal (80%) as seen in table 2.7. The classification accuracy for Bhavnagar is higher than that of Panchmahal. This could be attributed to the fact that Bhavnagar has a relatively flat terrain with uniform agricultural and open land cover, which makes the classification process less complex and more accurate. On the other hand, Panchmahal has a moderate undulating terrain with heterogeneous land cover, which makes it more challenging to accurately classify the different land cover types. This can be seen from the lower F1 score for Panchmahal compared to Bhavnagar. Therefore, the results suggest that the complexity of the terrain and the heterogeneity of the land cover can significantly affect the accuracy of the object-based classification.

The desertification process was identified in level 2 classification using RF and SVM algorithms. As discussed two different areas having different degradation processes were classified. In the classification of forest vegetation degradation, the models used the vegetation index to identify the extent of vegetation cover.



(a)



(b)

Figure 2.8: Level 1 classified Map using (a) SLIC segmentation comparison with (b) original GT class map for part of study area

Table 2.7: Object based classification result for two different study area(values in %)

| Area | Panchmahal | Bhavnagar |
|-----------|------------|-----------|
| Accuracy | 80 | 85 |
| Precision | 78 | 81 |
| Recall | 80 | 83 |
| F1 Score | 78 | 81 |

Similarly, in the classification of soil salinity, the models used a soil salinity index to identify the affected areas. The SVM and RF models used this index to classify both degradation areas. The classification results show that both models performed well in both areas, with RF showing higher accuracy in both cases. For forest vegetation degradation, the SVM model achieved an accuracy of 79%, while the RF model achieved a higher accuracy of 85%. The precision and recall scores were also higher for the RF model, indicating that it performed better in correctly identifying both positive and negative samples. For soil salinity, both models performed well with the RF model achieving a slightly higher accuracy of 81% compared to SVM's accuracy of 78%. The precision and recall scores were similar for both models, with RF having a slightly higher precision score and SVM having a slightly higher recall score.

Table 2.8: The classification results for forest vegetation degradation for two model performance (values in %)

| Model | SVM | RF |
|-----------|-----|----|
| Accuracy | 79 | 85 |
| Precision | 71 | 85 |
| Recall | 78 | 85 |
| F1 Score | 69 | 84 |

Overall, the use of vegetation and soil salinity indices in classification models provides a more objective and accurate way to identify areas of degradation. These indices can capture subtle changes in the vegetation and soil conditions that may not be apparent to the naked eye. The SVM and RF models were able to leverage these indices to achieve high accuracy in the classification of forest vegetation degradation and soil salinity.

Table 2.9: The classification results for soil salinity for two model performance (values in %)

| Model | SVM | RF |
|-----------|-----|----|
| Accuracy | 78 | 81 |
| Precision | 78 | 79 |
| Recall | 77 | 78 |
| F1 Score | 78 | 79 |

Table 2.10: The degraded areas change for two different year (Area in %)

| Study Area | Degradation Process | 2011 | 2019 |
|------------|------------------------|-------|-------|
| Bhavnagar | Salinity | 13.47 | 12.84 |
| Panchmahal | Vegetation Degradation | 16.74 | 17.28 |

Above level 1 and level 2 classification were performed on two dataset of year 2011 and 2019. As seen in table 2.10, The results indicate that in Bhavnagar, there was a decrease in the extent of salinity from 13.47% in 2011 to 12.84% in 2019. This may be attributed to improved agricultural practices or other irrigation resources in replace of ground water. On the other hand, in Panchmahal, there was an increase in the extent of vegetation degradation from 16.74% in 2011 to 17.28% in 2019. This may be due to various factors such as deforestation for agricultural land or urbanization leading to loss of vegetation cover.

CHAPTER 3

Predictive Soil modeling

The quality and productivity of soil have declined in recent decades due to population growth, climate change, and poor land management practices[24]. Soil degradation causes desertification and severely reduces potential land productivity. Large population growth, food security, land degradation, freshwater scarcity, threatened biodiversity, climate change, and sustainable development are just a few of the major issues the globe is facing in the twenty-first century [58]. A soil's physical and chemical properties affect plant growth and soil management practices. Agricultural production depends on healthy soils that can provide nutrients, support plant growth, and retain water. Soil degradation, which is the process of soil quality degradation, can lead to reduced crop yields and food security concerns[151].

India faces significant challenges in soil management due to its population growth, land degradation, and food demand, as well as the diversity of soil groups across the country[2]. With 86.08% of agricultural land holdings being small and marginal, conventional soil maps are not sufficient to manage the small land holdings[3].

Therefore, predictive soil modeling(PSM) using digital technology can be a valuable tool in assessing soil quality in India. Some important properties of soil are mineral content, texture, cation exchange capacity, bulk density, structure, porosity, organic matter content, carbon-to-nitrogen ratio, color, depth, fertility, and pH.

modeling soils is a crucial first step in decision-making for agriculture, but conventional soil modeling is based on expert knowledge[75] of soil surveyors

that ignore variations within classes and are time-consuming[89]. However, traditional soil mapping methods are labor-intensive, expensive, and primarily rely on taxonomic classification instead of quantifying soil properties. These maps are challenging to understand and utilize for non-soil scientists.

Predictive Soil Modeling (DSM)[106, 65, 140] is a new approach that uses predictive factors to estimate soil properties at locations where sampling has not occurred. Unlike conventional soil maps, digital soil maps provide estimates of accuracy and uncertainty and support precision agriculture[110]. These methods use statistical and digital mapping techniques to quantify and analyze soil properties, allowing for better land management decision-making[10].

For mapping regions with few soil observations, conventional techniques that employed soil maps as the foundation for soil properties estimates are still in use[18]. However, over the past few years, predictive soil modeling technology has advanced quickly, enabling it to be used for routine mapping across vast areas [106, 28, 65, 168], which combines soil survey data with geographic information systems, geostatistics, terrain analysis, machine learning, remote sensing, and high-performance computers to predict soil properties [110, 9]. By using numerical models to infer the geographical and temporal variations of soil types and soil properties from soil observation and knowledge of related environmental variables, PSM is defined by Lagacherie et.al. [90] as "the design and populating of spatial soil information systems."

Current approaches in modeling soil characteristics explicitly account for spatial variation of Jenny's soil forming factors and for possible spatial auto correlation. Such approaches were termed digital soil modeling and formalized in so-called "SCORPAN" models[106], where "SCORPAN" stands for

$$S = f(s, c, o, r, p, a, n) \quad (3.1)$$

where, s = soil, c = climate, o = organisms, r= relief, p= parent material, a = age/time, and n = spatial geographic location information(eg. latitude, longitude & altitude).

This approach has been used with machine learning algorithms for the pre-

diction of soil properties. As traditional methods based on field measurements cannot be performed in every region. ML models can be used to provide better spatial coverage based on limited field data points.

In this study, three important soil properties were selected for the prediction of Soil Organic Carbon (SOC), Soil pH and Soil Electric Conductivity (SEC).

The greatest known repository of terrestrial carbon is soil carbon[17, 152]. Comparing its global storage capacity to the carbon reserves in the atmosphere and plants, it is far greater. Soil organic carbon plays a significant part in many soil biological system administrations. Significant focus has been placed on soil's capacity to store large amounts of atmospheric carbon as a potential means of reducing the negative impacts of rising levels of greenhouse gases in the atmosphere[152, 36, 145]. Soil organic carbon influences the physical and chemical properties of the soil, for example, water permeability, soil moisture holding capacity, supplement accessibility, and microbial organic movement[123]. SOC also plays a dynamic role in the worldwide carbon cycle and environmental change[163]. Soil is a significant wellspring of carbon in earthly environments and manages soil well-being and usefulness.

Soils can be normally acid or alkaline, and this can be determined by testing their pH value. A pH number is really a proportion of hydrogen ion (H_3O^+) fixation. Most soils have pH values somewhere in the range of 3.5 and 10. In higher precipitation regions the regular pH of soils commonly goes from 5 to 7, while in drier regions the reach is 6.5 to 9. Soil pH influences the measure of nutrients and chemical substances that are dissolvable in soil water, and hence the measure of supplements accessible to plants. A few supplements are more accessible under acidic conditions while others are more accessible under alkaline conditions. In any case, most mineral supplements are promptly accessible to plants when soil pH is close to neutral[188].

Soil electrical conductivity (SEC) is a measure of the soil's ability to transmit or attenuate electrical current. It is a measure of the amount of dissolved salts in the soil. EC is measured in deciSiemens per meter(dS/m). Soil scientists have been using EC as an indicator of soil salinity. It is a significant indicator of soil well-

Table 3.1: Ranges for soil sample test values used in soil rating for SOC, pH, and SEC[33]

| Rating | OC range (%) | Rating | pH range | Rating | EC range (dS/m) |
|--------|--------------|------------------------|----------|--------------------------|-----------------|
| Low | <0.50 | Extremely acid | <4.5 | Normal | <1.0 |
| Medium | 0.50-0.75 | Very strongly acid | 4.5-5.0 | Tending to become saline | 1.0-2.0 |
| High | >0.75 | Strongly acid | 5.1-5.5 | Saline | 2.0-3.0 |
| | | Medium acid | 5.6-6.0 | Highly saline | >3.0 |
| | | Slightly acid | 6.1-6.5 | | |
| | | Neutral | 6.6-7.3 | | |
| | | Mildly alkaline | 7.4-7.8 | | |
| | | Moderately alkaline | 7.9-8.4 | | |
| | | Strongly alkaline | 8.5-9.0 | | |
| | | Very strongly alkaline | >9.0 | | |

being. as it affects crop productivity. The presence of salt in the soil is a major reason for soil degradation that can decrease soil fertility. The SEC can be used as a measure of soil salinity[124].

Table 3.1 provides ranges for soil sample test values used in soil rating for three parameters: soil organic carbon (SOC), pH, and electrical conductivity (EC)[33]. These ranges are used to categorize soil samples into different ratings based on their measured values.

3.1 Study Area

North Gujarat Agro climatic zones is selected as study area. Soil of this area is sandy loam to sandy soil. Land productivity is very low in this area; most common crops are wheat, vegetables, spices, oil seeds and tobacco. Rainfall is also average less than 1000 mm annual the climate of this area is arid to semi arid and soil is grey brown alluvial and more than 50% of area is cultivated and out of this area nearly 30% area is irrigated having main source of irrigation is groundwater.

3.2 Soil Sampling

In-situ soil samples were collected from Soil Health Card data from government for the north Gujarat agro-climate zone. Soil health card sampling data were used as soil sample. State agricultural university collected and processed these soil

samples. Total 1840 soil samples were used in this study. Soil samples were extensively collated over more than one year period and all the samples were than lab analysed for soil properties. Nearly 10 soil property information were analysed in lab including soil N, P, K, SEC, pH and SOC with soil micronutrients as well. For this study we used three main soil parameters eg. SEC, pH and SOC. These properties were categorized as per the rating limit given in Table 3.1[1] and Table 3.2 shows the statistics of the selected soil properties.

| Properties | class | Sample | Min | Max | Mean | St. Dev. | Median | Kurtosis | Skew |
|------------|-------|--------|------|------|------|----------|--------|----------|-------|
| OC (%) | Train | 1476 | 0.06 | 1.67 | 0.39 | 0.18 | 0.38 | 5.55 | 1.46 |
| | Test | 369 | 0.14 | 1.72 | 0.43 | 0.19 | 0.40 | 10.50 | 2.33 |
| pH | Train | 1476 | 3.95 | 9.51 | 7.82 | 0.61 | 7.85 | 1.18 | -0.42 |
| | Test | 369 | 6.19 | 9.76 | 7.88 | 0.55 | 7.87 | 0.54 | -0.01 |
| EC (ds/m) | Train | 1476 | 0.06 | 1.76 | 0.37 | 0.25 | 0.29 | 3.59 | 1.78 |
| | Test | 369 | 0.10 | 1.71 | 0.41 | 0.28 | 0.32 | 3.97 | 1.94 |

Table 3.2: Descriptive statistics of soil properties in the topsoil of the north Gujarat Agro-climatic zone

3.3 Covariates selection & preparation

Soil research has shown that soil profiles are influenced by five separate, yet interacting, factors: parent material, climate, topography, organisms, and time. Soil scientists call these the factors of soil formation. These factors give soil profiles their distinctive character. The soil genesis and geographic variation of soils could be explained by the combined activity of the five natural factors called as soil forming factors[51]. These factors form under complex interactions between climate, living organisms and anthropogenic influences, modified by relief and hydrological processes and operating over long periods of time. This has been clearly identified first by Jenny (1994) with his CLORPT factors of soil formation and subsequently extended by McBratney et. al.[106] with the SCORPAN formulation.

The equation 3.1 simply states that the soil type or attribute at an un-visited site (S) can be predicted from a numerical function or model (f) given the factors just described plus the locally varying.[106]

The relevant soil information obtained from remote sensing (RS), proximal sensing, and legacy soil maps can be used to represent soil in DSM. [192].

The DSM activities use the Worldclim database, which provides various climate-related variables such as temperature, precipitation, temperature seasonality, and precipitation seasonality[156, 155, 48, 46, 47, 121]. Lamichhane et al. have demonstrated the usefulness of automatic weather stations for DSM activities[94]. Organism-related covariates include variables related to vegetation, soil fauna, and human activities such as land use and management practices. Vegetation and land use[111, 74, 156]. In recent years, products derived from both conventional and remote sensing (RS) have been used to extract organism-related covariates. Land use land cover (LULC), Normalised Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) are some of the covariates commonly used to represent 'O' in the SCORPAN[156, 155, 47, 149, 74, 92, 121]. As potential covariates, band reflectance values and band ratios from various satellite imageries such as MODIS, Landsat, and Sentinel have been used[92, 47, 196, 197]. Relief (topography) is another important predictor of DSM, particularly at smaller mapping scales. Relief influences soil microclimate and thus soil characteristics. The majority of DSM studies have derived terrain attributes from various Digital Elevation Models (DEMs), such as the SRTM DEM, Cartosat DEM, ASTER DEM, and ALOS DEM[87, 47, 141, 111, 131, 92, 149, 187, 82].

Covariates representing parent material (P) account for very low of all environmental covariates used in DSM in India. Sreenivas et al.[155] and Mitran et al.[111, 155, 47] used geological data from the Geological Survey of India as covariates for DSM. Finally, spatial position (N) has only been used in a few studies. Studies that have considered spatial position as a covariate for DSM[149, 187].

In the present study of predictive soil modeling, the following sets of variables were taken into account as shown in Table 3.3.

Feature Selection is the strategy for reducing the input covariates by utilizing just relevant data and disposing of non relevant data. It is the process of deciding which features to include in your machine learning model according to the kind of issue you are attempting to resolve. It is accomplished by adding or removing

Table 3.3: Soil covariates used for study

| Category | Data description | Source |
|--------------|--|---|
| Soil Maps | Soil Texture, Parent Material | NBSS-LUP data |
| Vegetation | NDVI, FPAR, EVI, LAI | MODIS-Terra Sensor data |
| Climate | Temperature, Rainfall, Soil temperature Wind, Solar Radiation, | Terra Climate data from Google Earth Engine, Terraclimate, a high-resolution global dataset [179] of monthly climate and climatic water balance |
| Relief | Elevation, Aspect, Slope | CartoDEM (Bhuvan- ISRO) |
| Soil Samples | SEC, pH and SOC | Soil Health Card Data (State Agricultural University) |

significant features without altering them. It assists in minimising the amount of noise in raw data and the quantity of input data. The more features or elements there are in a feature set, the harder it is to visualise and operate with the training set. The fact that many of the features are often correlated is another crucial thing to keep in mind. Therefore, if you include every variable in the feature set in your training set, many of them will be redundant. Recursive feature elimination (RFE)[69] is essentially a regressive determination of the indicators. This algorithm begins by building a model in general course of action of covariates and calculating an importance score for each covariate. The most in-critical predictor(s) are then disposed of, the model is re-gathered, and significance scores are figured again. In this manner, the algorithm shows the quantity of covariates subsets to evaluate also every subset's size.

3.4 Methodology

Predictive soil modeling involves a series of steps to generate accurate soil maps. Figure 3.2 shows the flow chart of the methodology used in the study. The process begins with defining the study area and acquiring relevant data, including

soil samples and remote sensing and in-situ data. Data preprocessing is then performed to clean and prepare the stack of data for analysis. Next regression matrix is created by flattening the data and a predictive model is developed using techniques such as regression or machine learning. The model is trained and validated using available data and then applied to predict soil properties or classes across the entire study area. Spatial interpolation may be used to create continuous maps if necessary. The accuracy of the predicted soil maps is evaluated, and the results are interpreted and visualized to identify patterns and areas of interest. The process can be iterated and improved as needed.

Machine learning approaches, like multiple linear regression, random forest regression, Regression kriging, GAM and neural network, were applied for the prediction. Before applying these algorithms soil sampling data were split in two sets. In one set, data were used to develop a model and explore relationships among the covariates and the target soil parameters known as the training set. The second set is the final arbiter of the covariates or model combination performance known as the test set. All 1840 soil samples were separated into two parts, a training dataset and a testing dataset with a 70:30 ratio respectively using random sampling. The training set is utilized to create models and the test set is utilized distinctly at the end of these process for assessing a final and fair evaluation of the model's performance. It is important that the test set not be utilized preceding this point. Machine learning algorithms[106] like multiple linear regression [146, 26, 78], Generalized Additive Models (GAM)[105, 20], Regression kriging [71, 66, 21, 159, 107], Random forest [23, 97]. An Artificial Neural network algorithm was developed from scratch for soil properties prediction. The advantage of the neural network over the other methods is backpropagation. Through the backpropagation concept, the model relearns the weights for each input and predicts a more significant value[59, 96, 50]. Detailed prediction model approaches are discussed below. We are beginning by discussing the primitive ML models on this data set.

Multiple Linear Regression (MLR)

By fitting a line to the observed data, regression models are used to describe relationships between variables. Linear regression is simplest method for prediction. Numerous studies used this method of regression.[146, 26, 78] Using environmental covariates and soil properties the least square relation is to be established and weight matrix is calculated and the same weight matrix is used for unknown locations and soil properties are predicted. Relation between covariates and soil properties is assumed to be linear in this model.[26] To calculate the association between two or more independent variables and one dependent variable, multiple linear regression as shown in eq.3.2 is used.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (3.2)$$

where y = dependent variable predicted value, β_0 = the y-intercept (value of y when all other parameters are set to 0), β_1 = the regression coefficient of the first independent variable (X_1), β_n = the regression coefficient of the last independent variable, ϵ = model error

Generalized Additive Models (GAM)

A GAM is an extension of a generalized linear model that allows for non-linear relations between covariates and the response variable. They provide a modeling approach that combines powerful statistical learning with interpretability, smooth functions, and flexibility. It is made out of a number of smooth functions of covariates.

GAMs is by Trevor et.al.[70] and mathematically represented as:

$$y = \alpha + f_1(X_1) + f_2(X_2) + \dots + f_p(X_p) + \epsilon \quad (3.3)$$

where y = dependent variable predicted value, α = the y-intercept, X_1, X_2, \dots, X_p = the predictor variables, f_1, f_2, \dots, f_p = smooth functions that relate each predictor variable to the response variable, ϵ = model error.

the smooth functions f_1, f_2, \dots, f_p can be modeled using different types of smoothing techniques such as splines loess or generalized cross-validation. These techniques allow the functions to take on any shape including non-linear and non-monotonic shapes and can capture complex relationships between the predictor variables and the response variable.

Regression Kriging (RK)

Regression Kriging (RK) is an interpolation technique that predicts the spatial distribution of soil properties by combining geostatistical methods and regression models. After fitting a regression model with environmental covariates and the target variable, RK uses the Kriging method to interpolate the residuals[71]. A hybrid approach of regression and kriging of residuals has been performed in some of the articles[67, 112, 142, 49, 15]. These studies demonstrate that RK has the potential to predict soil properties and map the spatial distribution of soil properties. Other studies have found that prediction models combined with residual Kriging yielded better prediction accuracy for soil properties.[159, 21, 107].

Regression kriging is combination of two methods as describe above and they are explain as,

1. Regression modelling First the regression model fit on the data as explained in MLR section above using the equation,

$$y = X\beta + \epsilon \quad (3.4)$$

where y = dependent variable predicted value, X = the predictor variables vector, β = vector of regression coefficients, ϵ = vector of residuals. The regression coefficients were estimated using linear regression and the random forest method.

2. Residual modeling

The residuals, epsilon, from the regression model are modeled using kriging to capture any spatial correlation in the data. The kriging estimator of the

residual at an unknown location s is given by:

$$\epsilon(s) = \sum (w_i * [y_i - X_i * \beta]) \quad (3.5)$$

where, w_i is the kriging weight assigned to the i^{th} observation, y_i is the observed response variable at location i , and X_i is the vector of predictor variables at location i .

The next step is to combine the regression and kriging model predictions to produce the dependent variables y at an unknown location s :

$$y(s) = x(s) * \beta + \epsilon(s) \quad (3.6)$$

where $X(s)$ is the vector of predictor variables at site s , β is the vector of regression coefficients derived from the training data, and $\epsilon(s)$ is the residual kriging estimate at location s .

Random Forest (RF)

A popular machine learning method for classification and regression is the Random Forest algorithm[128, 161]. It builds an ensemble of decision trees based on different subsets of the training data and averages their predictions to improve accuracy and control overfitting. The algorithm is based on the bagging concept, which entails creating bootstrap samples of data and fitting decision trees to each sample. Because of its high accuracy and ability to handle complex data, it is a valuable tool for soil scientists and environmental practitioners and is used in many predictive studies[72, 88, 48, 47, 149].

A random forest's decision trees are constructed from the top down, with the algorithm recursively splitting the data into smaller subsets based on the features that best separate the classes or minimise the variance of the target variable. Each split is selected according to a criterion such as information gain or mean squared error. The trees are grown until they reach a stopping criterion, such as a maximum depth, a minimum number of observations per leaf, or a criterion improvement of a certain amount. After all of the trees have been built, they are used to

predict the class or value of a new instance. Based on the features and path it takes through the tree, each tree predicts a class or value. The predictions from all of the trees are then combined to yield the final prediction. The predicted class for classification tasks is the mode or most frequent class among the tree predictions. The mean or median of the data is used for regression tasks[97, 23].

Artificial Neural Network(NN)

An artificial neural network is a machine learning model that is based on the structure and operation of the human brain. It is made up of interconnected neurons that are organized into layers, with each neuron performing a simple calculation on its input and passing the result to the next layer until the final output is produced. ANN is also capable of establishing non-linear relationships among co-variates and handling complex datasets[120, 165, 81].

The neural network concept is founded on three major steps, which are as follows:

1. Multiply input x_i to weight w_i for each neuron in a layer. Each neuron in a layer receives inputs from the previous layer or input data, which are multiplied by their weights. After multiplying the inputs by the weights, the products are added together for each neuron in the layer, which is same as dot product of the row vector of input ($X = [x_1, x_2, \dots, x_n]$) with a row vector of weight ($w = [w_1, w_2, \dots, w_n]$). This yields a weighted sum of inputs, which serves as the activation function's input.

$$\sum = X \cdot w = (x_1 * w_1) + (x_2 * w_2) + \dots + (x_n * w_n) \quad (3.7)$$

2. Add offset/bias b to the dot product of row vectors of input and weights.

$$Y = X \cdot w + b \quad (3.8)$$

3. The weighted sum output is then passed through an activation function to introduce nonlinearity into the model and generate the neuron output. The

bias term is added to the weighted sum to shift the activation function's output. The activation function's output is used as the input to the next or final layer. The activation function's output is used as the input to the next layer or as the model's final output. Here we used ReLU activation function,

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases} \quad (3.9)$$

The neural network model proposed here has three hidden layers. The input feature dimension is reduced by selecting the most important features for learning from among the available features using RFE. The dimension of the input is 37. The hidden layers are made up of 29, 11, and 4 neurons, respectively. In the output layer, which consists of one neuron that indicates the predicted soil property value, the identity function is used. The ReLU activation function is used by all of the other layers to introduce non-linearity and improve predictive performance.

Dropout is also included in hidden layers to prevent overfitting. They each have a dropout probability of 0.3 or 0.4. To achieve the highest precision in predicting the values, various optimizers are tested.

The mean square error (MSE) is used to evaluate the final predictions. In addition, batch normalization was used in the model. During training, an adaptive learning rate was used to improve the model. To properly set the hyperparameters, different values of dropout probability and percent reduction in learning rate after certain epochs were tested. Overfitting occurs when the dropout value is too low, and underfitting occurs when the value is too high.

The model ensures that the network learns independently of the other layers. Backpropagation is one of the neural network's advantages over other methods. Backpropagation is a learning algorithm that adjusts the weights of the network's neurons based on the difference between the predicted and actual output. The network can learn to make better predictions by propagating the error back through the network and adjusting the weights.

3.5 Result & Discussion

It is critical to use appropriate metrics that accurately reflect the performance of the model when comparing various regression models. Three widely used metrics are used in this thesis: MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and R-squared. MAE, RMSE, and R-squared as predictor model evaluators metrics provide a comprehensive evaluation of the model's performance. MAE and RMSE provide information about the magnitude of the errors in the predictions, while R-squared (R^2) provides information about how well the model fits the data and the strength of the relationship between the variables.

The three tables presented the results of predicting soil properties using different models, with the assessment based on three evaluation metrics - Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-Squared (R^2) along with descriptive statistics, including minimum (Min), maximum (Max), mean, standard deviation (SD), median, kurtosis, and skewness, for the predicted values. These statistics provide an overall view of the distribution of predicted values and can help to identify any potential outliers or biases in the models. Soil pH was the best-predicted soil property. The MLR model had the highest RMSE value of 0.27, indicating a larger deviation from the actual values. The ANN model had the lowest MAE and RMSE values of 0.04 and 0.06 respectively, indicating the best performance in predicting soil pH. The R2 values for all models ranged from 0.42 to 0.58, with ANN having the highest value of 0.58, indicating a good fit between the predicted and actual values (Table. 3.4).

Table 3.4: The performance of different Soil pH prediction models using three indicators of model evaluation with descriptive statistics

| Models | MAE | RMSE | R-squared | Min | Max | Mean | StDev. | Median | Kurtosis | Skew |
|--------|------|------|-----------|------|------|------|--------|--------|----------|------|
| MLR | 0.39 | 0.49 | 0.25 | 6.52 | 9.42 | 7.86 | 0.39 | 7.89 | -0.09 | 1.11 |
| RF | 0.33 | 0.40 | 0.41 | 6.71 | 8.99 | 7.85 | 0.34 | 7.88 | -0.21 | 1.20 |
| RKLR | 0.34 | 0.43 | 0.38 | 6.70 | 9.00 | 7.84 | 0.35 | 7.88 | -0.28 | 1.11 |
| RKRF | 0.39 | 0.49 | 0.37 | 6.13 | 9.53 | 7.87 | 0.57 | 7.90 | -0.10 | 0.28 |
| GAM | 0.37 | 0.44 | 0.38 | 6.35 | 9.32 | 7.88 | 0.49 | 7.85 | 0.20 | 0.25 |
| ANN | 0.23 | 0.27 | 0.67 | 6.58 | 9.00 | 7.89 | 0.40 | 7.88 | -0.28 | 0.67 |

As seen in Table. 3.5, the ANN model had the lowest MAE and RMSE values

of 0.09 and 0.14 respectively, indicating the best performance in predicting soil EC. The R² values for all models ranged from 0.24 to 0.62, with ANN having the highest value of 0.62, indicating a good fit between the predicted and actual values.

Table 3.5: The performance of different Soil EC prediction models using three indicators of model evaluation with descriptive statistics

| Models | MAE | RMSE | R ² | Min | Max | Mean | StDev. | Median | Kurtosis | Skew |
|--------|------|------|----------------|------|------|------|--------|--------|----------|------|
| MLR | 0.17 | 0.25 | 0.24 | 0.02 | 0.75 | 0.36 | 0.14 | 0.35 | -0.16 | 0.24 |
| RF | 0.13 | 0.21 | 0.42 | 0.13 | 0.80 | 0.37 | 0.12 | 0.36 | 0.39 | 0.57 |
| RKLR | 0.17 | 0.26 | 0.37 | 0.02 | 1.04 | 0.35 | 0.20 | 0.30 | 1.30 | 1.11 |
| RKRF | 0.14 | 0.22 | 0.39 | 0.13 | 0.81 | 0.37 | 0.12 | 0.36 | 0.53 | 0.62 |
| GAM | 0.16 | 0.23 | 0.36 | 0.02 | 1.04 | 0.35 | 0.20 | 0.30 | 1.24 | 1.10 |
| ANN | 0.09 | 0.14 | 0.62 | 0.11 | 0.83 | 0.35 | 0.15 | 0.32 | 0.92 | 0.96 |

For Soil OC prediction, the ANN model had the lowest MAE and RMSE values of 0.07 and 0.10 respectively, indicating the best performance in predicting soil OC as shown in table. 3.6. The R² values for all models ranged from 0.23 to 0.65, with ANN having the highest value of 0.65, indicating a good fit between the predicted and actual values.

Table 3.6: The performance of different Soil OC prediction models using three indicators of model evaluation with descriptive statistics

| Models | MAE | RMSE | R-squared | Min | Max | Mean | StDev. | Median | Kurtosis | Skew |
|--------|------|------|-----------|------|------|------|--------|--------|----------|------|
| MLR | 0.13 | 0.17 | 0.23 | 0.00 | 0.87 | 0.39 | 0.11 | 0.39 | 2.26 | 0.43 |
| RF | 0.11 | 0.14 | 0.39 | 0.20 | 0.79 | 0.39 | 0.08 | 0.40 | 3.34 | 1.00 |
| RKLR | 0.11 | 0.16 | 0.36 | 0.21 | 0.79 | 0.39 | 0.08 | 0.40 | 3.27 | 0.99 |
| RKRF | 0.12 | 0.16 | 0.36 | 0.04 | 1.02 | 0.40 | 0.16 | 0.39 | 1.89 | 0.94 |
| GAM | 0.12 | 0.17 | 0.35 | 0.10 | 0.98 | 0.38 | 0.14 | 0.37 | 1.75 | 0.79 |
| ANN | 0.07 | 0.10 | 0.65 | 0.12 | 0.84 | 0.40 | 0.11 | 0.39 | 2.56 | 0.85 |

By looking at the performance evaluation indicators, we can see that for all three soil properties (Soil pH, Soil EC, and Soil OC), the Artificial Neural Network (ANN) model performs the best overall, with the lowest MAE and RMSE values, as well as the highest R-squared value. This indicates that the ANN model is able to accurately predict all three soil properties with the least amount of error and the highest amount of explained variation. Secondly, comparing the descriptive statistics for the predicted values, we can see that for all three soil properties, the ANN model generally has the smallest range of predicted values (from the minimum to the maximum), the smallest standard deviation, and the smallest

skewness and kurtosis values. This indicates that the ANN model is able to produce more consistent and less extreme predicted values compared to the other models.

This indicates that the ANN model is better at predicting soil properties compared to the other models tested. One of the reasons for the superior performance of the ANN model could be its ability to handle complex and nonlinear relationships between the soil properties and environmental variables. ANN models can capture intricate patterns and relationships between variables that may be missed by other models, making them highly flexible and effective in predicting soil properties.

Figure 3.4, 3.6 and 3.5 shows the scatterplot of predicted versus actual soil parameter values with a regression line for different algorithms.

The bivariate density curve in green color contours is a smooth curve that represents the joint distribution of the predicted and actual values. It can be useful in identifying patterns and trends in the data and can give insights into the relationship between the two variables. The univariate density curves with histograms (blue color) on the margins represent the marginal distributions of the predicted and actual values. The histograms show the frequency of values in each bin, while the density curve shows the probability density function of the values. These curves can provide additional insights into the distribution of the data and can be useful in identifying outliers or skewness in the data. The regression line is drawn through the data points so that it lies as close as possible to all the points, representing the best-fit line for the model. The trend indicates that the predictor environmental covariates still provide information about the response. R-square, even when small, can be significantly different from 0, indicating that the regression model has statistically significant explanatory power.

It was observed that the models are able to capture the underlying trend in the data, as evidenced by their relatively low MAE and RMSE values. This suggests that the models are useful for predicting soil properties based on the independent variables. Furthermore, the ANN model, which had the highest R2 value, also had the lowest MAE and RMSE values, indicating that it is the most accurate

model for predicting soil properties based on the independent variables.

In summary, predicting soil properties using environmental covariates is an important and challenging task in soil science. While there are many challenges associated with this process, it has the potential to provide valuable insights into soil-landscape relationships and support informed land management decisions.

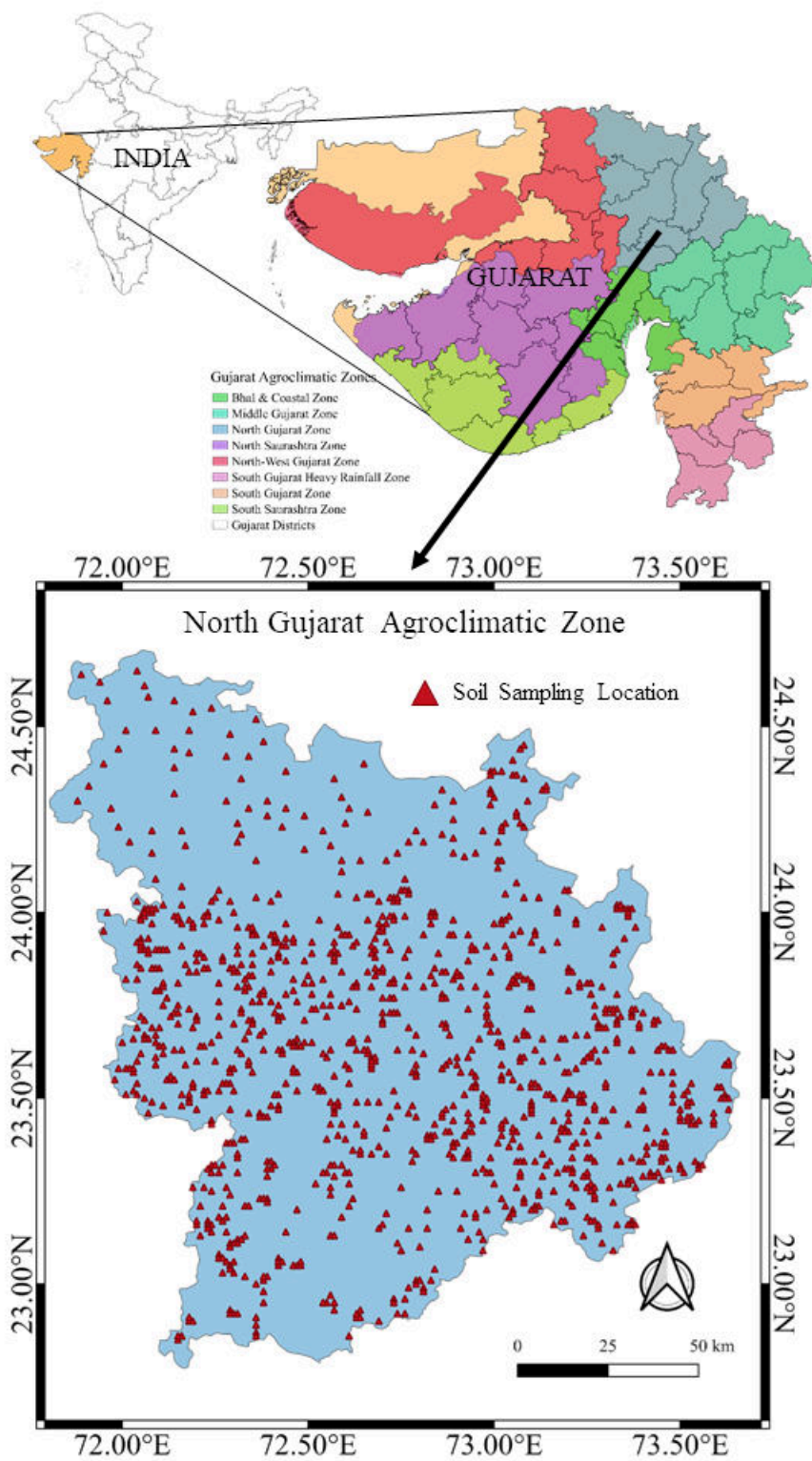


Figure 3.1: Study Area showing soil sample points in North Gujarat Agro-climatic area

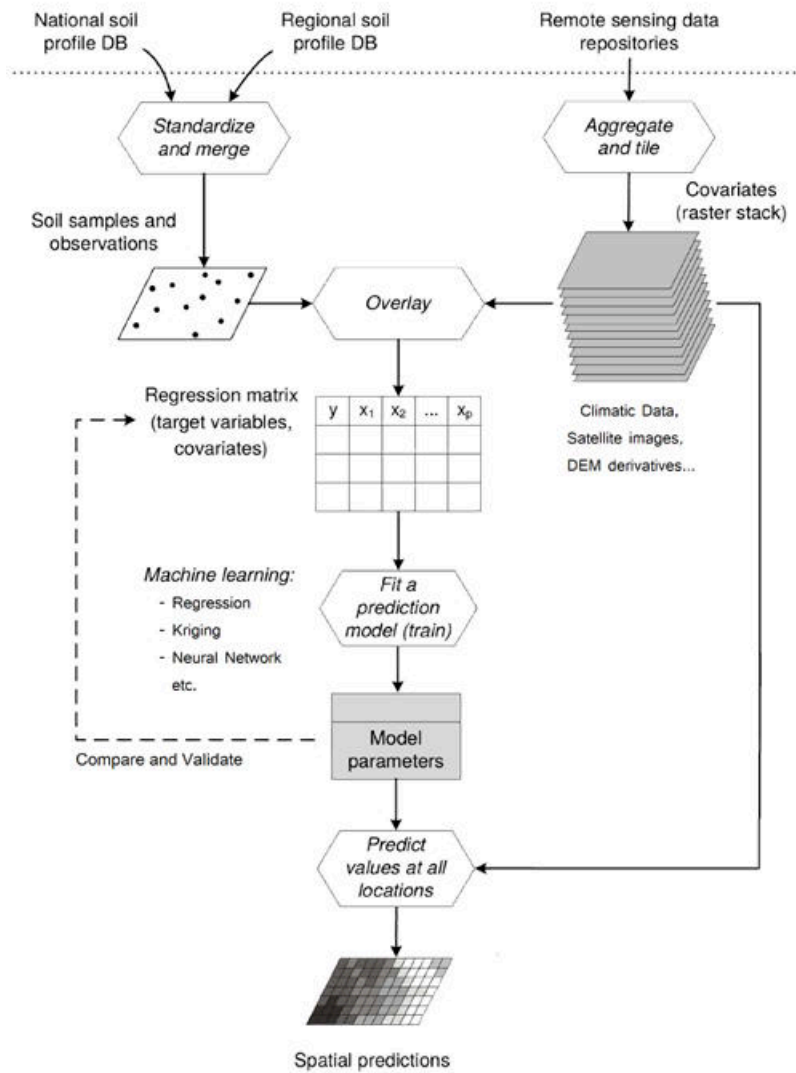


Figure 3.2: Predictive Soil modeling methodology flowchart

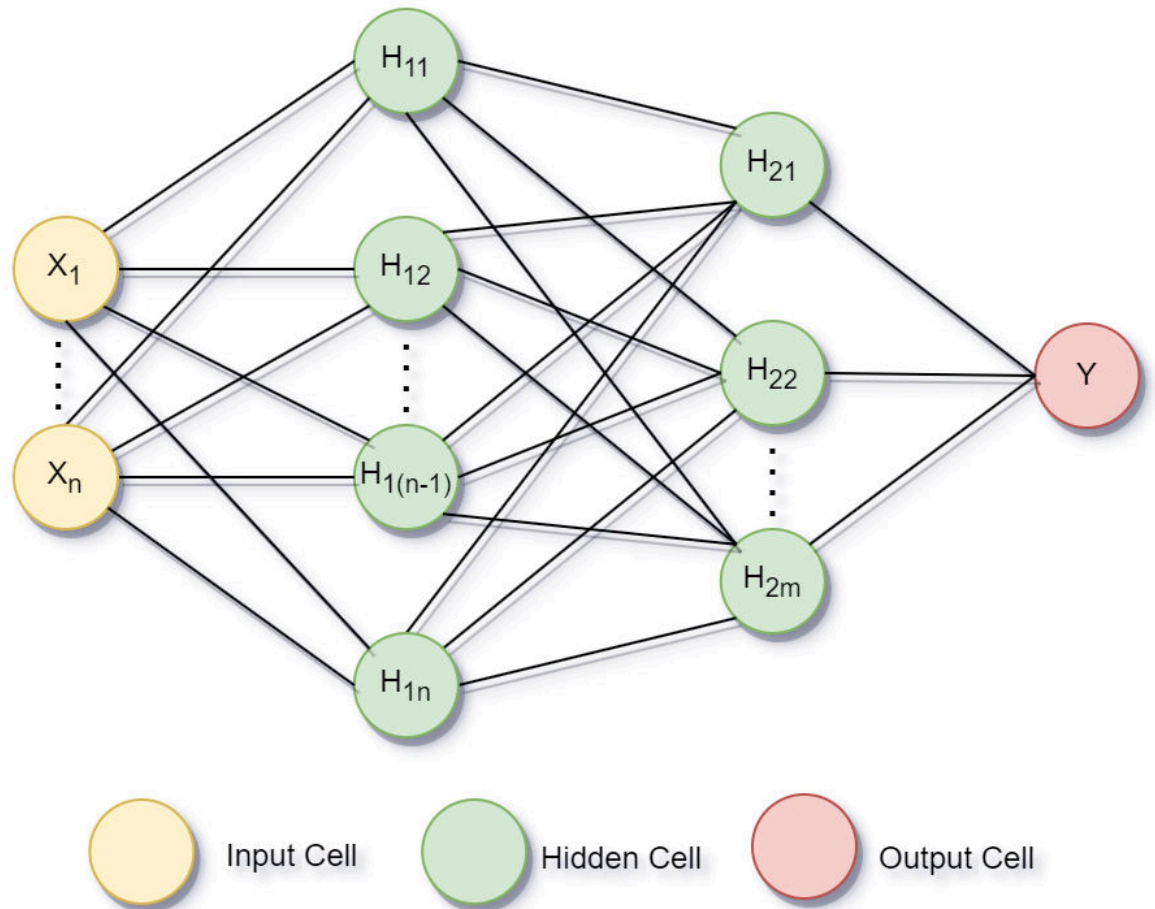
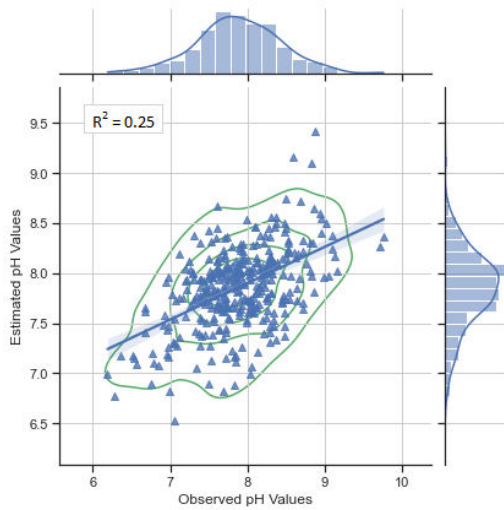
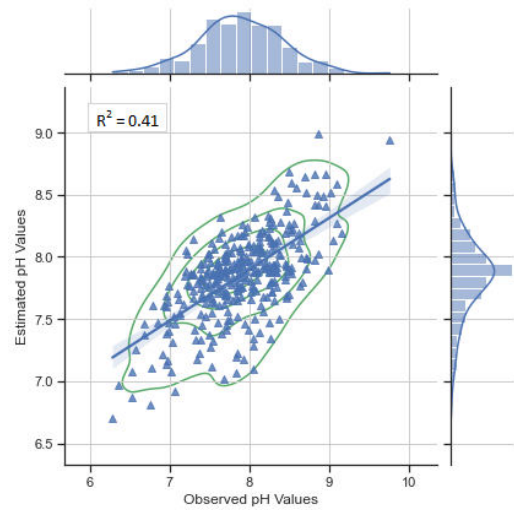


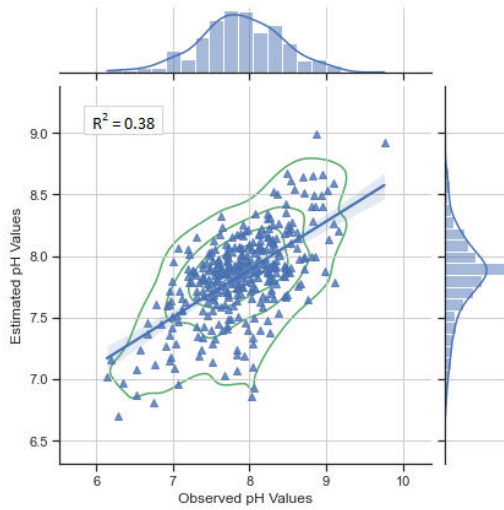
Figure 3.3: The above image shows a 3-layer artificial neural network: One input layer, two hidden layers, and one output layer.



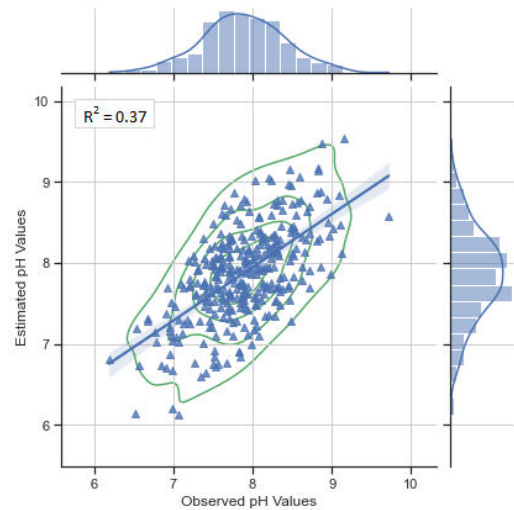
(a)MLR



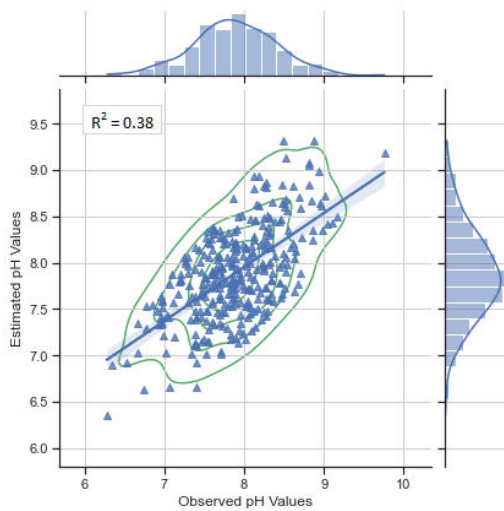
(b)RF



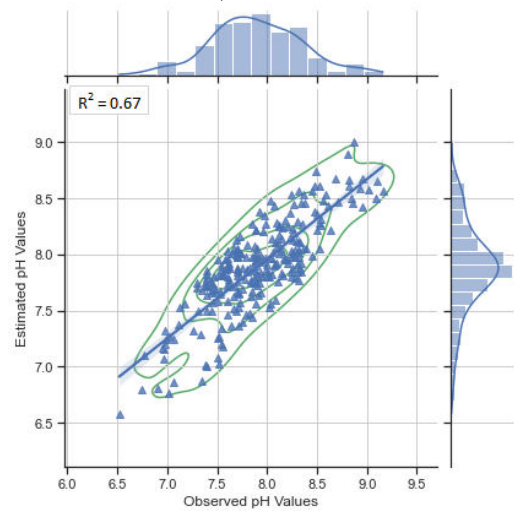
(c)RK - Linear regression



(d)RK - RF

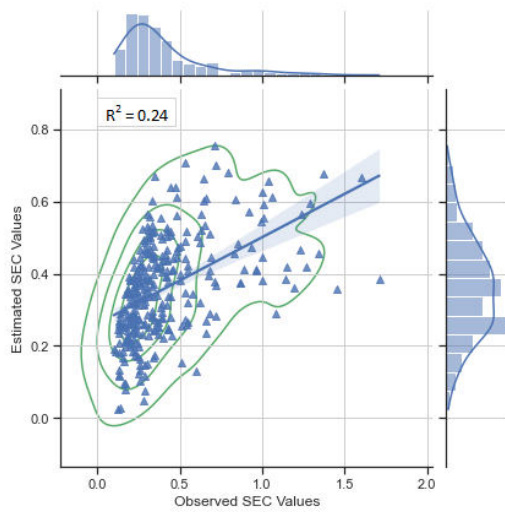


(e) GAN

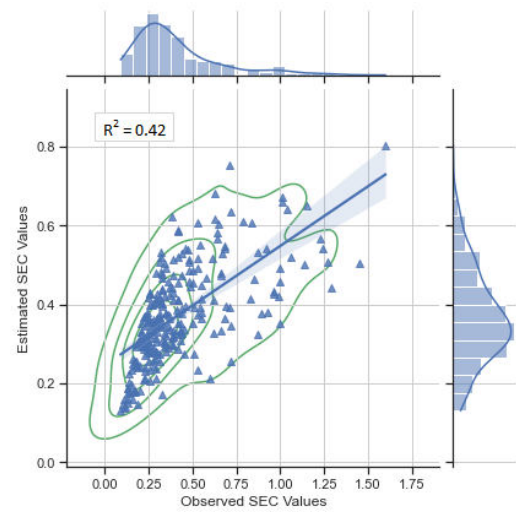


(f)ANN

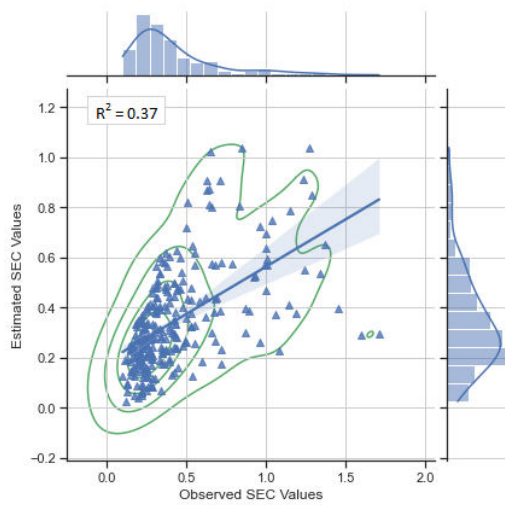
Figure 3.4: Performance of Predicted Soil pH property for six different prediction models



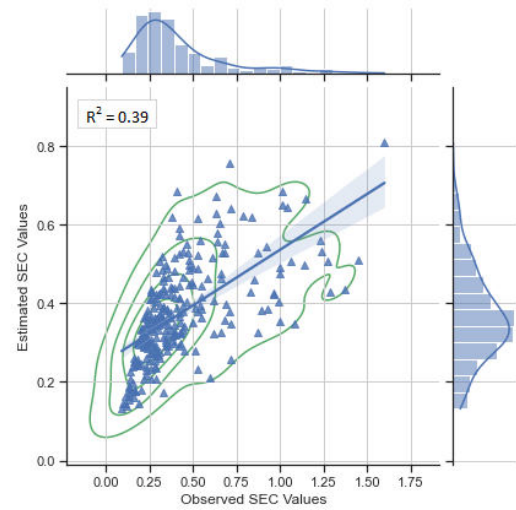
(a)MLR



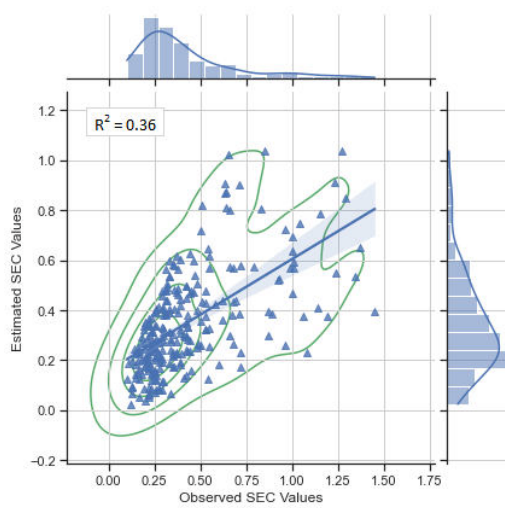
(b)RF



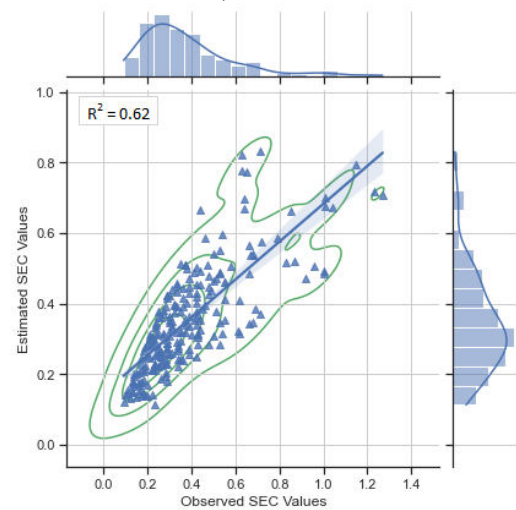
(c)RK - Linear regression



(d)RK - RF



(e) GAN



(f)ANN

Figure 3.5: Performance of Predicted Soil EC property for six different prediction models

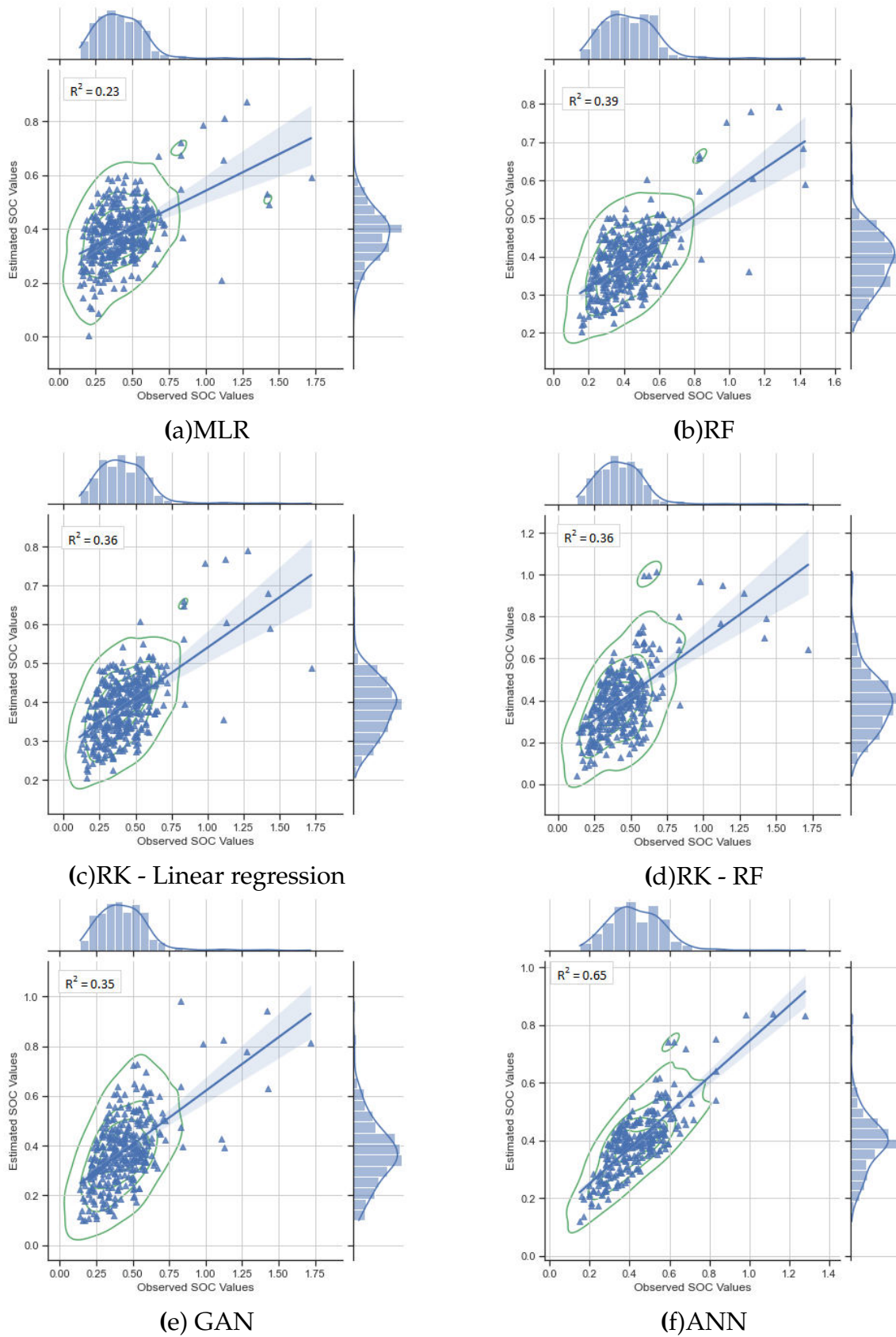


Figure 3.6: Performance of Predicted Soil OC property for six different prediction models

CHAPTER 4

Desertification Vulnerability Assessment

The objective of this study was to characterize an area vulnerable to desertification and its severity level. In order to slow down or stop the progress of desertification. The flaw of any system can be assessed by considering the basic concept of vulnerability. In the context of the present study, the vulnerability assessment of desertification was conducted using remote sensing and geographic information system (GIS) techniques to map out sensitive areas. Areas, where interventions are required, will also be identified using a vulnerability assessment. This might involve initiatives like water harvesting, sustainable land management techniques, or reforestation.

Multidisciplinary approaches have also been made in an effort to integrate various data sources and methodologies in order to study the implications of land degradation in semi-arid India due to the complex and multilayered relationship between land degradation and socioeconomic development [125]. An integrated approach can be used to portray the progress of desertification in a thorough and simple manner by taking into account quantitative and qualitative methods and by employing certain indicators[153]. Rashid et al. [130] applied an indicator technique to comprehend regional land degradation. Slope, land use, and vegetation cover were employed as indicators in this study. However, it was noted that at time periods longer than a year, signs of land degradation are dynamic. Longer-term records are required to be investigated in order to provide a better assessment. In order to assess an area vulnerable to desertification, a number of elements, including climate, vegetation, soil, and land use, must be examined. Natural and anthropogenic factors accelerate the Land degradation process, which

can be assessed using a multidisciplinary approach based on environmental and human indices.

Many different mathematical models have been developed and used in different areas to identify areas vulnerable to desertification. [16, 139, 41, 6, 27] The conventional methods to identify vulnerability were to classify basic parameters for land degradation and then define classes with a certain range of values and assign particular classes. The Mediterranean Desertification and Land Use (MEDALUS) approach is a methodology used to assess the risk of land degradation in Mediterranean regions. It involves the analysis of various environmental and socio-economic factors to develop an environmental sensitive index along with the social and economic index of settlements. All the factors were indexed statistically and using expert knowledge in different severity classes. The MEDALUS approach has been used in various regions, including the Mediterranean, to assess the risk of land degradation and desertification. It is an effective methodology for identifying the areas that are most vulnerable to land degradation and can be used to develop targeted interventions to mitigate the risk of degradation. [85, 6, 22, 91, 126, 160]

Another risk assessment model globally used is the fuzzy logic model (FL) integrated within geospatial environment [143, 41, 13]. In FL model, fuzzy membership function with statistical mean and standard deviation are used to identify risk areas. Environmental and socio-economic data values were found, which represent not a particular class of value but a transitional zone. These zones are most important in terms of degradation. They have the lowest probability of being in a certain consistent class and the highest probability of being in the most vulnerable classes [198, 38]. These transitional zones are critical in assessing the risk of land degradation because they are the areas where preventative measures can be most effective in mitigating the risk of degradation. By identifying the transitional zones, FL models can provide a more accurate assessment of the risk of land degradation than traditional classification methods. By focusing on the transitional zones, the models can identify areas that may not be at immediate risk of degradation but are susceptible to degradation in the future. This allows

policymakers and land managers to implement targeted interventions to prevent future degradation and preserve the health of the ecosystem.

The study's outcome would help identify the vulnerable areas to desertification/land degradation and the severity of degradation in natural and socio-economic parameters. Additionally, remote sensing techniques and GIS to evaluate the degree of risk would help the expert in very efficient planning of resource allocation and decision-making to mitigate desertification and layout proper management plans. It is possible to implement conservation and management plans to lower this risk and save the land and the people who depend on it, once the regions at risk of desertification have been identified. The approaches could consist of:

- The preservation and restoration of vital natural resources like vegetation, soil, and water;
- The adoption of sustainable land use and land cover change management practices;
- The promotion of alternate livelihood options and income-generating ventures that ease the strain on natural resources;
- The promotion of education and training on desertification and effective conservation practices.

4.1 Study Area

Panchmahal is located in the Eastern Part of Gujarat. It lies between the parallels of latitude 22.26° and 23.11° and the meridians of longitude 73.34° and 73.95° . Godhra is the district headquarter. The district is located on eastern end of the state. The length from north to south of the district is about 95 km and from east to west about 67 km. It is bounded on the north by the Mahisagar district, on the to the north-east and east by Dahod district. Chhota Udaipur district to the southeast and on the west by Kheda district and Vadodara district to the southwest. Agriculture is main source of income of this district. Panchmahal is rich

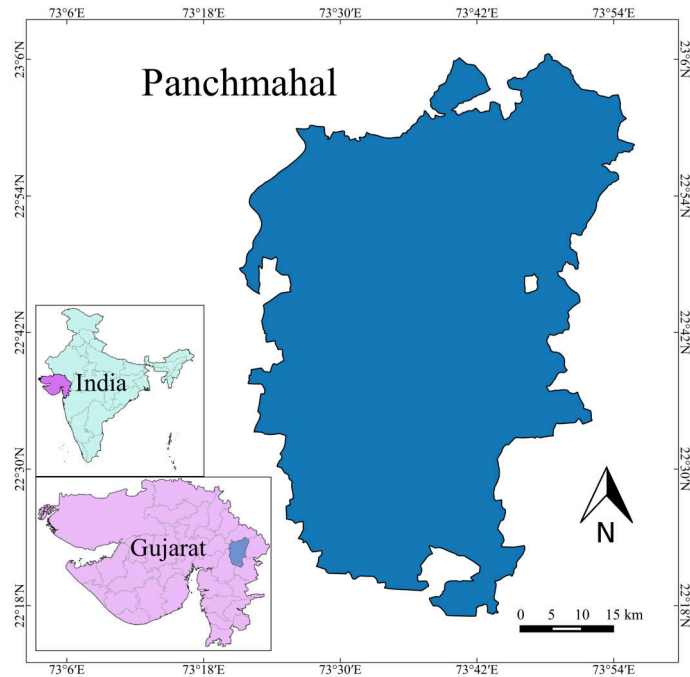


Figure 4.1: Panchmahal District of Gujarat

in forestry resources and quartz and manganese are the two largest mineral resources in the district.[31]

4.2 Datasets

Natural parameters like climate, soil, vegetation, terrain, land cover, etc. and anthropogenic parameters like population, land use, economy, amenities, etc. directly or indirectly affect land degradation. These parameters are considered as inputs for vulnerability assessment. NDVI is used for Vegetation Index (VI). Satellite images of three-season (*Kharif, Rabi and Zaid*) were used to calculate NDVI. Soil texture, soil erosion, soil pH, soil depth and drainage maps were used as soil input data extracted from soil map prepared by National Bureau of Soil Survey and Land Use Planning (NBSS-LUP)[116]. Consultative Group on International Agricultural Research - Consortium for Spatial Information (CGIAR-CSI) global aridity index database of aridity index was used for climate index[167]. This aridity index modeled using the data available from the WorldClim Global Climate Data from 1950-2000 as input parameters[73]. Land use land cover (LULC) maps used

Table 4.1: Dataset Specifications

| Index | Resolution | Layer format | Source |
|------------------|---------------|--------------|---------------------------|
| Climate | 865 mtr | Raster | Global Aridity Index[167] |
| Soil | 1:250000 | Vector | NBSS-LUP[116] |
| Land Utilisation | 56 mtr | Raster | Bhuvan-ISRO |
| Elevation&Slope | 30 mtr | Raster | Bhuvan-ISRO |
| Vegetation | 30 mtr | Raster | Landsat Data |
| Socio-Economic | Village level | Vector | 2011 Census of India[31] |

for land utilization index, and used a digital elevation model for elevation and slope index. India Census 2011 [31] data were used to generate a socio-economic index for anthropogenic input for desertification risk analysis. Detail dataset specification, including its data type, source, and spatial resolution given in table 4.1.

4.3 Methodology

Land degradation refers to the gradual deterioration of the quality of the land and its ability to support various ecological and economic functions. It is a complex phenomenon that involves a range of interconnected processes that can lead to a decline in land productivity and biodiversity. It is then essential to generate different indices, which can be utilized to recognize the weak areas prone to desertification. Climate, vegetation, land use and soil play a huge part in the desertification of any area. Subsequently, these indices are used, Climate index (CI), Soil index (SI), Elevation index (EI), Slope index (SlopI), Vegetation index (VI), and land utilization index (LUI). These information sources have been taken as spatial layers and have coordinated in a GIS environment to give naturally weak areas to desertification. Human intervention prompts critical changes in the environment, making socioeconomic factors another significant contribution to assessing vulnerability to desertification. The incorporation of the Cumulative Amenities index (CAI) and Economic Development Index (EDI) into a Socio-economic index (SEI) is a useful approach to analyzing the social and economic factors that contribute to land degradation. In this thesis, two unique methodologies were

Table 4.2: Vulnerability index definition in MEDALUS model

| Category | Severity Level |
|----------|----------------|
| 1 | Very Low |
| 2 | Low |
| 3 | Moderate |
| 4 | High |

taken for calculating vulnerability. MEDALUS model is applied as one technique, and the second methodology is fuzzy logic for quantifying the information and finding the risk areas for desertification.

4.3.1 Medalus Model

The procedure used depends on the basic model of MEDALUS, created in a large project set up by the European Commission.[85] The MEDALUS is one of the most generally utilized models in observing desertification affectability. The MEDALUS model's ability to adapt the amount of variables used to determine the vulnerability to desertification is one of its primary advantages. For instance, in a study in Serbia, three criteria were employed for the soil quality index, three for the climatic quality index, two for the management quality index, and three for the vegetation quality index[80]. While in another study, three criteria were utilized for the climatic quality index, two for the management quality index, and six were used for the soil quality index[43]. Because users can add, remove, and modify the sub-indicators as needed, the methodology is also flexible.

All the variables as described in the previous section were indexed into four categories from 1 to 4, where 1 represents the low vulnerability and 4 represents the high vulnerability as shown in table 4.2 The geometric mean of all the indices was calculated to find out the final desertification vulnerability index. Eq. 4.1 is representing the desertification vulnerability index(DVI).

$$DVI = (CI * SI * EI * SIpI * VI * LUI * CAI * EDI)^{\frac{1}{8}} \quad (4.1)$$

where, CI = Climate Index, SI = Soil Index, EI = Elevation Index, SIpI = Slope

Index, VI = Vegetation Index, LUI = Land Utilization Index, CAI = Cumulative Amenities Index, EDI = Economic Development Index. All the indices used to calculate the final DVI are discussed in detail in the following sections.

Climate Index (CI)

The values of aridity parameters are used to calculate the climate Index. Climate classes based on aridity index values quantify in four index values, as given in table 4.3.

Table 4.3: Quantization of Climate Index

| Climate classification | Aridity Value | Index Value |
|------------------------|---------------|-------------|
| Arid | < 0.2 | 4 |
| Semi-arid | 0.2-0.5 | 3 |
| Dry sub-humid | 0.5-0.65 | 2 |
| Humid | > 0.65 | 1 |

Soil Index (SI)

Soil is the main factor for land degradation study. Soil index was coordinated from a few soil parameters that consolidate soil erosion, soil drainage, soil depth, soil pH, and soil texture. All parameters were categorized as shown in table 4.4 and the geometric mean was determined for soil index. This soil index was then quantized in four severity classes using statistical mean and standard deviation, as shown in table 4.5.

$$SoilIndex = (Texture * Depth * pH * Drainage * Erosion)^{\frac{1}{5}} \quad (4.2)$$

Elevation Index (EI) & Slope Index (SlopI)

Elevation and slope are generally identified with land degradation other than in any remarkable case. Elevation values are not normally distributed. Instead of using mean and standard deviation for categories, median and quantile values

Table 4.4: Quantization of Soil parameter

| Soil Properties | Class | Categories |
|-----------------|---------------------------------|------------|
| Soil Texture | Clayey | 4 |
| | Very Fine, Fine | 3 |
| | Loamy, Fine Loamy, Coarse Loamy | 2 |
| | Loamy Skeletal, Clayey Skeletal | 1 |
| Soil Depth | Ext. Shallow (<10 cm) | 4 |
| | Shallow (10-75 cm) | 3 |
| | Mod. Deep (75-100 cm) | 2 |
| | Deep, Very Deep (>100 cm) | 1 |
| Soil pH | Strongly alkaline (>9.5) | 4 |
| | Mod. alkaline (8.5-9.5) | 3 |
| | Slightly alkaline (7.5-8.5) | 2 |
| | Neutral (6.5-7.5) | 1 |
| Soil Drainage | Very poor, Imperfect | 4 |
| | Somewhat excessive, Excessive | 3 |
| | Moderately well | 2 |
| | Well | 1 |
| Soil Erosion | Very severe | 4 |
| | Severe | 3 |
| | Slight Severe | 2 |
| | Moderate | 1 |

Table 4.5: Quantization of Soil Index

| Rule | SI Value | Index Value |
|------------------------|-------------|-------------|
| $< (\mu - \sigma)$ | < 1.48 | 1 |
| $(\mu - \sigma) - \mu$ | 1.48 - 1.79 | 2 |
| $\mu - (\mu + \sigma)$ | 1.79 - 2.10 | 3 |
| $> (\mu + \sigma)$ | > 2.10 | 4 |

were utilized for Indexing. Accordingly, the slope angle is significant for examination. The more extreme the inclination angle, more prominent the erosion, thus speeding water-stream and wind streams in a descending way increases with the slope. The slope index in degrees was listed in four classes, as appeared in table 4.6.

Table 4.6: Quantization of Slope Index

| Slope (in degrees) | Index Value |
|-----------------------------------|-------------|
| Mod. Steep sloping (15-30) | 4 |
| Moderately sloping (8-15) | 3 |
| Gently Sloping (3-8) | 2 |
| Very gently or nearby level (0-3) | 1 |

Vegetation Index (VI)

The fundamental factor concerning the security of soil richness and its profitability is vegetation. Destruction of vegetation, oftentimes, by human activities stimulates soil degradation to a tremendous degree, accordingly provoking desertification. Higher photosynthetic action will bring about lower reflectance in the red frequency channel and higher reflectance in the close-to-infrared frequency channel in satellite remote sensing data. Green plants absorb more red light and reflect more NIR light, while unhealthy or water-stressed vegetation reflects more red light and absorbs more NIR light. Therefore, NDVI is used as an indicator of vegetation health and can be used to identify areas of healthy or stressed vegetation. A high NDVI value indicates healthy vegetation, while a low NDVI value indicates stressed vegetation. In other words, the NDVI signature is significantly identified concerning the green plants, and it gets inverse if there is an event of unhealthy or water-stressed vegetation. To get a normal vegetation life, three years mean NDVI was thought about. Using this primary mechanism probability distribution of NDVI values vulnerable to desertification was indexed as table 4.7.

Socio-Economic Index(SEI)

Cumulative Amenities Index(CAI)

Table 4.7: Quantization of Vegetation Index

| NDVI Value | Index Value |
|-------------|-------------|
| < 0 | 4 |
| 0 - 0.17 | 3 |
| 0.17 - 0.38 | 2 |
| > 0.38 | 1 |

The other important factor for assessing any degradation in nature is the progress of society. As society grows, there may be increased awareness and concern about environmental issues, but it is not always the case that this leads to significant investigations and a reduction in harm to the environment. Amenities available at the village level, such as medical, education, communication, etc. used to calculate the village-level amenities index based on the cumulative weighted index model (eq.4.3). The villages have more amenities; the harm will be less. Equation 5 shows the cumulative amenities index for a particular village.

$$I_a = \frac{(\sum A_i W_i)}{\sum W_i} \quad (4.3)$$

where,

$i = 1$ to n

I_a = index for a particular settlement vis-à-vis class of amenity

n = Number of amenities in a category (e.g. 2 or 4 nos.)

$A_i = 0$ or 1 (Not available, Available)

W_i = Weight of the amenity within category/class facility and it is defined as,

$$W_i = \frac{N - K_i}{K_i} * 100 \quad (4.4)$$

N = Total no. of Settlements

K_i = No. of Settlements having amenity i Amenities index

Cumulative amenities index for a particular village was calculated using

$$CAI = \sum I_a \quad (4.5)$$

where

$a = 1$ to m ; m is the number of amenity categories

Economic Development Index(EDI)

The economic development of a village is derived based on its working and non-working population as given in Eq.4.6. The cumulative amenities index and economic development index was assembled into three classes subject to its mean and standard deviation.

$$EDI = \sqrt{D * W(1 - A)} \quad (4.6)$$

where,

D = Population density of the village (Total population/village area)

W = Employed population / Total population

A = Unskilled workers proportion (i.e. (agricultural labors + marginal workers) / Total population)

Land Utilization Index(LUI)

Land use/land cover data were procured from ISRO [118] which was prepared using knowledge-based supervised classification and probability classifier calculation. The information demonstrated that there were 16 diverse land use and land cover in Gujarat state. All the classes have been allocated a numerical value from 1 to 4 according to their weakness towards land degradation to the expert's decision. Table 4.8 displays the effect of vulnerability for different land utilization classes. Same LUI is used in both models.

4.3.2 Fuzzy logic approach

Data and decisions are firmly connected, the raw data was associated by fuzzy logic to the quantitative decision rules. Quantitative evaluation using fuzzy logic theory was investigated in order to consolidate the decision rules with both quantitative and subjective manners for vulnerability assessment. In the fuzzy logic

Table 4.8: Land Utilization Index

| LULC Class* | Index Value |
|-----------------------|-------------|
| Scrub/Degraded Forest | |
| Littoral swamp | |
| Other wasteland | 4 |
| Gullied | |
| Scrubland | |
| Current fallow | 3 |
| Kharif only | |
| Rabi only | |
| Zaid only | 2 |
| Double / triple | |
| Plantation/orchard | |
| Evergreen forest | |
| Deciduous forest | 1 |
| Water bodies | |
| Build up | NA |
| Rann | |

*Bhuvan-ISRO LULC

framework, The distribution of data of a single parameter was assumed to be a normal distribution with mean μ and standard deviation σ . Normal probability density function(pdf) was used to obtain the membership of individual variables to be in a particular class, following the statistical formula,

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.7)$$

where,

x = an individual from a class of picked boundaries,

μ = class mean

σ = the standard deviation of the class.

The data was partitioned in three classes as given in table 4.9. The fuzzy areas between these classes were identified as vulnerable areas and given the severity index as very low, low, moderate and high as indicated for input parameters. Besides, recognizable proof of weak regions is extremely straightforward from this model. The qualities lying in the middle of $(\mu + \sigma)$ of the past class and $(\mu -$

Table 4.9: Classes definition in Fuzzy logic

| | |
|---------|--|
| Class 1 | $X \leq \mu - \sigma$ |
| Class 2 | $(\mu - \sigma) < X \leq (\mu + \sigma)$ |
| Class 3 | $X > (\mu + \sigma)$ |

Table 4.10: Vulnerability index definition in Fuzzy logic

| Value | VI, EDI, CAI, | CI, SI, EI, SlopI |
|---|---------------|-------------------|
| $X \leq (\mu_1 - \sigma_1)$ | High | Very Low |
| $(\mu_1 + \sigma_1) \leq X \leq (\mu_2 - \sigma_2)$ | Moderate | Low |
| $(\mu_2 + \sigma_2) \leq X \leq (\mu_3 - \sigma_3)$ | Low | Moderate |
| $X \geq (\mu_3 - \sigma_3)$ | Very Low | High |
| Other | NA | NA |

σ) of the following class were considered vulnerable and recognized as the danger territory as depicted in the table. Areas apart from these were not considered as vulnerable and given NA as an index(table 4.10).

As an example, the vegetation index is taken in Fig. 4.2. Here three bell-shaped pdf plots represent three classes based on NDVI data mean and standard deviation and the fuzzy area, shown in red, yellow, green and blue color rectangle, between $\mu + \sigma$ of the previous class and $\mu - \sigma$ of the following class is called the transition zone and considered as a vulnerable class and given its severity index from high to very low.

$$[h]DVI = MAX(CI, SI, EI, SlopI, VI, LUI, CAI, EDI) \quad (4.8)$$

Composition of all indices was generated and the maximum value is taken for DVI for the fuzzy logic algorithm as shown in Eq. (4.8).

4.4 Results & Discussion

4.4.1 Climate Index (CI)

The climate of this area is portrayed by a blistering summer and dryness in the non-rainy seasons. Humidity of Panchmahal district is higher in July, August and

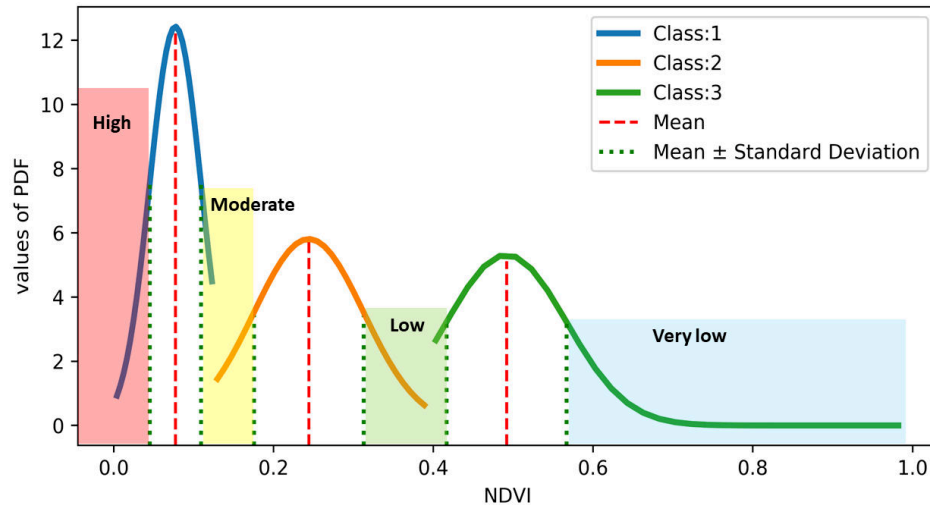


Figure 4.2: Severity index for vegetation using fuzzy logic

September months. The district has arid and semi-arid climate having aridity index from 0.2 to 0.6. Fig. 4.3(a) shows the climate index for Panchmahal district of Gujarat using Medalus model. Panchmahal have nearly two equal classes as moderate and low severe climate index. As Panchmahal is mostly covered with forest area and it has good amount of rainfall around the year gives low aridity index area more than 50% using fuzzy logic(Fig. 4.3(b)).

4.4.2 Soil Index (SI)

The study area has great assortments of soils. The soil varies in its fruitfulness here and there.

The western zones have richness in soils, while the eastern territories have shallow sandy soils. In the northwest, the soil is alluvial; south of this there is a belt of dull black soil. Toward the northeast, a rich medium-dark soil called Besar is helpful for wheat and gram. Southern has nearly all through goradu soil which is more fruitful. In the southern stretches of rich dark soil with all through goradu soil which is more fruitful in the centre region. 60% of the district area falls under a very low severity zone using the Medalus approach(Fig. 4.4(a)). While in fuzzy logic 15% area for both low and very low severity zone(Fig. 4.4(b)). As soil is very rich in Panchmahal district there is no area in high or moderate one

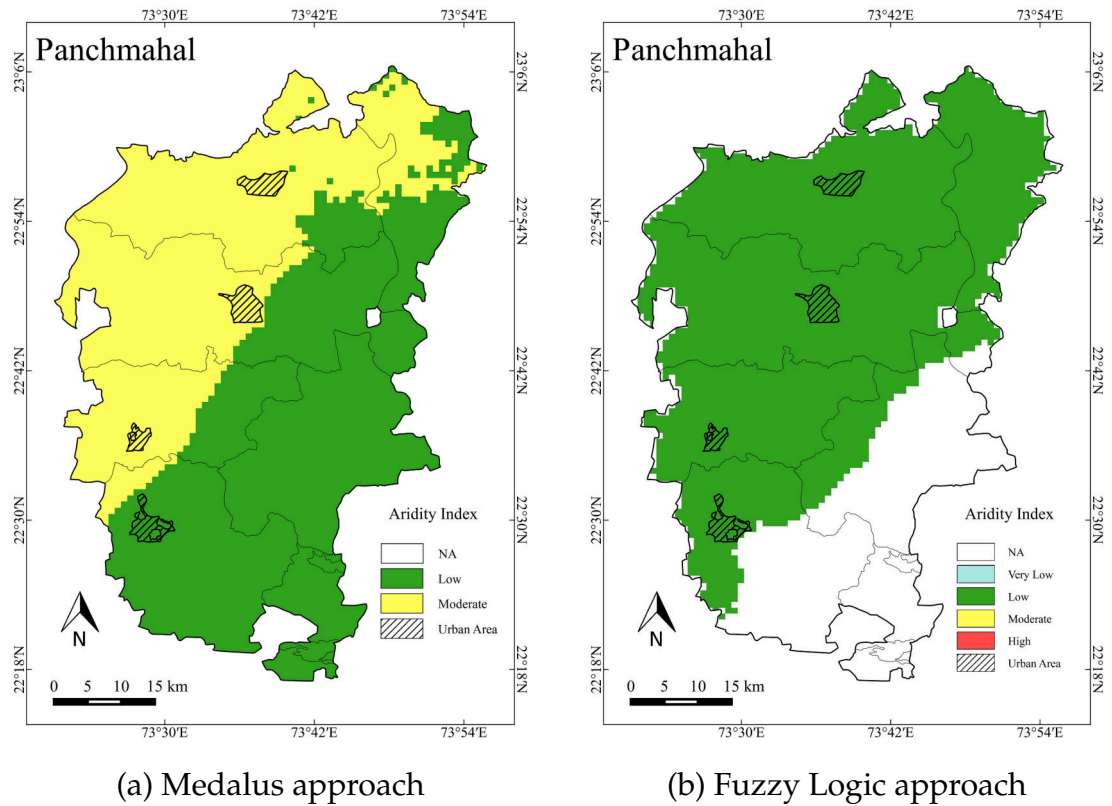
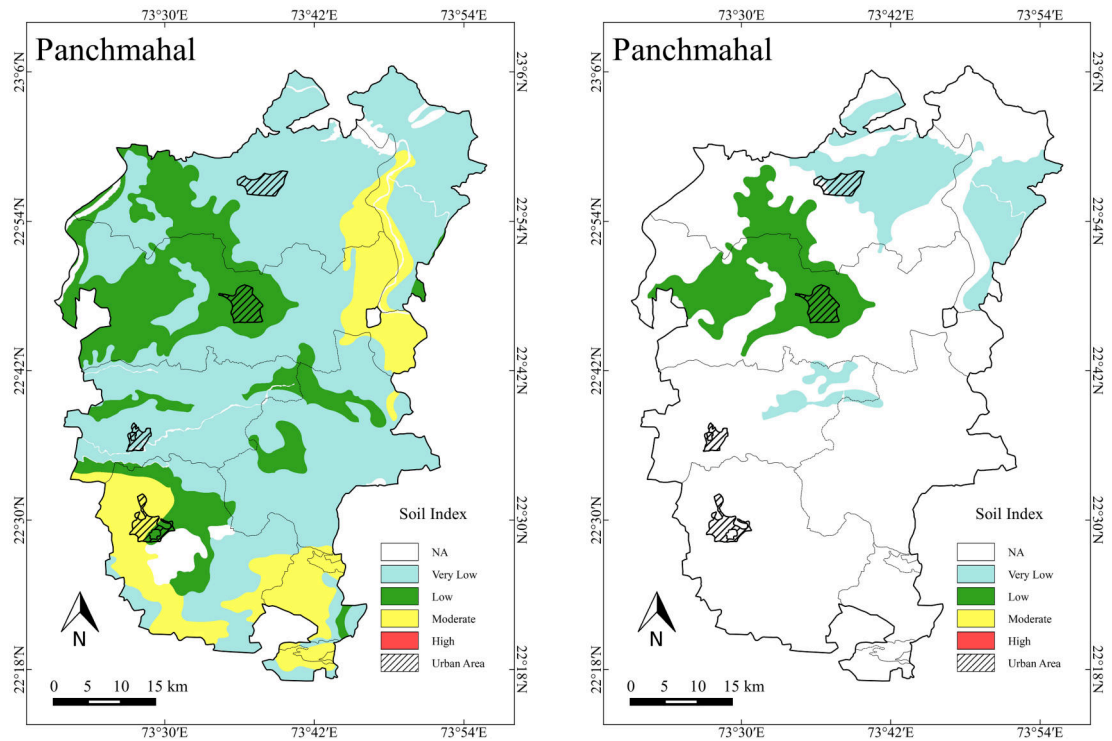


Figure 4.3: Aridity Index Map

for desertification vulnerability.

4.4.3 Elevation Index(EI) Slope Index

Panchmahal has little scattered sedimentary and volcanic dissected hills. The southern and eastern part is covered with hills. Elevation gradually decreases from east to west. Pavagadh hill rises suddenly to a stature of more than 800m and with high slope in the southern part of the region. The western part comprises a plain region. The central part of the district is covered by slopes with forest and plain cultivated lands in villages situated in the stream valley. The western piece of the locale is generally flat. Fig. 4.5 and Fig. 4.6 shows the elevation index and slope index map for the Panchmahal region.



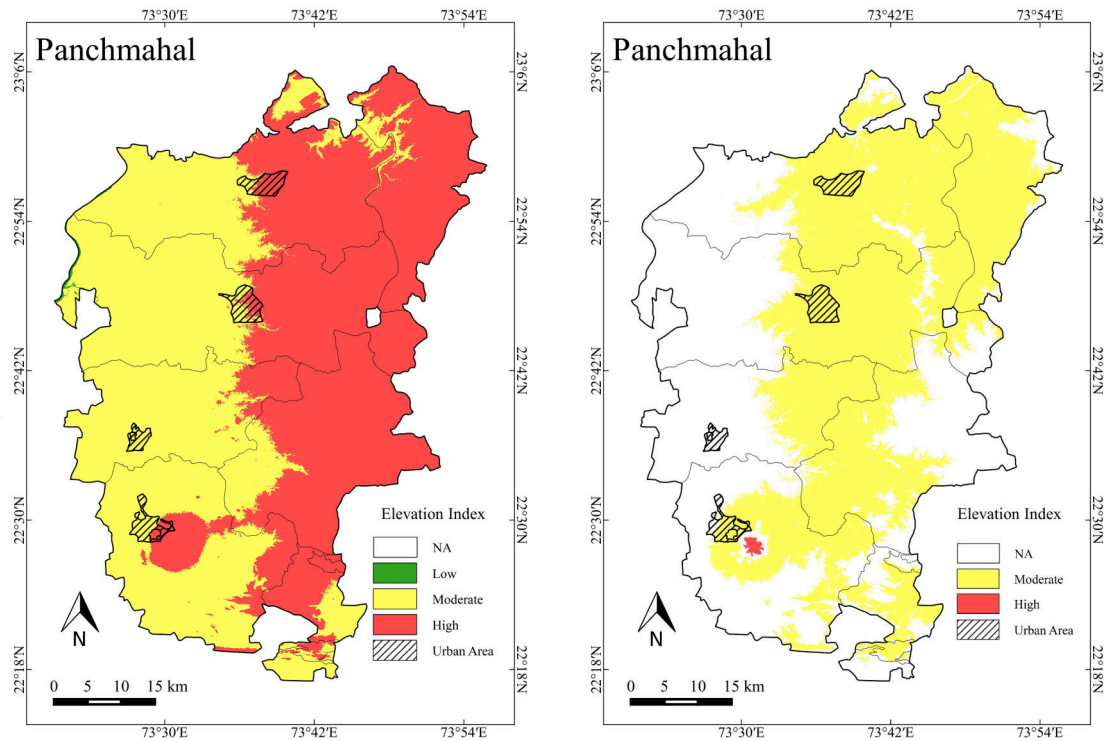
(a) Medalus approach

(b) Fuzzy Logic approach

Figure 4.4: Soil Index Map

4.4.4 Vegetation Index (VI)

Living of millions of individuals in India relies upon farming which is overall affected by locally accessible natural resources. In Gujarat as in other states of India, farming turns out to be the primary wellspring of means for most individuals. NDVI's high values represent high vegetation and it contributes less to the land degradation process while the sparse vegetation has a high commitment to the land degradation and henceforth fell into the low and moderate zone. The forest cover of the Panchmahal comprises around 23% of an absolute topographical regions of the district. 86.49% of the region is under a low weak zone for the area. Fig. 4.7(a) shows the vegetation index map of Panchmahal district for medalus model.



(a) Medalus approach

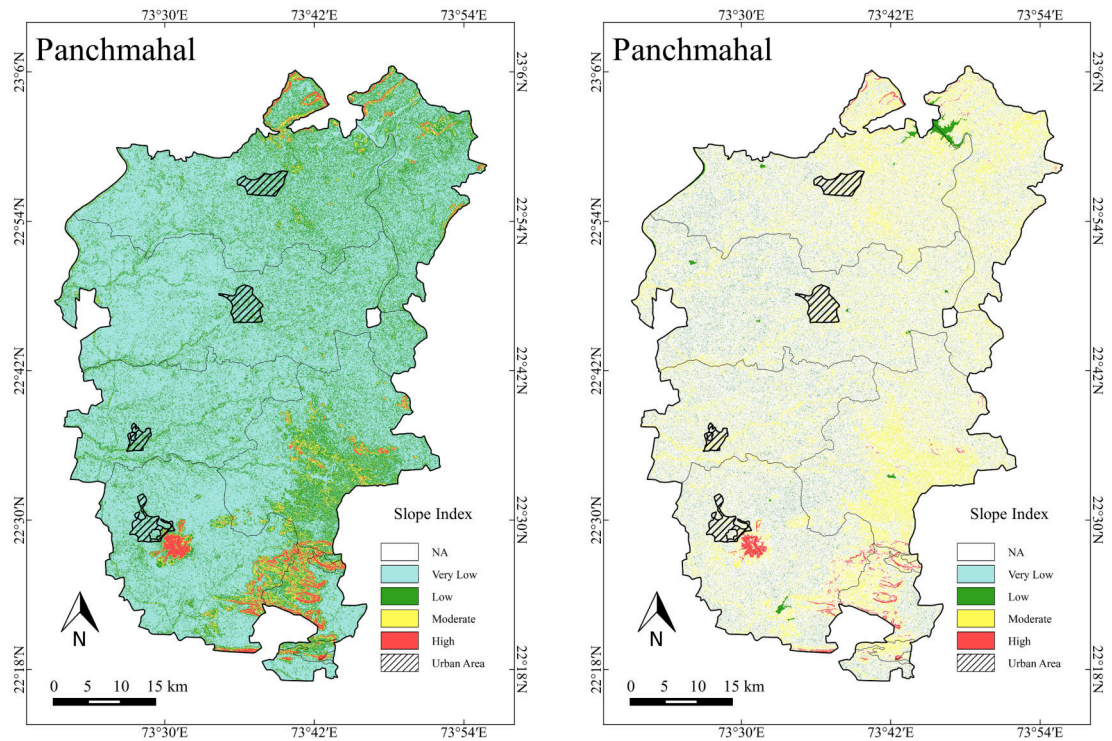
(b) Fuzzy Logic approach

Figure 4.5: Elevation Index Map

4.4.5 Socio-Economic Index (SEI)

Amenities like education, medical, communication facility, and transport network facilities are considered administrative necessities. Moreover, good amenities are a prerequisite for the social and economic development of any district. Facilities like education and healthcare were used to calculate the amenities index. Working and non-working populations of all the villages were considered for economic development index calculation.

The area of the Panchmahal district is mostly under low and very low vulnerability classes for the amenities index. 36.06% of area is under the moderate vulnerable zone for the district and less than 1% area is under high class for amenities index. Fig. 4.8 shows the amenities index map of Panchmahal district. Fig. 4.9 shows economic development index map in which more than 10% of the area is under the high-risk zone and 44.26% area under moderate class.



(a) Medalus approach

(b) Fuzzy Logic approach

Figure 4.6: Slope Index Map

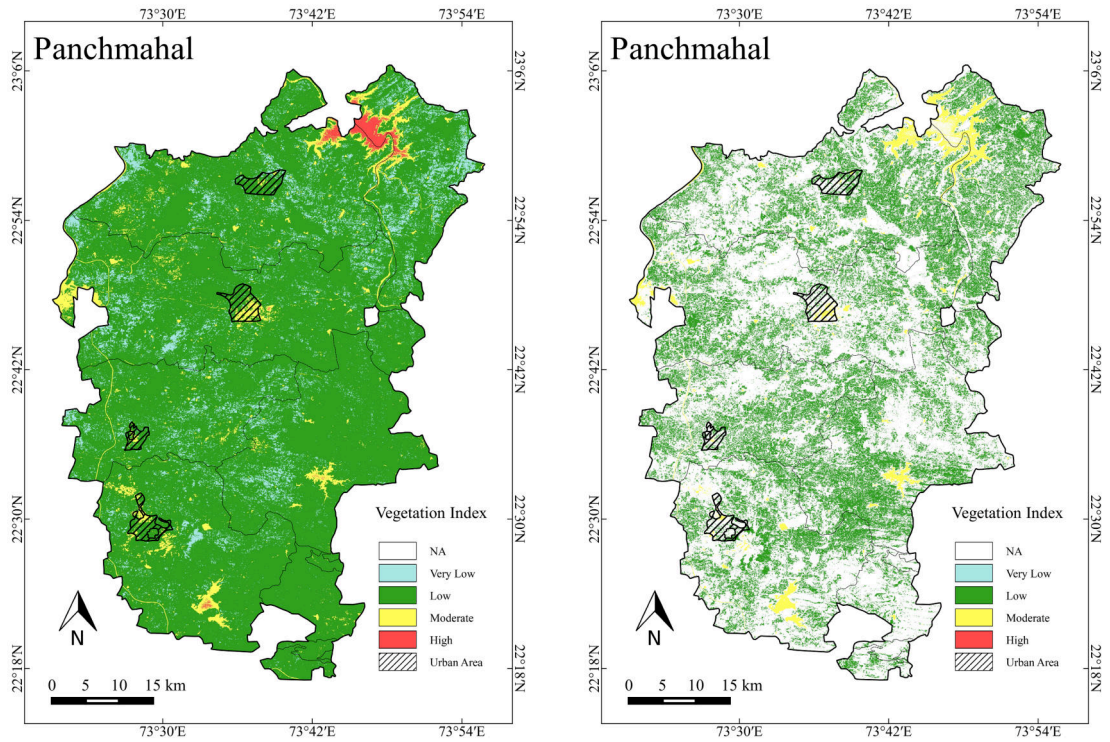
4.4.6 Land Utilization Index (LUI)

The area of the Panchmahal district is mostly covered by agricultural land and gives 67.52% of the total area of a district under a low vulnerability class. 21.28% of the area is under high vulnerable zone for the district. Fig. 4.10 shows the land utilization index map of Panchmahal district.

4.4.7 Desertification Vulnerability Index (DVI)

Environmental sensitive index (CI, SI, EI, VI, SloI, EI) and socio-economic index (CAI, EDI) were represented spatially in GIS environment to know the areas at risk of desertification with their severity.

Table. 4.11 represents the percentage area under different severity conditions for desertification vulnerability using two approaches: Medalus Approach and Fuzzy Logic Approach. Urban areas, water bodies, and rivers were masked in DVI map and marked as NA.



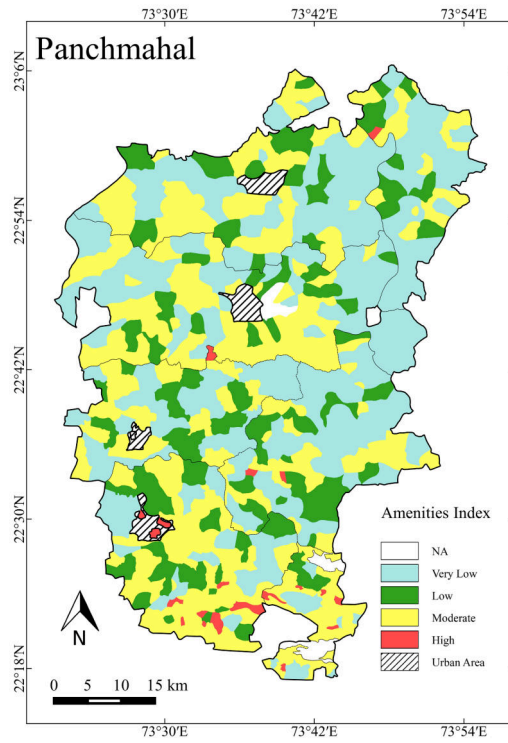
(a) Medalus approach

(b) Fuzzy Logic approach

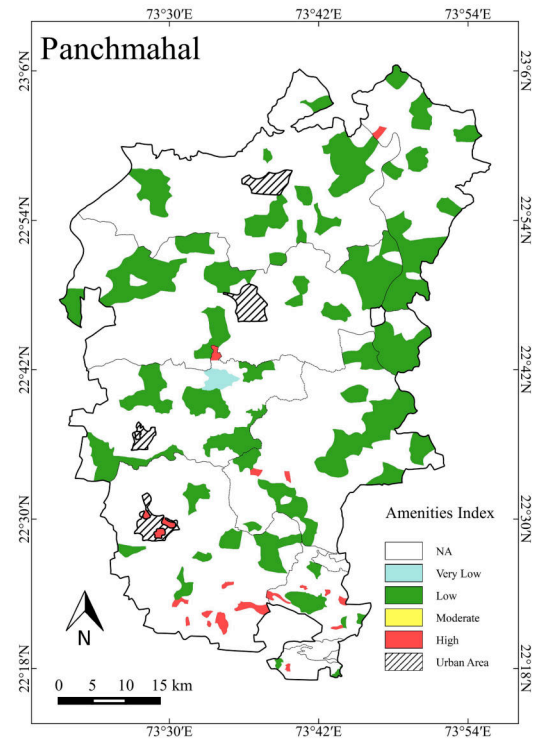
Figure 4.7: Vegetation Index Map

Table 4.11: Percentage area under different severity condition for desertification vulnerability

| | Medalus Approach | Fuzzy logic approach |
|----------|------------------|----------------------|
| NA | 9.67% | 9.77% |
| Very Low | 10.72% | 0.25% |
| Low | 35.53% | 24.76% |
| Moderate | 30.92% | 44.29% |
| High | 13.16% | 20.93% |

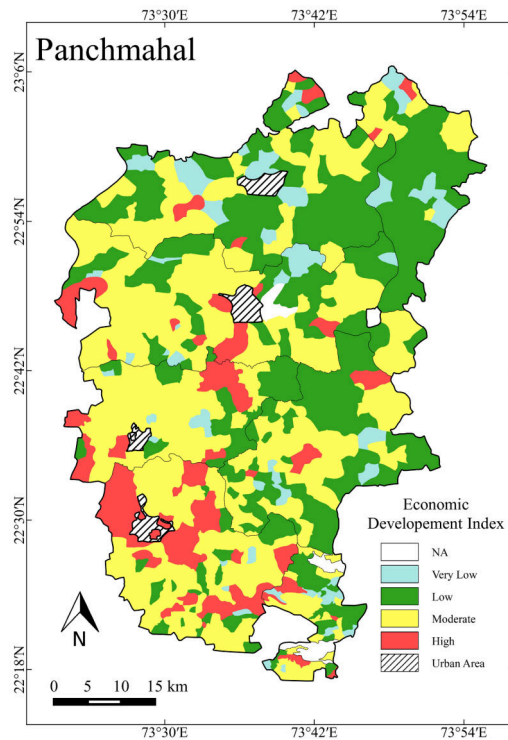


(a) Medalus approach

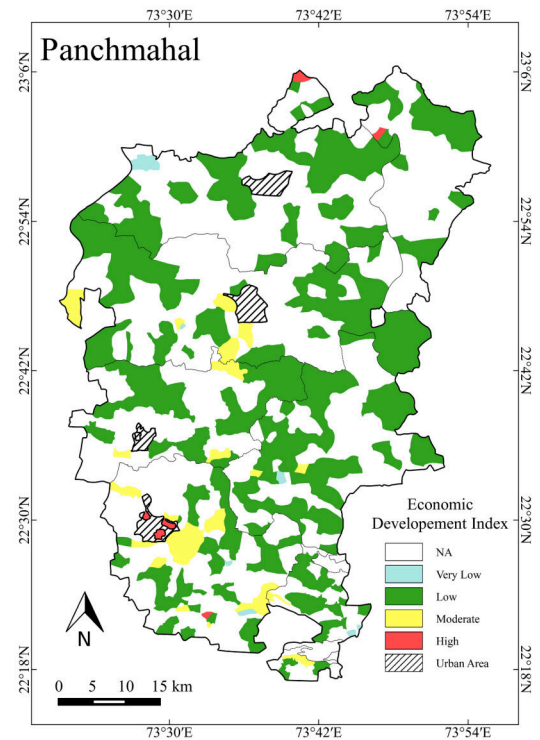


(b) Fuzzy Logic approach

Figure 4.8: Cumulative Amenities Index Map



(a) Medalus approach



(b) Fuzzy Logic approach

Figure 4.9: Economic Development Index Map

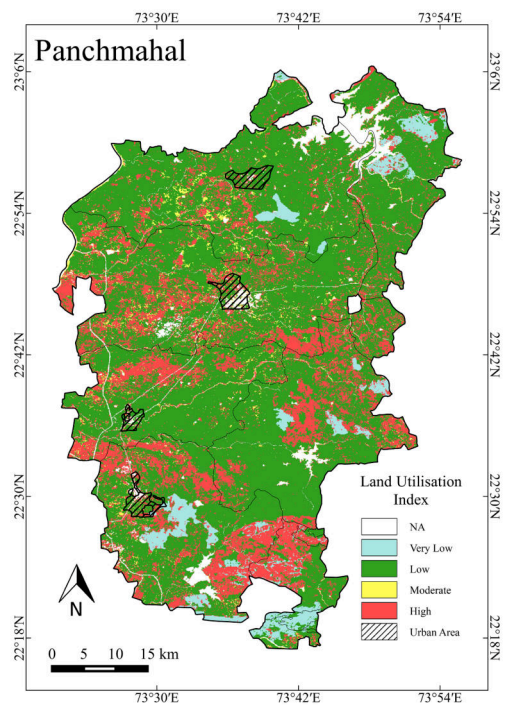
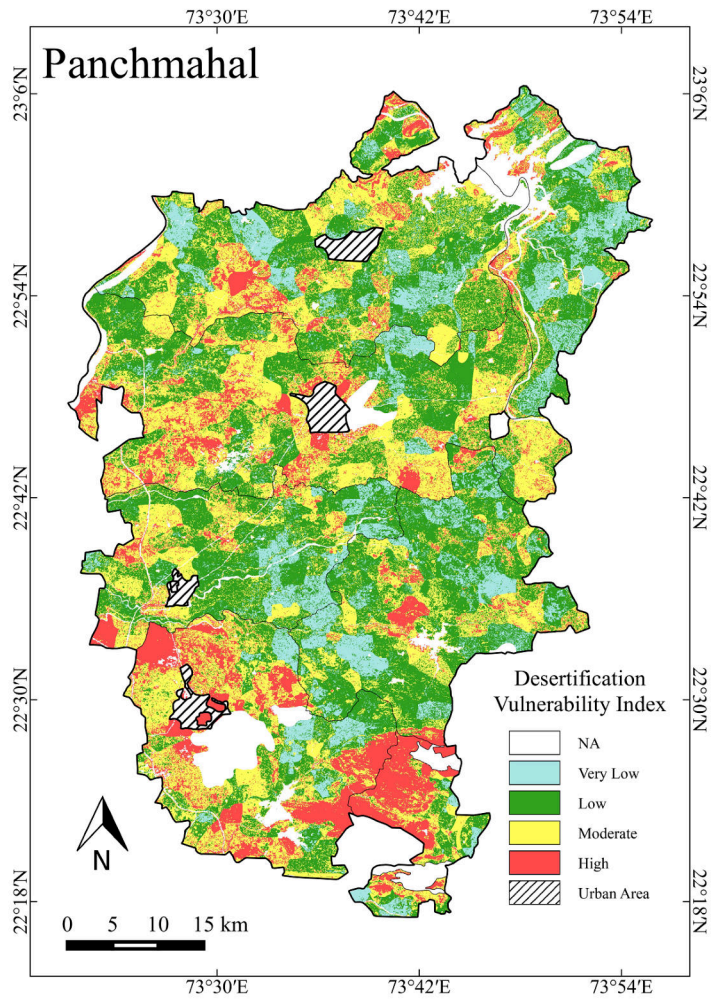
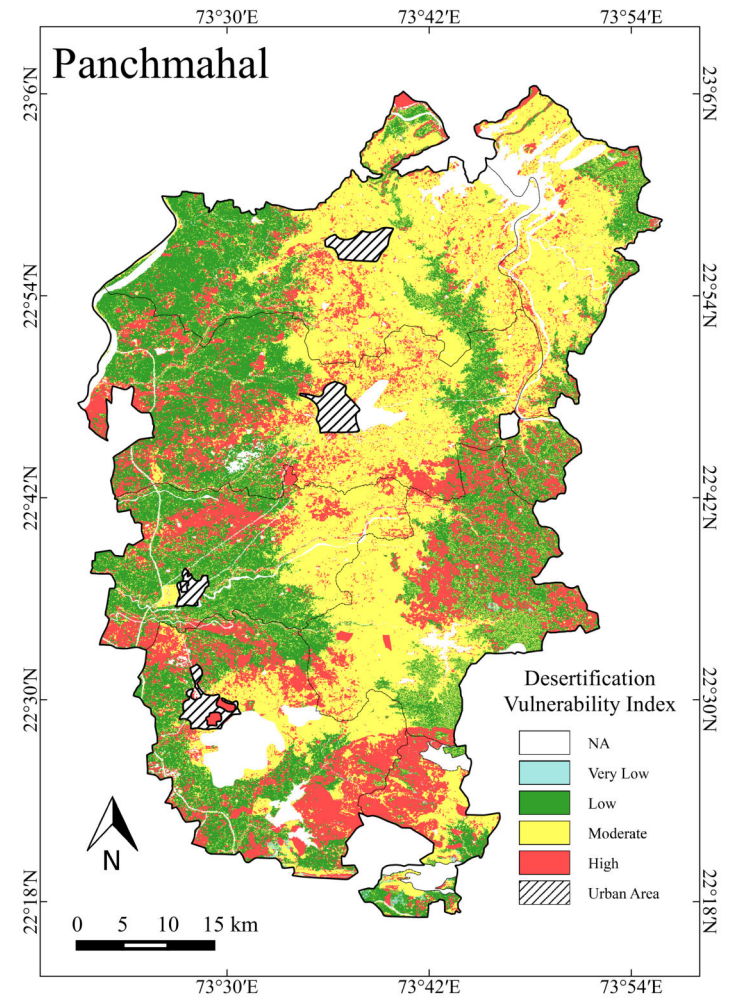


Figure 4.10: Land Utilization Index Map



(a) Medalus approach



(b) Fuzzy Logic approach

Figure 4.11: Desertification Vulnerability Index Map

The area of the Panchmahal district is mostly under a low severity level with around 35% of the total area. 13.16% of the area under high risk zone for desertification. 30.92% and 10.72% of area under moderate and very low-risk zone for desertification using Medalus model. In the fuzzy logic approach, mostly Panchmahal district is under moderate vulnerability with 44.29% area since lower altitude area is primarily out of risk comparatively having low vulnerability with nearly 25% area. and high alarming areas are in patches scattered with 21% of the total area as seen visible in Fig. 4.11 which shows the DVI map of the Panchmahal district for two different methods. Here using both methods the area is different for the high-risk zone. Still, the geographic location of these areas is the same, which indicates that this severity condition is easier to predict or map accurately. However, the lower severity conditions have more significant differences between the two approaches.

The results show that the Medalus approach tends to identify larger areas as being vulnerable to desertification in lower severity conditions (e.g., Very Low and Low). This may be due to the algorithm used to calculate the DVI, which may be more sensitive to socioeconomic indicators that are more prevalent in these areas. The results of the Fuzzy Logic approach for different severity conditions show some variation in the percentage area for each vulnerability class. The results show that the Fuzzy Logic approach tends to identify larger areas as being vulnerable to desertification in higher severity conditions (e.g., Moderate and High). This may be due to the use of fuzzy logic rules, which can handle imprecise or uncertain data and may be more sensitive to certain biophysical indicators like elevation and climate that are more prevalent in these areas.

Overall, both approaches have their strengths and weaknesses and can be useful in assessing desertification vulnerability.

The Medalus approach relies on the knowledge and expertise of local stakeholders, which can provide valuable insights into the specific conditions and factors contributing to desertification vulnerability. However, it can also be subjective and dependent on the quality and reliability of the expert input.

The Fuzzy Logic approach, on the other hand, relies on quantitative data anal-

ysis, which can provide a more objective and reproducible assessment of desertification vulnerability. However, it can be limited by the availability and quality of the data, and may not capture the full complexity of the system being studied.

The choice of approach may depend on the specific needs of the study and the level of detail required. It is also important to note that the accuracy and precision of the results may vary depending on the quality and availability of data, the assumptions and parameters used in the models, and the methods used to generate the maps.

CHAPTER 5

Desertification hot-spot using Aridity Index

Desertification is a phenomenon that occurs when an area that was once covered with productive vegetation turns into degraded land due to various natural and human reasons. Reduced precipitation and soil moisture, which can result in vegetation loss, soil degradation, and a drop in agricultural output, are the main causes of desertification. The risk of desertification in a particular area can be determined using the aridity index. Due to increased water stress and decreased plant cover, areas with a high aridity index are more likely to develop desertification. Desertification hot spots can be identified by regions with aridity indexes indicating arid, semi-arid, or hyper-arid conditions. The aridity index can also be used to pinpoint places that are already experiencing desertification as well as those that could be at risk in the future. These regions often experience significant land degradation and erosion, leading to the expansion of desertification[39, 150].

The aridity index is the numerical representation of the dryness of the climate for a particular location. The aridity index (AI) is a useful parameter to study desertification conditions and its pattern. The AI formulation has been adopted by UNEP, FAO, and UNCCD for explaining different situations. Eq.5.1 shows the definition of Aridity index[16].

$$AridityIndex(AI) = \frac{Precipitation(P)}{Potential\ evapotranspiration(PET)} \quad (5.1)$$

Potential evapotranspiration is defined as the idealized quantity of water evaporated and transpired by vegetation, soil and ecosystem, per-unit area, per unit time from an idealized, sufficient water surface under existing atmospheric con-

Table 5.1: The aridity zones based on Aridity index values

| Zone | Aridity Index |
|----------------|----------------|
| Hyperarid: | <0.05 (Desert) |
| Arid: | 0.05 to 0.20 |
| Semi-arid: | 0.20 to 0.5 |
| Dry sub-humid: | 0.5 to 0.65 |
| Humid: | >1 |

ditions. Precipitation defines the water vapor from the atmosphere which falls on the earth due to gravity. As per this index, different aridity zones are classified as given in Table 5.1[16]. Areas, other than polar and sub-polar regions, in which the ratio of annual precipitation to potential evapotranspiration falls within the range from 0.05 to 0.65 is called "arid, semi-arid, and dry sub-humid areas". Fig. 5.1 depicts the climatic zones delineated by Raju et al.[129] with the district area as a unit of study.

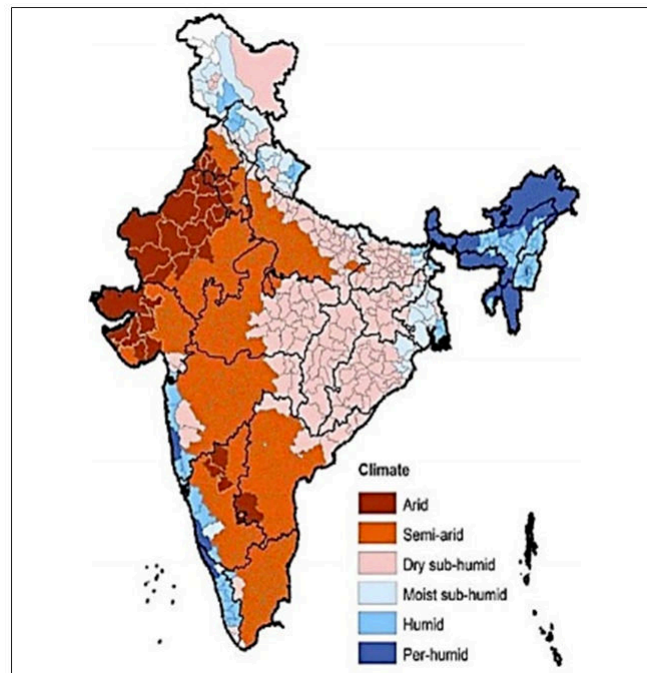


Figure 5.1: District level climate classification of aridity zones of India.[129]

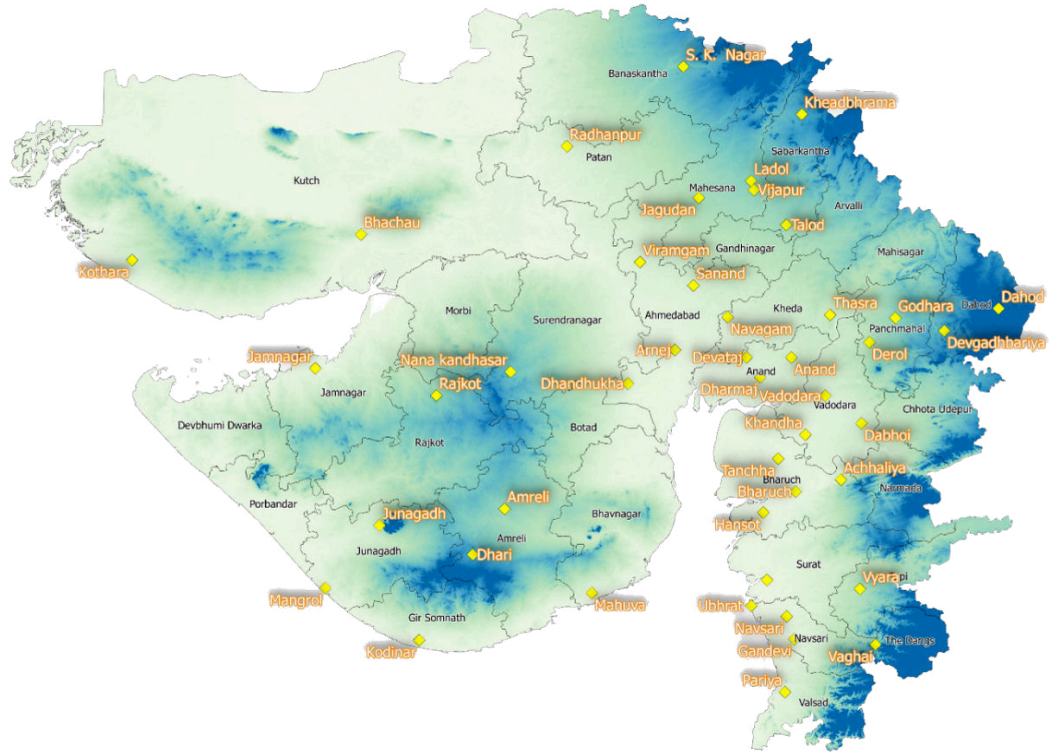


Figure 5.2: Meteorological observatories in Gujarat

5.1 Datasets

Weather data of Gujarat from meteorology observations for more than 18 locations during 1995-2015 has been used in this study. Fig.5.2 shows the locations of the meteorology stations in Gujarat. Meteorological parameters like minimum temperature, maximum temperature, rainfall, relative humidity, wind speed and wind direction and bright sunshine hour parameters were collected along with the geographic location of each station. In addition, the satellite-derived product like MODIS-Terra Normalized Difference Vegetation Index (NDVI) product for the past 20-year period of rabi season was used for the analysis.

5.2 Methodology

Data for the weather parameters were collected on daily basis. Some ambiguity and randomness were corrected in preprocessing and cleaning of data. Then daily data were transformed to weekly, monthly and yearly data to calculate the AI. Potential Evapotranspiration was calculated for all stations of Gujarat using FAO Penman-Monteith equation explained in eq. 5.3.[8] PET and rainfall were used to calculate AI for different locations. Spatial interpolation was performed to get continuous map from point data and an annual AI map for the whole state of Gujarat was generated using these values. Spatial interpolation is a technique used in GIS and remote sensing to estimate values of a variable (such as temperature, precipitation, or elevation) at locations where no direct measurements are available. This is done by using known values of the variable at nearby locations to estimate values at the unknown locations. kriging is a geostatistical interpolation method that uses a linear combination of nearby known data points to estimate the value of a variable at an unknown location. The general formula for the kriging interpolator is:

$$Z(s_0) = \sum_{i=1}^N \lambda_i Z(s_i) \quad (5.2)$$

where,

$Z(s_i)$ = the measured value at the i^{th} location, λ_i = an unknown weight for the measured value at the i^{th} location, s_0 = the prediction location, N = the number of measured values, $Z(s_0)$ = the predicted value at s_0 .

To develop an aridity index map for the Gujarat state from in-situ weather data, first ET_0 was calculated on yearly mean data for the stations in which required meteorological parameters were available. The FAO Penman-Monteith method was developed by adopting the Penman-Monteith combination method for reference evapotranspiration. The method gives more significant values with actual crop water use data worldwide. The FAO Penman-Monteith [8] method to estimate ET_0 is expressed as:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34 u_2)} \quad (5.3)$$

where,

ET_0 = reference evapotranspiration [mm/day],

R_n = net radiation at the crop surface [MJ/ m^2 day],

G = soil heat flux density [MJ/ m^2 day],

T = mean daily air temperature at 2 m height [$^{\circ}$ C],

u_2 = wind speed at 2 m height [m/s],

e_s = saturation vapour pressure [kPa],

e_a = actual vapour pressure [kPa],

D = slope vapour pressure curve [kPa/ $^{\circ}$ C],

g = psychrometric constant [kPa/ $^{\circ}$ C].

Utilizing rainfall data collected from observatories, AI was calculated for different stations of Gujarat for available stations. From point data, kriging interpolation was performed to get the AI map of Gujarat. Using this method the AI maps for the Gujarat region over the 20 years are developed. The calculated AI result was compared with the satellite-derived product NDVI. The Modis-Terra NDVI product for the same year's as of AI was taken and compared. The NDVI for the Gujarat region is processed and formatted to match at the same resolution as that of the AI map. The AI results are checked and correlated with the comparison for different times under drought and normal years according to meteorology throughout Gujarat State.

5.3 Results & Discussion

The state of Gujarat in India is characterized by different climatic zones based on aridity, which divided the region into three zones: arid, semi-arid, and sub-humid zones as explained in Table 5.1. Fig. 5.3(a)-(d) shows the calculated average potential evapo-transpiration, rainfall and aridity index map of Gujarat. The

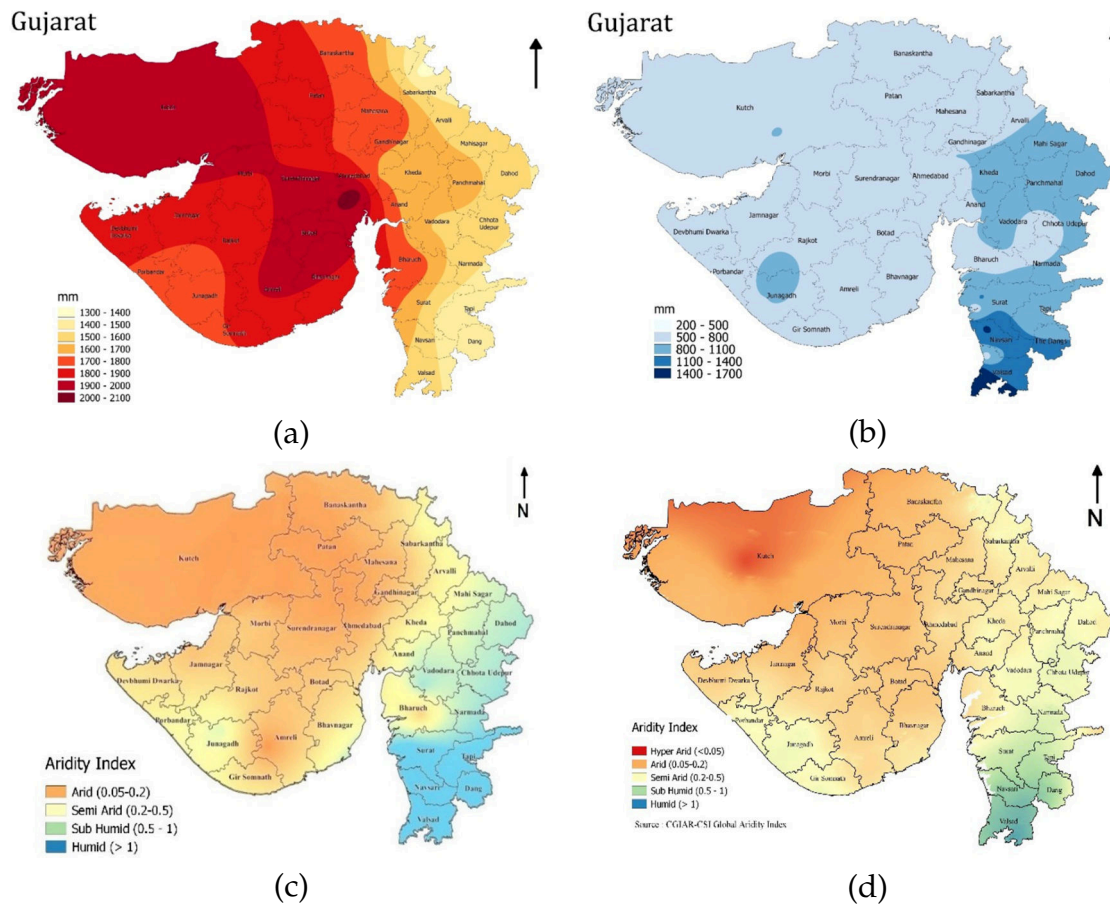


Figure 5.3: Average (a) PET, (b) rainfall and (c) aridity index map of Gujarat (d) CGIAR-CSI global aridity index

AI map represents the climatic zones of Gujarat. Aridity zones were classified as explained. Fig. 5.3(d) shows the aridity index map of Gujarat using the CGIAR-CSI Global Aridity Index database. The Global-Aridity is modeled using the data available from the WorldClim Global Climate Data as input parameters, for 1950-2000.[41] Visual comparison of the developed Aridity Index (AI) map with the globally available CGIAR-CSI AI map revealed small discontinuous patches of arid zones in Bharuch, Amreli and Ahmedabad districts although these districts are located in a semi-arid zone. These areas can be considered as desertification-prone zones. Similar pockets were also noted in north Gujarat and some parts of Saurashtra despite of them being located in semi-arid zone. Both the aridity index maps as shown in the Fig. 5.3(c) and 5.3(d) are estimated by the Penman-Monteith method but the developed map Fig. 5.3(c) has used real observatory data over eighteen meteorological observatories of Gujarat as shown in Fig.5.2.

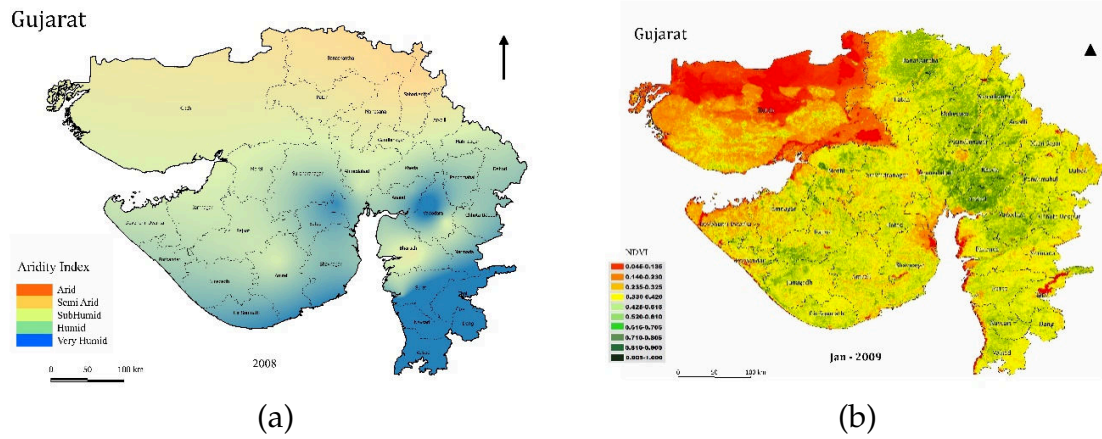


Figure 5.4: Comparison of (a) normalized difference vegetation index(NDVI)during rabi season of 2008 and (b) aridity index map of 2008

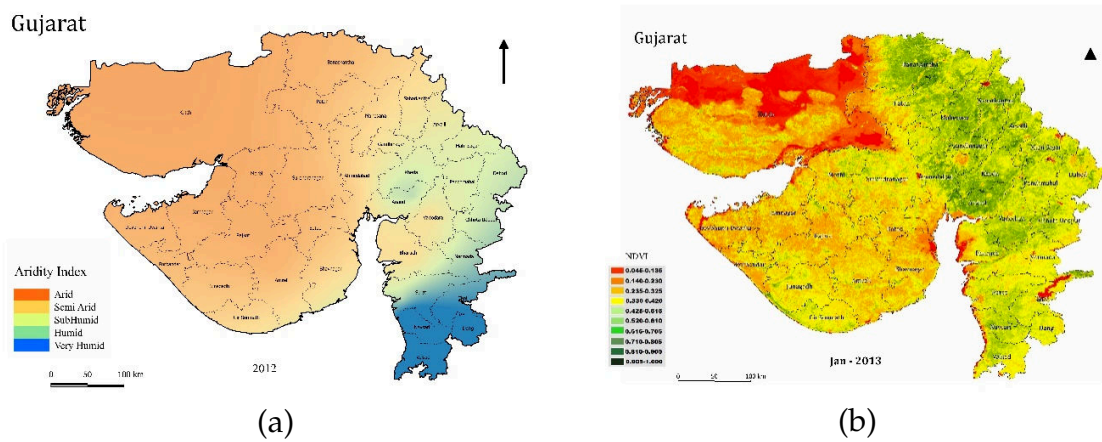


Figure 5.5: Comparison of (a) normalized difference vegetation index(NDVI) during rabi season of 2012 and (b) aridity index map of 2012

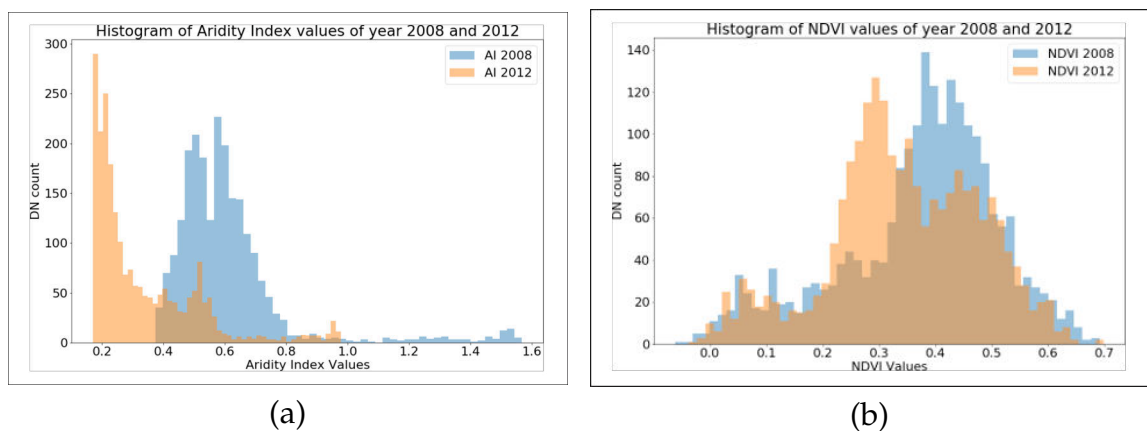


Figure 5.6: Comparison of Histogram of rabi NDVI and AI map for year (a) 2008 and (b) 2012.

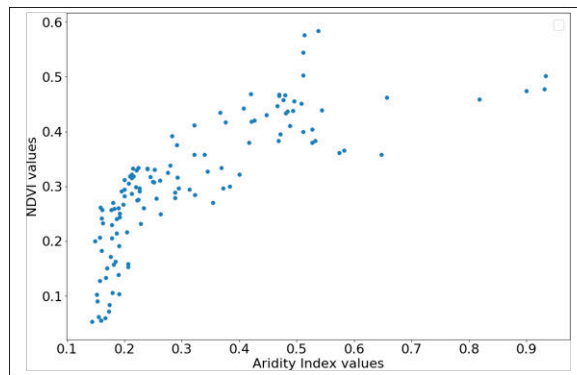


Figure 5.7: Relation between NDVI and AI value of random points for the year 2008 and 2012

As there is a very strong relationship between NDVI and rainfall. [54] So Annual AI and NDVI maps were used for comparing the trend under normal (2008) and drought (2012) situation as depicted in Fig. 5.4 and Fig. 5.5, respectively. In Fig. 5.4 NDVI results are compared with rabi data of normal rainfall year (2008). Comparing the AI map with NDVI map using pixel-based correlation gives a correlation coefficient value from 0.48 to 0.7. The results of the NDVI are compared with AI map for different years based on histogram method as shown in Fig. 5.6. The mean AI in the year 2008 is 0.73 and NDVI is 0.33 and in the year 2012 mean AI is 0.49 and NDVI is 0.31. The NDVI value decreases as the AI value decreases. To get 1:1 relation between NDVI and AI value of random points (sixty) were selected from the 2008 and 2012 year data (Fig. 5.7). Positive correlation of NDVI with AI, was observed because as the AI values increased the NDVI is also increases.

CHAPTER 6

Conclusion

The study conducted an analysis of desertification in two different areas, Panchmahal and Bhavnagar in Gujarat, India. In this study, a three-stage algorithm was proposed to calculate the severity and status of desertification. In the first stage, three different segmentation methods (Region Growing, Mean-shift, and SLIC) were used for the classification of Level 1 land cover. The accuracy comparison showed that SLIC was the most accurate method for this classification. Three different algorithms were then used for the classification: Support Vector Machine (SVM), Patch-based Convolutional Neural Network (CNN), and SLIC combined with Random Forest (RF). The results showed that all three algorithms achieved relatively high accuracy scores, with the SLIC+RF approach performing the best at 80%. The precision and recall scores for all three models were also quite high, with the SLIC+RF approach outperforming the other two in terms of precision and recall. Object-based classification was used for level 1 land cover classification, and the results showed that the overall accuracy for both areas was good, with Bhavnagar having a slightly higher accuracy (85%) than Panchmahal (80%). The results suggested that the complexity of the terrain and the heterogeneity of the land cover significantly affect the accuracy of the object-based classification. The study identified the desertification process in level 2 classification using RF and SVM algorithms. The models used the vegetation index to identify the extent of vegetation cover in the classification of forest vegetation degradation, and soil salinity index to classify the soil salinity. The study also compared the visual results of SLIC with RF classified results with the ground truth (GT) class map for level 1 land cover classification. The comparison showed that object-based

classification resulted in more cohesive and smooth polygons representing different land cover classes, as compared to pixel-based classification. The study concluded that the SLIC+RF approach is a promising method for level 1 classification and Random forest for level 2, particularly in terms of its ability to achieve high precision and recall scores. Overall, the study provided valuable insights into the severity of desertification in the study areas and demonstrated the effectiveness of the proposed three-stage algorithm in classifying land cover and assessing the desertification status. Covariates such as soil texture, climate-related variables, vegetation-related variables, and relief (topography) can be used to represent soil in predictive soil mapping (DSM) activities. Covariates selection involves reducing the input covariates by utilizing only relevant data and disposing of non-relevant data. Recursive feature elimination (RFE) is an algorithm that can be used to determine the importance of each covariate in the model and eliminate the redundant ones. Overall, The study discusses the use of machine learning in predicting soil properties by comparing different regression models in predicting soil pH, soil OC and soil EC, with the ANN model having the best performance in all cases. SCORPAN approach based soil forming factors were used as covariates. Covariates such as soil texture, climate-related variables, vegetation-related variables, and relief (topography) can be used to represent soil in predictive soil mapping (DSM) activities. RFE with linear regression was used to identify the most relevant covariates or features that are most predictive of the target soil properties. By removing non-relevant features, the model's complexity is reduced, and it becomes easier to interpret as proper covariates selection and preparation are essential for accurate and efficient. Additionally, reducing the number of features can help reduce the risk of overfitting and improve the model's generalization performance on new data. The results suggest that machine learning models can be effective in predicting soil properties. However, it is important to use appropriate evaluation metrics to accurately compare the performance of different models. The use of descriptive statistics such as minimum, maximum, mean, standard deviation, median, kurtosis, and skewness can also provide an overall view of the distribution of predicted values and help to identify potential outliers or biases in

the models. The study analyzed the desertification vulnerability of the Panchmahal district of Gujarat using various environmental and socio-economic indices. The study found that the district has a semi-arid climate with an aridity index ranging from 0.2 to 0.6. The soil of the area varies in their fruitfulness, and the western zones have richness in soils, while the eastern territories have shallow sandy soils. The district has little scattered sedimentary and volcanic dissected hills, and the elevation gradually diminishes from east to west. The forest cover of the district comprises around 23% of the total topographical area. The study assessed desertification vulnerability using two approaches: Medalus Approach and Fuzzy Logic Approach. The Medalus approach is a participatory approach that relies on the knowledge and expertise of local stakeholders, while the Fuzzy Logic approach is a data-driven approach that uses quantitative data analysis. The study used two sets of data to calculate the Desertification Vulnerability Index (DVI): the environmental sensitive index (ESI) and the socio-economic index (SEI). These two indices were represented spatially in a GIS environment to identify the areas at risk of desertification and their severity. The findings showed that the Panchmahal district is mostly under low severity level using the Medalus approach, while the Fuzzy Logic approach identified the district as primarily under moderate vulnerability. The high-risk zone had a similar geographic location using both methods, but there were significant differences in lower severity conditions. Lastly the study aimed to assess the aridity zones and potential desertification-prone areas based on aridity index maps for the Gujarat state. The region was divided into three zones based on aridity: arid, semi-arid, and sub-humid. The study used in-situ meteorological observatory data and the Penman-Monteith method to develop aridity index maps for Gujarat. The comparison of the developed map with the globally available CGIAR-CSI AI map revealed small patches of arid zones in some districts, indicating the potential for desertification. To assess the impact of aridity on vegetation, the study compared the aridity index maps with the normalized difference vegetation index (NDVI) during normal and drought situations in 2008 and 2012. The results showed a positive correlation between NDVI and aridity, as the AI values increased, the NDVI also increased. The

study found a correlation coefficient value is nearly 0.7 when comparing the AI map with the NDVI map using pixel-based correlation. The study highlights the importance of using aridity index maps to identify potential desertification-prone areas and assess the impact of aridity on vegetation.

6.1 Key Findings and Implications of the thesis

- The study analyzed desertification in two areas of Gujarat, India, using a three-stage algorithm that included object-based and pixel-based classification methods. The SLIC with RF approach was found to be the most accurate method for classifying level 1 land cover, achieving an accuracy score of 80%.
- Object-based classification resulted in more cohesive and smooth polygons representing different land cover classes, as compared to pixel-based classification.
- The study identified the desertification process in level 2 classification using RF and SVM algorithms for two different desertification processes eg. vegetation degradation in forest and salinity degradation in agriculture area. And found Random forest model using features focusing on the particular property is efficient for identifying the degradation process.
- Environment covariates used for predictive soil mapping and parameters such as relief (topography), climate-related variables, and vegetation-related variables are most important for soil properties prediction.
- The study highlighted the importance of proper covariates selection and preparation for accurate and efficient predictive soil mapping (DSM) activities. RFE with linear regression was used to identify the most relevant covariates or features that are most predictive of the target soil properties.
- The study found that machine learning models can be effective in predicting soil properties, but it is important to use appropriate evaluation metrics to

accurately compare the performance of different models.

- The study analyzed the desertification vulnerability of the Panchmahal district of Gujarat using various environmental and socio-economic indices, identifying the areas at risk of desertification and their severity.
- The Medalus approach and Fuzzy Logic approach were used to calculate the Desertification Vulnerability Index (DVI) based on environmental sensitivity and socio-economic indices. The Panchmahal district was mostly under low severity level using the Medalus approach, while the Fuzzy Logic approach identified the district as primarily under moderate vulnerability.
- The study assessed the aridity zones and potential desertification-prone areas based on aridity index maps for Gujarat state, identifying three zones: arid, semi-arid, and sub-humid.
- The study demonstrated the effectiveness of machine learning algorithms in analyzing and assessing desertification in different regions, providing valuable insights into the severity of desertification.

6.2 Future Work

- The study conducted a desertification assessment using remote sensing data for two areas in Gujarat, India. Future work could expand the analysis to other regions in India or other countries with similar environmental conditions to assess the effectiveness of the proposed algorithm for desertification assessment. This could be implemented on a state or regional level to monitor the desertification process and its severity over time continuously.
- To improve the accuracy of desertification pattern recognition, additional desertification processes such as water erosion, wind erosion, water-logging, urbanization and mass movement can be included in level 2 classification. This would enable a more comprehensive assessment of the extent and severity of desertification.

- Future work could explore different segmentation techniques and machine learning algorithms for classification to improve the classification and desertification assessment accuracy. Ensemble models, which combine other models, could also be explored having less complexity and faster computation power. Ensemble models could help reduce errors and improve the overall performance of the algorithm.
- The study used medium-resolution data, and future work could explore the use of high-resolution data, such as satellite images, to improve the accuracy of the classification and desertification assessment. High-resolution data can provide more detailed information about land cover and vegetation, which is crucial for assessing desertification.
- The study showed that machine learning models especially artificial neural networks can be effective in predicting soil properties. Future work could use these models for soil property mapping, which is essential for sustainable land management and agriculture. Soil property mapping can help identify areas of land degradation and assist in developing appropriate soil management strategies.
- The study used a limited number of meteorological observatory station points to calculate the aridity index for Gujarat. However, the aridity index is a useful tool for identifying and monitoring desertification hotspot regions. To locate more desertification-prone zones with greater accuracy, village-level climate data will be essential. The study also found a positive correlation between the aridity index and the Normalized Difference Vegetation Index (NDVI), indicating that vegetation indices are useful for monitoring desertification over time.
- Desertification is a global issue, and climate change is expected to exacerbate its effects. Future work could investigate the impact of climate change on desertification using machine learning and remote sensing techniques. This could help predict future desertification trends and assist in developing appropriate mitigation and adaptation strategies.

- The study used two approaches for desertification assessment. Future work could explore other approaches, such as machine learning-based or participatory approaches, for desertification assessment. Comparing different approaches could help determine the most effective approach for desertification assessment.
- The study assessed desertification vulnerability using environmental and socio-economic indices. Future work could investigate the impact of other environmental indices and socio-economic factors such as population growth, livestock information, transportation information, land use changes, and migration on desertification. This could help identify the drivers of desertification and assist in developing appropriate mitigation strategies.
- Develop a monitoring system: Desertification is a gradual process that can take years or even decades to manifest. Future work could develop a monitoring system using machine learning and remote sensing techniques to detect the early signs of desertification and prevent its spread. A monitoring system could help track the progress of desertification over time and assist in developing appropriate mitigation strategies.

References

- [1] Department of agriculture & cooperation ministry of agriculture government of india. In *Soil Testing in India*. New Delhi, India, 2011.
- [2] United nations, department of economic and social affairs, population division, world fertility report 2013: Fertility at the extremes. *Popul. Dev. Rev.*, 41(3):555–555, Sept. 2015.
- [3] All india report on number and area of operational holdings. In *Agricultural Census*, pages 2015–2016. New Delhi, 2019.
- [4] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.
- [5] P. Agyeman. Source apportionment, contamination levels, and spatial prediction of potentially toxic elements in selected soils of the czech republic. *Environmental Geochemistry and Health*, 43:601–620, 2020.
- [6] A. Ait Lamqadem, B. Pradhan, H. Saber, and A. Rahimi. Desertification sensitivity analysis using medalus model and gis: a case study of the oases of middle draa valley, morocco. *Sensors*, 18(7):2230, 2018.
- [7] Ajai, A. S. Arya, P. S. Dhinwa, S. K. Pathan, and K. G. Raj. Desertification/land degradation status mapping of india. *Current Science*, pages 1478–1483, 2009.
- [8] R. G. Allen, L. S. Pereira, D. Raes, M. Smith, et al. Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *Fao, Rome*, 300(9):D05109, 1998.

- [9] D. Arrouays, M. Grundy, N. Saby, N. McKenzie, J. Hempel, M. Martin, L. Boulonne, A. Bispo, O. Briand, A. C. Richer-de Forges, et al. Soil legacy data rescue via globalsoilmap and other international and national initiatives. *GeoResJ*, 14:1–19, 2017.
- [10] D. Arrouays, C. Le Bas, A. C. Richer-de Forges, M. Grundy, N. P. A. Saby, and C. Schwartz. Soil legacy maps of france: a new map representation for soil information. *European Journal of Soil Science*, 65(1):71–80, 2014.
- [11] E. Asfaw, K. Suryabhagavan, and M. Argaw. Soil salinity modeling and mapping using remote sensing and gis: The case of wonji sugar cane irrigation farm, ethiopia. *Journal of the Saudi Society of Agricultural Sciences*, 17(3):250–258, 2018.
- [12] A. Aubréville et al. Climats, forêts et désertification de l’afrique tropicale. 1949.
- [13] H. Bahrami, H. VAGHEEI, B. VAGHEEI, N. TAHMASEBIPOUR, and T. F. TALIEY. A new method for determining the soil erodibility factor based on fuzzy systems. 2005.
- [14] Z. G. Bai, D. L. Dent, L. Olsson, and M. E. Schaepman. Global assessment of land degradation and improvement 1. identification by remote sensing. Technical report, Report 2008, 2008.
- [15] S. A. Bangroo, G. R. Najar, E. Achin, and P. N. Truong. Application of predictor variables in spatial quantification of soil organic carbon and total nitrogen using regression kriging in the north kashmir forest himalayas. *Catena*, 193(104632):104632, Oct. 2020.
- [16] C. Barrow. World atlas of desertification (united nations environment programme), edited by n. middleton and dsg thomas. edward arnold, london, 1992. isbn 0 340 55512 2,£ 89.50 (hardback), ix+ 69 pp. *Land Degradation & Development*, 3(4):249–249, 1992.

- [17] N. Batjes. Soil vulnerability to diffuse pollution in central and eastern eu-
rope (soveur project) (version 1.0). Technical Report Report 2000/03, Food
and Agriculture Organization of the United Nations (FAO) and Interna-
tional Soil Reference and Information Centre (ISRIC), Wageningen, 2000.
- [18] N. H. Batjes. Mapping soil carbon stocks of central africa using soter. *Geo-
derma*, 146(1):58–65, 2008.
- [19] L. S. Bins, L. M. G. Fonseca, G. J. Erthal, and F. Mitsuo II. Satellite imagery
segmentation: A region growing approach. In *Proceedings of VII Simposio
Brasileiro de Sensoriamento Remoto*, pages 677–680, Salvador, May 1996.
- [20] T. Bishop and A. McBratney. A comparison of prediction methods for
the creation of field-extent soil property maps. *Geoderma*, 103(1-2):149–160,
2001.
- [21] B. R. Bonfatti, A. E. Hartemink, E. Giasson, C. G. Tornquist, and K. Adhikari.
Digital mapping of soil carbon in a viticultural region of southern brazil.
Geoderma, 261:204–221, 2016.
- [22] F. Boudjemline and A. Semar. Assessment and mapping of desertification
sensitivity with medalus model and gis–case study: basin of hodna, algeria.
Journal of water and land development, 36(1):17–26, 2018.
- [23] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [24] E. C. Brevik, A. Cerd'a, J. Mataix-Solera, L. Pereg, J. N. Quinton, J. Six, and
K. Van Oost. The interdisciplinary nature of soil. *SOIL*, 1:117–129, 2015.
- [25] E. M. Bridges and L. R. Oldeman. Global assessment of human-induced soil
degradation. *Arid soil research and rehabilitation*, 13(4):319–325, 1999.
- [26] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. Geographically
weighted regression: a method for exploring spatial nonstationarity. *Ge-
ographical analysis*, 28(4):281–298, 1996.

- [27] M. Budak, H. Günal, İ. Çelik, H. Yıldız, N. Acir, and M. Acar. Environmental sensitivity to desertification in northern mesopotamia; application of modified medalus by using analytical hierarchy process. *Arabian Journal of Geosciences*, 11(17):481, 2018.
- [28] E. N. Bui, C. J. Moran, and A. B. McBratney. Pedometrics 2007: Recent developments in quantitative pedology. *Geoderma*, 152(3-4):165–166, 2009.
- [29] M. Carranza-García, J. García-Gutiérrez, and J. C. Riquelme. A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sensing*, 11(3):274, 2019.
- [30] L. Census. *Government of India*. 19th Livestock Census-2012, 2012.
- [31] CensusofIndia. Village & town directory, gujarat, part xii-a & b, series-25, gujarat. 2011.
- [32] N. R. S. Centre-ISRO. Resourcesat-1 handbook. https://bhuvan.nrsc.gov.in/bhuvan/PDF/Resourcesat-1_Handbook.pdf, 2008.
- [33] P. Chaudhari, N. Desai, P. Chaudhari, and K. Rabari. Status of chemical properties and available major nutrients in soils of patan district of gujarat, india. *Crop Research*, 53(3and4):147–153, 2018.
- [34] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [35] A. Conacher. Land degradation and desertification: history, nature, causes, consequences, and solutions. *Theme*, 6:10, 2004.
- [36] R. T. Conant, M. G. Ryan, G. I. Ågren, H. E. Birge, E. A. Davidson, P. E. Eliasson, S. E. Evans, S. D. Frey, C. P. Giardina, F. M. Hopkins, et al. Temperature and soil organic matter decomposition rates—synthesis of current knowledge and a way forward. *Global change biology*, 17(11):3392–3404, 2011.

- [37] G. Connette, P. Oswald, M. Songer, and P. Leimgruber. Mapping distinct forest types improves overall forest identification based on multi-spectral landsat imagery for myanmars tanintharyi region. *Remote Sensing*, 8(11):882, 2016.
- [38] A. Dasgupta, K. Sastry, P. Dhinwa, V. Rathore, and M. Nathawat. Identifying desertification risk areas using fuzzy membership and geospatial technique—a case study, kota district, rajasthan. *Journal of Earth System Science*, 122(4):1107–1124, 2013.
- [39] V. Dave, M. Pandya, and R. Ghosh. Identification of desertification hot spot using aridity index. *Ann. Arid Zone*, 58(1–2):39–44, 2019.
- [40] V. A. Dave, M. Pandya, and R. Ghosh. An assessment of the desertification vulnerability based on medalus model. In *2019 International Conference on Intelligent Computing and Remote Sensing (ICICRS)*, pages 1–6. IEEE, 2019.
- [41] V. A. Dave and K. Sur. Fuzzy integrated desertification vulnerability model. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2018.
- [42] R. M. David, N. J. Rosser, and D. N. Donoghue. Improving above ground biomass estimates of southern africa dryland forests by combining sentinel-1 sar and sentinel-2 multispectral imagery. *Remote Sensing of Environment*, 282:113232, 2022.
- [43] F. De Paola, D. Ducci, and M. Giugni. Desertification and erosion sensitivity. a case study in southern italy: the tusciano river catchment. *Environmental earth sciences*, 70(5):2179–2190, 2013.
- [44] R. S. DeFries and J. Townshend. Ndvi-derived land cover classifications at a global scale. *International journal of remote sensing*, 15(17):3567–3586, 1994.
- [45] A. Dehni and M. Lounis. Remote sensing techniques for salt affected soil mapping: application to the oran region of algeria. *Procedia Engineering*, 33:188–198, 2012.

- [46] S. Dharumarajan, , M. Lalitha, R. Vasundhara, R. Hegde, , and and. The major biophysical indicators of desertification in arid and semi arid regions of india. *agroped*, 26(2), jun 2019.
- [47] S. Dharumarajan and R. Hegde. Digital mapping of soil texture classes using random forest classification algorithm. *Soil Use and Management*, 38(1):135–149, oct 2020.
- [48] S. Dharumarajan, R. Hegde, and S. Singh. Spatial prediction of major soil properties using random forest techniques - a case study in semi-arid tropics of south india. *Geoderma Regional*, 10:154–162, sep 2017.
- [49] S. Dharumarajan, B. Kalaiselvi, A. Suputhra, M. Lalitha, R. Vasundhara, K. S. A. Kumar, K. M. Nair, R. Hegde, S. K. Singh, and P. Lagacherie. Digital soil mapping of soil organic carbon stocks in western ghats, south india. *Geoderma Reg.*, 25(e00387):e00387, June 2021.
- [50] X. Ding, Z. Zhao, Q. Yang, L. Chen, Q. Tian, X. Li, and F.-R. Meng. Model prediction of depth-specific soil texture distributions with artificial neural network: A case study in yunfu, a typical area of udults zone, south china. *Computers and Electronics in Agriculture*, 169:105217, 2020.
- [51] V. V. Dokuchaev. Russian chernozem. 1883.
- [52] H. E. Dregne and N.-T. Chou. Global desertification dimensions and costs. *Degradation and restoration of arid lands*, 1:73–92, 1992.
- [53] I. Dronova, P. Gong, and L. Wang. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at poyang lake, china. *Remote sensing of Environment*, 115(12):3220–3236, 2011.
- [54] S. Dubey, G. Pranuthi, and S. Tripathi. Relationship between ndvi and rainfall relationship over india. *International Journal of Water Resources and Environmental Sciences*, 1(4):102–108, 2012.

- [55] O. Dubovyk. The role of remote sensing in land degradation assessments: opportunities and challenges. *European Journal of Remote Sensing*, 50(1):601–613, 2017.
- [56] D. C. Duro, S. E. Franklin, and M. G. Dubé. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using spot-5 hrg imagery. *Remote sensing of environment*, 118:259–272, 2012.
- [57] G. Fischer, H. van Velthuizen, M. Shah, and F. Nachtergaele. Global agro-ecological assessment for agriculture in the 21 century. Technical report, IIASA Research Report, IIASA, Laxenburg, 2002.
- [58] Food and Agriculture Organization of the United Nations. *The State of the World's Land and Water Resources for Food and Agriculture*. FAO, Rome, Italy, 2011.
- [59] S. Freire, N. de Lisboa, I. Fonseca, R. Brasil, J. Rocha, and J. A. Tenedório. Using artificial neural networks for digital soil mapping—a comparison of mlp and som approaches. In *AGILE*, 2013.
- [60] K. Fukunaga and L. D. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1):32–40, 1975.
- [61] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist. Novel algorithms for remote estimation of vegetation fraction. *Remote sensing of Environment*, 80(1):76–87, 2002.
- [62] M. Glantz. Desertification: environmental degradation in and around lands. *Westview Special Studies in Natural Resources and Energy Management (USA)*, 1977.
- [63] M. H. Glantz. Desertification. *Encyclopedia of World Climatology*, page 318324.

- [64] M. H. Glantz and N. S. Orlovsky. Desertification: A review of the concept. *Desertification Control Bulletin*, 9:15–22, 1983.
- [65] S. Grunwald and A. B. McBratney. Geographic information science in soil science: Trends and opportunities. *Geoderma*, 162(3-4):197–198, 2011.
- [66] M. Guevara, G. F. Olmedo, E. Stell, Y. Yigini, Y. Aguilar Duarte, C. Arel-lano Hernández, G. E. Arévalo, C. E. Arroyo-Cruz, A. Bolivar, S. Bunning, et al. No silver bullet for digital soil mapping: country-specific soil organic carbon estimates across latin america. *Soil*, 4(3):173–193, 2018.
- [67] P.-T. Guo, M.-F. Li, W. Luo, Q.-F. Tang, Z.-W. Liu, and Z.-M. Lin. Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach. *Geoderma*, 237-238:49–59, jan 2015.
- [68] S. K. Gurjar and V. Tare. Estimating long-term LULC changes in an agriculture-dominated basin using CORONA (1970) and LISS IV (2013–14) satellite images: a case study of ramganga river, india. *Environ Monit Assess*, 191(4), mar 2019.
- [69] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1):389–422, 2002.
- [70] T. Hastie and R. Tibshirani. Generalized additive models: Some applica-tions. *J. Am. Stat. Assoc.*, 82(398):371, June 1987.
- [71] T. Hengl, G. B. Heuvelink, and A. Stein. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 120(1-2):75–93, 2004.
- [72] T. Hengl, M. Nussbaum, M. N. Wright, G. B. Heuvelink, and B. Gräler. Ran-dom forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6:e5518, 2018.

- [73] R. Hijmans, S. Cameron, J. Parra, P. Jones, and A. Jarvis. The worldclim interpolated global terrestrial climate surfaces. version 1.3, 2004.
- [74] G. Hinge, R. Y. Surampalli, and M. K. Goyal. Prediction of soil organic carbon stock using digital mapping approach in humid india. *Environ Earth Sci*, 77(5), feb 2018.
- [75] B. D. Hudson. The soil survey as paradigm-based science. *Soil Science Society of America Journal*, 56(3):836–841, may 1992.
- [76] K. Jia, S. Liang, L. Zhang, X. Wei, Y. Yao, and X. Xie. Forest cover classification using landsat etm+ data and time series modis ndvi data. *International Journal of Applied Earth Observation and Geoinformation*, 33:32–38, 2014.
- [77] Z. Jiang, X. Ni, and M. Xing. A study on spatial and temporal dynamic changes of desertification in northern china from 2000 to 2020. *Remote Sensing*, 15(5):1368, 2023.
- [78] M. Jones. The organic matter content of the savanna soils of west africa. *Journal of Soil Science*, 24(1):42–53, 1973.
- [79] R. Jones, Y. L. Bissonnais, J. S. Diaz, O. Duwel, L. Oygarden, P. Prasuhn, Y. Yordanov, P. Strauss, B. Rydell, J. B. Uveges, G. Loj, M. Lane, and L. Vandekerckhove. *EU Soil Thematic Strategy: Technical working group on erosion, Work Package 2: Nature and extent of soil erosion in Europe*. 2003.
- [80] R. Kadović, Y. A. M. Bohajar, V. Perović, S. B. Simić, M. Todosijević, S. Tošić, M. Anđelić, D. Mlađan, and U. Dovezenski. Land sensitivity analysis of degradation using medalus model: Case study of deliblato sands, serbia. *Archives of Environmental Protection*, 2016.
- [81] J. G. Kalambukattu, S. Kumar, and R. Arya Raj. Digital soil mapping in a himalayan watershed using remote sensing and terrain parameters employing artificial neural network model. *Environmental earth sciences*, 77:1–14, 2018.

- [82] J. G. Kalambukattu, S. Kumar, and R. A. Raj. Digital soil mapping in a himalayan watershed using remote sensing and terrain parameters employing artificial neural network model. *Environ Earth Sci*, 77(5), mar 2018.
- [83] N. Kariminejad, H. R. Pourghasemi, M. Hosseinalizadeh, M. Rossi, and A. Mondini. Evaluating land degradation by gully erosion through soil erosion indices and rainfall thresholds. *Natural Hazards*, pages 1–17, 2023.
- [84] N. M. Khan, V. V. Rastoskuev, Y. Sato, and S. Shiozawa. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agricultural Water Management*, 77(1-3):96–109, 2005.
- [85] C. Kosmas, A. Ferrara, H. Briasouli, and A. Imeson. Methodology for mapping environmentally sensitive areas (esas) to desertification. *The medalus project Mediterranean desertification and land use. Manual on key indicators of desertification and mapping Environmentally Sensitive Areas to desertification*, pages 31–47, 1999.
- [86] C. Kosmas, O. Kairis, C. Karavitis, C. Ritsema, L. Salvati, S. Acikalin, M. Alcalá, P. Alfama, J. Atlhopheng, J. Barrera, et al. Evaluation and selection of indicators for land degradation and desertification monitoring: methodological approach. *Environmental management*, 54:951–970, 2014.
- [87] A. Krishna. Land cover change dynamics of a himalayan watershed utilizing indian remote sensing satellite (IRS) data. In *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*. IEEE.
- [88] P. Lagacherie, D. Arrouays, H. Bourennane, C. Gomez, and L. Nkuba-Kasanda. Analysing the impact of soil spatial sampling on the performances of digital soil mapping models and their evaluation: A numerical experiment on quantile random forest using clay contents obtained from vis-nir-swir hyperspectral imagery. *Geoderma*, 375:114503, 2020.
- [89] P. Lagacherie and A. McBratney. Chapter 1 spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. In *Developments in Soil Science*, pages 3–22. Elsevier, 2006.

- [90] P. Lagacherie, A. B. McBratney, and M. Voltz, editors. *Digital Soil Mapping: An Introductory Perspective*. Elsevier, 2007.
- [91] H. Lahlaoui, H. Rhinane, A. Hilali, S. Lahssini, and S. Moukrim. Desertification assessment using medalus model in watershed oued el maleh, morocco. *Geosciences*, 7(3):50, 2017.
- [92] M. Lalitha, S. Dharumarajan, A. Suputhra, B. Kalaiselvi, R. Hegde, R. Reddy, C. S. Prasad, C. Harindranath, and B. Dwivedi. Spatial prediction of soil depth using environmental covariates by quantile regression forest model. *Environ Monit Assess*, 193(10), sep 2021.
- [93] E. F. Lambin, H. J. Geist, and E. Lepers. Dynamics of land-use and land-cover change in tropical regions. *Annual review of environment and resources*, 28(1):205–241, 2003.
- [94] S. Lamichhane, L. Kumar, and B. Wilson. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma*, 352:395–413, oct 2019.
- [95] L. Lavauden. La desertification de la france méridionale. *Les forêts du Sahara*. Berger-Levrault, Paris, France, 1927.
- [96] X. Li and G. W. McCarty. Application of topographic analyses for mapping spatial patterns of soil properties. *Geospatial Analyses of Earth Observation (EO) data*, 2019.
- [97] A. Liaw, M. Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.
- [98] W. Lockeretz. The lessons of the dust bowl: Several decades before the current concern with environmental problems, dust storms ravaged the great plains, and the threat of more dust storms still hangs over us. *American Scientist*, 66(5):560–569, 1978.
- [99] MA. *Ecosystems and Human Well-being: Synthesis Report*. Island Press, Washington DC, 2005.

- [100] MA. *Ecosystems and Humans Well-Being: Desertification Synthesis*. Island Press, Washington DC, 2005.
- [101] B. Mannan, J. Roy, and A. K. Ray. Fuzzy ARTMAP supervised classification of multi-spectral remotely-sensed images. *International Journal of Remote Sensing*, 19(4):767–774, jan 1998.
- [102] J. F. Mas and J. J. Flores. The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, 29(3):617–663, 2008.
- [103] A. Mathur and G. Foody. Multiclass and binary SVM classification: Implications for training and classification users. *IEEE Geosci. Remote Sensing Lett.*, 5(2):241–245, apr 2008.
- [104] A. E. Maxwell, T. A. Warner, and F. Fang. Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, 39(9):2784–2817, feb 2018.
- [105] A. B. McBratney, I. O. Odeh, T. F. Bishop, M. S. Dunbar, and T. M. Shatar. An overview of pedometric techniques for use in soil survey. *Geoderma*, 97(3-4):293–327, 2000.
- [106] A. B. McBratney, M. M. Santos, and B. Minasny. On digital soil mapping. *Geoderma*, 117(1-2):3–52, 2003.
- [107] M. D. d. Menezes, S. H. G. Silva, C. R. d. Mello, P. R. Owens, and N. Curi. Spatial prediction of soil properties in two contrasting physiographic regions in brazil. *Scientia Agricola*, 73:274–285, 2016.
- [108] C. Michael, H. Charles, R. James, J. HILL, S. Stefan, and V. M. Graham. World atlas of desertification. 2018.
- [109] N. J. Middleton and D. S. G. Thomas. *World Atlas of Desertification: United Nations Environmental Programme*. Arnold, 1992.

- [110] B. Minasny and A. B. McBratney. Digital soil mapping: a brief history and some lessons. *Geoderma*, 264:301–311, 2016.
- [111] T. Mitran, U. Mishra, R. Lal, T. Ravisankar, and K. Sreenivas. Spatial distribution of soil carbon stocks in a semi-arid region of india. *Geoderma Regional*, 15:e00192, dec 2018.
- [112] A. Mondal, D. Khare, S. Kundu, S. Mondal, S. Mukherjee, and A. Mukhopadhyay. Spatial soil organic carbon (soc) prediction by regression kriging using remote sensing data. *The Egyptian Journal of Remote Sensing and Space Science*, 20(1):61–70, 2017.
- [113] G. Mountrakis, J. Im, and C. Ogole. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247–259, 2011.
- [114] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote sensing of environment*, 115(5):1145–1161, 2011.
- [115] F. O. F. Nachtergaele and C. Licona-Manzur. The land degradation assessment in drylands (LADA) project: Reflections on indicators for land degradation assessment. In *The Future of Drylands*, pages 327–348. Springer Netherlands.
- [116] S. NBSS&LUP. Soils of india. *National Bureau of Soil Survey and Land Use Planning, Nagpur, India, NBSS Publ*, 94:130, 2002.
- [117] K.-A. Nguyen, Y.-A. Liou, H.-P. Tran, P.-P. Hoang, and T.-H. Nguyen. Soil salinity assessment by using near-infrared channel and vegetation soil salinity index derived from landsat 8 oli data: a case study in the tra vinh province, mekong delta, vietnam. *Progress in Earth and Planetary Science*, 7:1–16, 2020.

- [118] L. NRSC. Land use/land cover database on 1: 50,000 scale. *Natural Resources Census Project, LUCMD, LRUMG, RSAA, National Remote Sensing Centre, ISRO, Hyderabad. Natural Resource Census-Land Use Land Cover, Ver, 1:1–11, 2014.*
- [119] G. of India. *State of Environment Report India-2009*. Environmental Information System (ENVIS), Ministry of Environment Forests, 2009.
- [120] E. I. Okonkwo, R. Corstanje, and G. J. Kirk. Digital soil assessment for quantifying soil constraints to crop production: a case study for rice in punjab, india. *Soil Use and Management*, 34(4):533–541, 2018.
- [121] E. I. Okonkwo, R. Corstanje, and G. J. D. Kirk. Digital soil assessment for quantifying soil constraints to crop production: a case study for rice in punjab, india. *Soil Use Manage*, 34(4):533–541, sep 2018.
- [122] D. Omaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.
- [123] P. Panagos, C. Ballabio, Y. Yigini, and M. B. Dunbar. Estimating the soil organic carbon content for european nuts2 regions based on lucas data collection. *Science of the Total Environment*, 442:235–246, 2013.
- [124] M. Pandya, V. Dave, and R. Ghosh. Artificial neural network (ann) based soil electrical conductivity (sec) prediction. In *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 581–586, 2020.
- [125] P. Pani and P. Carling. Land degradation and spatial vulnerabilities: a study of inter-village differences in chambal valley, india. *Asian Geographer*, 30(1):65–79, jun 2013.
- [126] R. Prăvălie, I. Săvulescu, C. Patriche, M. Dumitrașcu, and G. Bandoc. Spatial assessment of land degradation sensitive areas in southwestern romania using modified medalus method. *Catena*, 153:114–130, 2017.

- [127] U. N. E. Programme. *Desertification control bulletin number 21, 1992: A bulletin of world events in the control of desertification restoration of degraded lands and reforestation*. Jan 1992.
- [128] M. R. P. Rad, N. Toomanian, F. Khormali, C. W. Brungard, C. B. Komaki, and P. Bogaert. Updating soil survey maps using random forest and conditioned latin hypercube sampling in the loess derived soils of northern iran. *Geoderma*, 232-234:97–106, nov 2014.
- [129] B. M. Raju, K. V. S. Rao, B. Venkateswarlu, A. V. M. S. Rao, C. A. R. Rao, V. U. M. Rao, B. B. Rao, N. R. Kumar, R. Dhakar, N. Swapna, and P. S. Latha. Revisiting climatic classification in india : a district-level analysis. 2013.
- [130] M. Rashid, M. A. Lone, and S. A. Romshoo. Geospatial tools for assessing land degradation in budgam district, kashmir himalaya, india. *Journal of earth system science*, 120(3):423–433, 2011.
- [131] N. N. Reddy, P. Chakraborty, S. Roy, K. Singh, B. Minasny, A. B. McBratney, A. Biswas, and B. S. Das. Legacy data-based national-scale digital mapping of key soil properties in india. *Geoderma*, 381:114684, jan 2021.
- [132] V. R. Reddy. Land degradation in india: Extent, costs and determinants. *Economic and Political Weekly*, 38(46):4700–4713, 2003.
- [133] Z. Ren, H. Zheng, X. He, D. Zhang, X. Yu, and G. Shen. Spatial estimation of urban forest structures with landsat tm data and field measurements. *Urban Forestry & Urban Greening*, 14(2):336–344, 2015.
- [134] M. E. Romero-Sanchez and R. Ponce-Hernandez. Assessing and monitoring forest degradation in a deciduous tropical forest in mexico via remote sensing indicators. *Forests*, 8(9):302, 2017.
- [135] SAARC. *Strategies for Arresting Land Degradation in South Asian Countries*. SAARC Agriculture Centre, BARC Complex, Farmgate Dhaka, Bangladesh, 2011.

- [136] I. SAC. Desertification and land degradation atlas of india (based on irs awifs data of 2011–13 and 2003–05). *Space Applications Centre, ISRO, Ahmedabad, India, Ahmedabad*, 2016.
- [137] U. N. Safriel. The assessment of global trends in land degradation. In *Climate and Land Degradation*, pages 1–38. Springer Berlin Heidelberg.
- [138] A. Saha, S. C. Pal, I. Chowdhuri, A. R. M. T. Islam, P. Roy, and R. Chakraborty. Land degradation risk dynamics assessment in red and lateritic zones of eastern plateau, india: a combine approach of k-fold cv, data mining and field validation. *Ecological Informatics*, 69:101653, 2022.
- [139] L. Salvati and M. Zitti. Convergence or divergence in desertification risk? scale-based assessment and policy implications in a mediterranean country. *Journal of environmental planning and management*, 52(7):957–971, 2009.
- [140] P. A. Sanchez and E. C. Brevik. Digital soil mapping across the globe. *Geoderma*, 152(3-4):253–258, 2009.
- [141] P. Santra, R. Singh, M. C. Sarathjith, N. R. Panwar, P. Varghese, and B. S. Das. Reflectance spectroscopic approach for estimation of soil properties in hot arid western rajasthan, india. *Environ Earth Sci*, 74(5):4233–4245, apr 2015.
- [142] S. Sarkar, A. K. Roy, and T. R. Martha. Soil depth estimation through soil-landscape modelling using regression kriging in a himalayan terrain. *Geogr. Inf. Syst.*, 27(12):2436–2454, Dec. 2013.
- [143] K. Sasikala, M. Petrou, and J. Kittler. Fuzzy classification with a gis as an aid to decision making. *EARSel Advances in remote sensing*, 4:97–105, 1996.
- [144] M. Scaioni, L. Longoni, V. Melillo, and M. Papini. Remote sensing for landslide investigations: An overview of recent achievements and perspectives. *Remote Sensing*, 6(10):9600–9652, oct 2014.

- [145] J. P. Scharlemann, E. V. Tanner, R. Hiederer, and V. Kapos. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, 5(1):81–91, 2014.
- [146] P. Scull, J. Franklin, O. A. Chadwick, and D. McArthur. Predictive soil mapping: a review. *Progress in Physical geography*, 27(2):171–197, 2003.
- [147] J. Sehgal and I. Abrol. *Soil degradation in India: status and impact*. Oxford IBH Publishing Co., 1994.
- [148] H.-c. Shih, D. A. Stow, and Y. H. Tsai. Guidance on and comparison of machine learning classifiers for landsat-based land cover and land use mapping. *International Journal of Remote Sensing*, 40(4):1248–1274, 2019.
- [149] G. Shukla, R. D. Garg, H. S. Srivastava, and P. K. Garg. An effective implementation and assessment of a random forest classifier as a soil spatial predictive model. *International Journal of Remote Sensing*, 39(8):2637–2669, jan 2018.
- [150] R. B. Singh and Ajai. A composite method to identify desertification hotspots and brightspots. *Land Degradation & Development*, 30(9):1025–1039, 2019.
- [151] S. K. Singh and S. Chatterji. Land resource inventory for agricultural land use planning using geospatial techniques. In *Geotechnologies and the Environment*, pages 163–183. Springer International Publishing, 2018.
- [152] P. Smith. How long before a change in soil organic carbon can be detected? *Global Change Biology*, 10(11):1878–1883, 2004.
- [153] S. Sommer, C. Zucca, A. Grainger, M. Cherlet, R. Zougmore, Y. Sokona, J. Hill, R. Della Peruta, J. Roehrig, and G. Wang. Application of indicator systems for monitoring and assessment of desertification from national to global scales. *Land Degradation & Development*, 22(2):184–197, 2011.

- [154] Space Applications Centre, ISRO. *Desertification and Land Degradation Atlas of India (Assessment and analysis of changes over 15 years based on remote sensing)*. Ahmedabad, India, 2021.
- [155] K. Sreenivas, V. Dadhwal, S. Kumar, G. S. Harsha, T. Mitran, G. Sujatha, G. J. R. Suresh, M. Fyzee, and T. Ravisankar. Digital mapping of soil organic and inorganic carbon status in india. *Geoderma*, 269:160–173, may 2016.
- [156] K. Sreenivas, G. Sujatha, K. Sudhir, D. V. Kiran, M. A. Fyzee, T. Ravisankar, and V. K. Dadhwal. Spatial assessment of soil organic carbon density through random forests based imputation. *J Indian Soc Remote Sens*, 42(3):577–587, feb 2014.
- [157] E. P. Stebbing. The encroaching sahara: the threat to the west african colonies. *The Geographical Journal*, 85(6):506–519, 1935.
- [158] W. L. Stefanov and M. Netzband. Assessment of ASTER land cover and MODIS NDVI data at multiple scales for ecological characterization of an arid urban center. *Remote Sensing of Environment*, 99(1-2):31–43, nov 2005.
- [159] Z. Sun, J. Wang, R. Li, and C. Tong. A new kriging based learning function and its application to structural reliability analysis. *Reliability Engineering & System Safety*, 157:152–165, 2017.
- [160] E. Symeonakis, N. Karathanasis, S. Koukoulas, and G. Panagopoulos. Monitoring sensitivity to land degradation and desertification with the environmentally sensitive area index: The case of lesvos island. *Land Degradation & Development*, 27(6):1562–1573, 2016.
- [161] R. Taghizadeh-Mehrjardi, K. Schmidt, N. Toomanian, B. Heung, T. Behrens, A. Mosavi, S. S. Band, A. Amirian-Chakan, A. Fathabadi, and T. Scholten. Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models. *Geoderma*, 383:114793, feb 2021.

- [162] S. Talukdar, P. Singha, Shahfahad, S. Mahato, B. Praveen, and A. Rahman. Dynamics of ecosystem services (ESs) in response to land use land cover (LU/LC) changes in the lower gangetic plain of india. *Ecological Indicators*, 112:106121, may 2020.
- [163] Z. Tan, R. Lal, N. Smeck, and F. Calhoun. Relationships between surface soil organic carbon pool and site variables. *Geoderma*, 121(3-4):187–195, 2004.
- [164] P. Teluguntla, P. S. Thenkabail, A. Oliphant, J. Xiong, M. K. Gumma, R. G. Congalton, K. Yadav, and A. Huete. A 30-m landsat-derived cropland extent product of australia and china using random forest machine learning algorithm on google earth engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144:325–340, 2018.
- [165] S. K. Tiwari, S. K. Saha, S. Kumar, et al. Prediction modeling and mapping of soil carbon content using artificial neural network, hyperspectral satellite data and field spectroscopy. *Advances in Remote Sensing*, 4(01):63, 2015.
- [166] S. I. Toure, D. A. Stow, H. chien Shih, J. Weeks, and D. Lopez-Carr. Land cover and land use change analysis using multi-spatial resolution data and object-based image analysis. *Remote Sensing of Environment*, 210:259–268, jun 2018.
- [167] A. Trabucco and R. Zomer. Global potential evapo-transpiration (global-pet) and global aridity index (global-aridity) geo-database. *CGIAR consortium for spatial information*, 2009.
- [168] J. Triantafilis and S. Buchanan. Exploring the use of geophysics for assessing soil properties. *Journal of Environmental Management*, 91(4):808–815, 2010.
- [169] C. J. Tucker. Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, 8(2):127–150, 1979.
- [170] UNCCD. United nations convention to combat desertification in countries experiencing serious drought and/or desertification, particularly in africa. a/ac. 241/27, paris, france. 1994.

- [171] UNCED. United nations conference on environment and development, rio de janeiro, brazil, 3-14 june 1992.
- [172] United Nations Development Programme. Undp annual report 2014, 2014.
- [173] S. van der Esch, A. Sewell, M. Bakkenes, E. Berkhout, J. Doelman, E. Stehfest, C. Langhans, L. Fleskens, A. Bouwman, and B. Ten Brink. The global potential for land restoration: Scenarios for the global land outlook 2. *PBL Netherlands Environmental Assessment Agency: The Hague, The Netherlands*, 2022.
- [174] G. Van Lynden, H. Liniger, and G. Schwilch. The wocat map methodology, a standardized tool for mapping degradation and conservation. In *Proceedings of ISCO Conference 2002, Vol. IV*, pages 11–16, Beijing, 2002.
- [175] G. Van Lynden and L. Oldeman. *The Assessment of the Status of Human-induced Soil Degradation in South and South East Asia*. UNEP, FAO and ISRIC, Wageningen, 1997.
- [176] N. Varghese and N. P. Singh. Linkages between land use changes, desertification and human development in the thar desert region of india. *Land Use Policy*, 51:18–25, feb 2016.
- [177] M. M. Verstraete. Defining desertification: a review. *Climatic Change*, 9(1):5–18, 1986.
- [178] C. M. Viana, I. Girão, and J. Rocha. Long-term satellite image time-series for land use/land cover change detection using refined open source data in a rural region. *Remote Sensing*, 11(9):1104, may 2019.
- [179] S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7):1696–1718, 2010.
- [180] J. S. Walker and T. Blaschke. Objectbased landcover classification for the phoenix metropolitan area: optimization vs. transportability. *Int. J. Remote Sens.*, 29(7):2021–2040, Apr. 2008.

- [181] J. Wang. Spatio-temporal pattern of land degradation from 1990 to 2015 in mongolia. *Environmental development*, 34:100497, 2020.
- [182] W. Wei. Spatiotemporal changes of land desertification sensitivity in north-west china from 2000 to 2017. *Journal of Geographical Sciences*, 31:46–68, 2021.
- [183] E. A. Wentz, D. Nelson, A. Rahman, W. L. Stefanov, and S. S. Roy. Expert system classification of urban land use/cover for delhi, india. *International Journal of Remote Sensing*, 29(15):4405–4427, jul 2008.
- [184] S. Wijitkosum. Factor influencing land degradation sensitivity and desertification in a drought prone watershed in thailand. *International Soil and Water Conservation Research*, 9:217–228, 2021.
- [185] S. Wittke, X. Yu, M. Karjalainen, J. Hyyppä, and E. Puttonen. Comparison of two-dimensional multitemporal sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *International Journal of Applied Earth Observation and Geoinformation*, 76:167–178, 2019.
- [186] S. Wood, K. Sebastian, and S. J. Scherr. *Pilot analysis of global ecosystems: Agro ecosystems*. World Resources Institute, 2000.
- [187] Y. Xu, S. E. Smith, S. Grunwald, A. Abd-Elrahman, and S. P. Wani. Effects of image pansharpening on soil total nitrogen prediction models in south india. *Geoderma*, 320:52–66, June 2018.
- [188] M. Yang, D. Xu, S. Chen, H. Li, and Z. Shi. Evaluation of machine learning approaches to predict soil organic matter and ph using vis-nir spectra. *Sensors*, 19(2):263, 2019.
- [189] S. Yedla and S. Peddi. Agriculture and environment. In S. Ray, editor, *Oxford Handbook of Agriculture*, pages 57–104. Oxford University Press, 2007.
- [190] X. Yi, R. Zhang, H. Li, and Y. Chen. An mff-slic hybrid superpixel segmentation method with multi-source rs data for rock surface extraction. *Applied Sciences*, 9(5):906, 2019.

- [191] S. Yousefi, H. R. Pourghasemi, M. Avand, S. Janizadeh, S. Tavangar, and M. Santosh. Assessment of land degradation using machine-learning techniques: A case of declining rangelands. *Land Degradation & Development*, 32(3):1452–1466, 2021.
- [192] M. Zeraatpisheh, A. Jafari, M. B. Bodaghabadi, S. Ayoubi, R. Taghizadeh-Mehrjardi, N. Toomanian, R. Kerry, and M. Xu. Conventional and digital soil mapping in iran: Past, present, and future. *CATENA*, 188:104424, may 2020.
- [193] Y. Zhang, Y. Tian, Y. Li, D. Wang, J. Tao, Y. Yang, J. Lin, Q. Zhang, and L. Wu. Machine learning algorithm for estimating karst rocky desertification in a peak-cluster depression basin in southwest guangxi, china. *Scientific Reports*, 12(1):19121, 2022.
- [194] Y. Zhang, J. Zhang, and X. Zhao. Land cover classification using object-based region growing and random forest algorithm with high-resolution satellite imagery. *Remote Sensing*, 13(3):443, 2021.
- [195] Z. Zhang, J. Chen, and Z. Liu. Slic segmentation method for full-polarised remote-sensing image. *Journal of Engineering*, 2019(20):6404–6407, 2019.
- [196] T. Zhou, Y. Geng, J. Chen, J. Pan, D. Haase, and A. Lausch. High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, sentinel-1 and sentinel-2 data based on machine learning algorithms. *Science of The Total Environment*, 729:138244, aug 2020.
- [197] T. Zhou, Y. Geng, C. Ji, X. Xu, H. Wang, J. Pan, J. Bumberger, D. Haase, and A. Lausch. Prediction of soil organic carbon and the c:n ratio on a national scale using machine learning and satellite data: A comparison between sentinel-2, sentinel-3 and landsat-8 images. *Science of The Total Environment*, 755:142661, feb 2021.
- [198] P. Zlateva, G. Kirov, and K. Stoyanov. Fuzzy logic application for eco-tourism potential assessment of villages. *Automatics & Informatics*, 4:20–23, 2005.

List of Publications

- **Journal**

1. **Dave Viral**, Pandya M, and Ranendu Ghosh. "Identification of desertification hot spot using aridity index."

Published in *Ann. Arid Zone* 58, no. 12 (2019): 39-44.

Citation : 13

- **Atlas**

1. Desertification and Land Degradation Atlas of India (Assessment and analysis of changes over 15 years based on remote sensing). Space Applications Centre, ISRO. Ahmedabad, India, 2021. ISBN: 978-93-82760-39-9

- **Full length conference papers**

1. **Dave Viral A.**, and Koyel Sur. "Fuzzy integrated desertification vulnerability model." in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018 ISPRS TC V Mid-term Symposium Geospatial Technology Pixel to People* Volume XLII-5, 2023 November 2018, Dehradun, India.

Citation : 2

2. **Dave, Viral A.**, M Pandya, and Ranendu Ghosh. "An Assessment of the Desertification Vulnerability based on MEDALUS model." In *2019 International Conference on Intelligent Computing and Remote Sensing (ICI-CRS)*, pp. 1-6. IEEE, 2019.

3. Pandya Megha, **Viral Dave**, and Ranendu Ghosh. "Artificial neural network (ANN) based soil electrical conductivity (SEC) prediction." in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 581-586. IEEE, 2020.

Citation : 2

4. Saha, Arnav, Srikumar Sastry, **Viral A. Dave**, and Ranendu Ghosh. "Evaluation of Tree Species Classification Methods using Multi-Temporal Satellite Images." In *2020 IEEE Latin American GRSS ISPRS Remote Sensing Conference (LAGIRS)*, pp. 40-43. IEEE, 2020.

Citation : 2

- **Abstract**

1. **Viral A. Dave**, Abhijeet Ghodgaonkar, and Ranendu Ghosh. "Desertification status classification using Gabor filters" in *National Symposium on Innovations in Geospatial Technology for Sustainable Development with special emphasis on NER*, Shillong, November 20-22, 2019.

- **Manuscript ready to be published**

1. Viral A. Dave and Ranendu Ghosh, "Predictive soil mapping using machine learning"
2. Viral A. Dave and Ranendu Ghosh, "Desertification Vulnerability using different approaches"
3. Viral A. Dave and Ranendu Ghosh, "Hierarchical pattern recognition of Desertification process"

CHAPTER A

Desertification Status classification using Gabor filters

A.1 Abstract

Desertification is one of the most main environmental problems in the world. Pattern recognition is one of the crucial step for monitoring of desertification process. Texture of the image is a primary characteristics for retrieving visually similar patterns in remote sensing images. Remote senmagesing applications are data rich as they rely increasingly on high dimensional imagery. Gabor filter have been found to be particularly appropriate for texture representation and discrimination. In the present studies, desertification status classification was carried out using Gabor filtering. Multi-temporal IRS AWiFs dataset were used for classification input data. Gabor filter were applied on input dataset and new features were generated for classification. Support vector classification algorithm is used for classifying desertification status using both original and filtered dataset and compared. Experimental results shows considerate improvement in accuracy for Gabor filtered dataset as compared to original dataset.

A.2 Introduction

Desertification and Land Degradation risk is one of the major environmental and socio-economic which constantly affects the global environment. Understanding of the desertification status help us The Gabor filter is a linear filter used in nu-

Table A.1: Dataset Specification

| Dataset | Resolution | Date |
|-----------|------------|---------------|
| | | 7/Oct/2010 |
| IRS AWiFS | 56m | 28/Feb/2011 |
| | | 5/June/2011 |
| DSM | 56m | 2011-13 cycle |

merous image processing application for edge detection, texture analysis, feature extraction, etc. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. Simple cells in the visual cortex of mammalian brains are modeled by Gabor functions. Thus, image analysis with Gabor filters is thought to be similar to perception in the human visual system. Gabor filters are special classes of band-pass filters, i.e., they allow a certain band of frequencies and reject the others. A Gabor function is viewed as a sinusoidal plane modulated by a Gaussian envelope. Gaussian provides weight and sinusoid provides direction.

A.3 Methodology

The study area used is the part of Surendranagar district of Gujarat state. It lies between the parallels of latitude 22° 8' and 23° 38' and the meridians of longitude 71° 00' and 72° 02'. It is bounded on the north by the little Rann of Kutch. IRS AWiFS Datasets were collected for all three seasons and its specifications are given in table 1. Desertification status map (DSM) of Surendranagar district prepared under Desertification status mapping of India 2nd cycle was used as a classification output map for training [136].

The areas of desertification are divided into patches or group of pixels which are segregated as training and testing samples. Gabor filter as shown in equation (1) was used on input data for spatial feature extraction. The Gabor filter is a linear filter used in numerous image processing application for edge detection, texture analysis, feature extraction, etc. Gabor filters have been found to be

particularly appropriate for texture representation and discrimination. Extracted features from Gabor filter then used for SVM supervised classification.

$$G(x, y) = e^{\left(\frac{-x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \cos\left(\frac{2\pi x'}{\lambda}\right) \quad (\text{A.1})$$

$$x' = x \cos \theta + y \sin \theta \quad (\text{A.2})$$

$$y' = -x \sin \theta + y \cos \theta \quad (\text{A.3})$$

where σ is the variance of the Gaussian function, λ is the wavelength of the sinusoidal function, θ is the orientation of the normal to the parallel stripes of the Gabor function and γ is the spatial aspect ratio specifies the ellipticity of the support of the Gabor function. For, $\gamma = 1$ the support is circular. For $\gamma < 1$ the support is elongated in the orientation of the parallel stripes of the function.

A.4 Result & Discussion

Figure A.1 demonstrates the classification accuracy achieved using Support Vector Machines (SVM). The results indicate that the SVM approach alone yielded an accuracy of 87% with a kappa value of 0.47. However, when Gabor filtering was applied prior to SVM classification, the accuracy improved to 91% with a kappa value of 0.59. This suggests that the inclusion of Gabor filtering significantly enhances the accuracy of image classification compared to the standard approach without filtering.

A.5 Conclusion

In this abstract, an efficient image classification technique has been proposed for multispectral remote sensed satellite images with the aid of clustering and Support Vector Machines (SVM). The proposed method classification technique comprises Gabor feature extraction and classification of the resultant image using

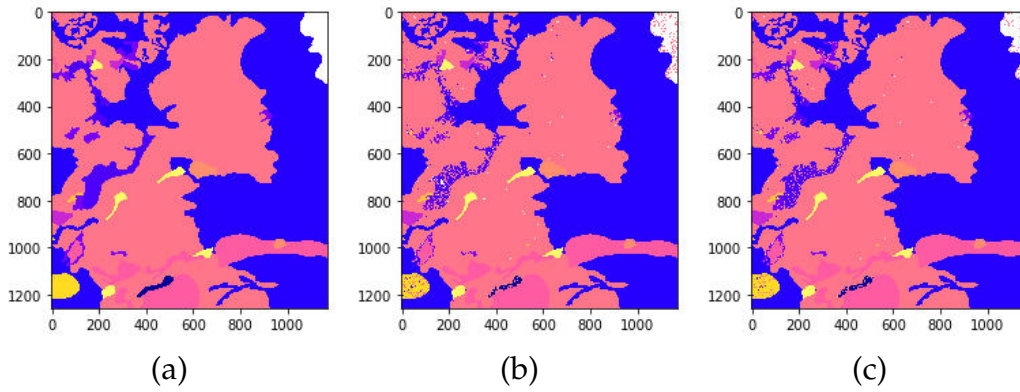


Figure A.1: (a) Original Classification Training Image, (b) Classification using SVM, (c) Classification using Gabor + SVM

SVM. The Gabor filter improves the classification performance as it is able to get the texture features from the data accurately.

CHAPTER B

Action plan generation for combating land degradation at micro-watershed level

Desertification is land degradation in arid, semi-arid, and dry sub-humid areas resulting from various factors, including climatic variations and human activities, said by UNCCD. 30% of India's total geographical area being affected by land degradation and to prevent it combating plan for land degradation is necessary. From multi-date satellite data and ancillary information with other thematic maps, e.g. Land use/ land cover, vegetation cover, land capability and slope are prepared. Realizing the importance of adopting an integrated approach, and recognizing the mutual interdependence of natural resources, thematic information is integrated using python programming platform. In order to integrate various themes, firstly, land use/ land cover layer is integrated with vegetation cover layer. The resultant of these two themes is unionized with slope and finally with land capability. The resultant coverage has the basic information of all the four themes- land use/ land cover, vegetation cover, land capability and slope and referred in resource data base. Various map units known as composite land development units (CLDU) are created in this composite layer. Overall methodology for combating desertification is given in figure 2 below. Based on the CLDU characteristics various measures are suggested for conservation and protection of natural resources. The decision rules, given in Table 1 are used to suggest the appropriate action plan for each of the CLDU.

B.1 Methodology

To develop CLDU first area has been identified which comes under the semi-arid to desertified zone. From that region micro-watershed has been chosen to develop the action plan of combating desertification. To generate CLDU, along with satellite images, ancillary data also need to stack with different thematic maps. After that, unique combination of the all stacked data has been identified. And from the thousands of the unique combinations, it becomes so difficult to decide the action plan. So, less than 5Ha area has been unionized to the nearby region and try to classify the region with heterogeneity. So, the different action plan can be generated against each unique combination or the process that lead the region to desertification. With the help of expert advice and knowledge some of the action plan has been generated to increase overall vegetation cover in terms of increasing the NDVI of the region and cutting off or reduce the run-off of the region.

The lookup table B.1 developed with expert advice serves as a crucial tool for proactive decision-making and planning in the face of various challenges and informs policy recommendations for tackling desertification and land degradation. By deriving conclusions from the data, it guides effective policy implications, aiding strategies to counter these challenges through informed decision-making and sustainable land management practices. It empowers authorities to make well-informed choices that can mitigate risks, enhance resilience, and contribute to the overall well-being of the affected region and its inhabitants.

Table B.1: Strategic Framework: Micro-Watershed Level Action Plan for Combating Land Degradation

| CLDU*_Code | DSM-CODE | LULC* | Slope (%) | LCAP* | GWP* | Combat Action/Prescription |
|------------|----------|--------------------|------------------------------|-----------|----------|---|
| 1 | Fv1 | Open forest | 15-25, >25 | 2, 6, 7 | Moderate | Afforestation |
| | Fv2 | Open forest | 15-25, >25 | 2, 6, 7 | | |
| | Fv3 | Degraded forest | 0-5, 5-10, 10-15, 15-25, >25 | 1,2,6,7 | | |
| 2 | Fw1 | Open forest | 5-10, 15-25 | 6 | Moderate | Afforestation and soil conservation measures |
| | Fw2 | Degraded forest | >25 | | | |
| 3 | Fw4 | Open forest | 15-25, >25 | 7 | Moderate | Afforestation/grasslands and soil conservation measures |
| | Fw4 | Degraded forest | 0-5, 10-15,15-25, >25 | 6, 7 | | |
| | Fw3 | Degraded forest | 0-5, 5-10, 10-15, 15-25, >25 | 6, 7 | | |
| 4 | Iw4 | Permanent fallow | 15-25, >25 | 6, 7 | Moderate | Contour bunding and orchard development |
| | Iw2 | Permanent fallow | 15-25 | 7 | | |
| 5 | NAD | Permanent fallow | 5-10, 10-15, 15-25, >25 | 2, 6, 7 | Moderate | Contour bunding/ Grassland land development |
| 6 | Sv3 | Land without scrub | 10-15,15-25,>25 | 6,7 | Moderate | Contour bunding/Social Forestry development |
| | Sv4 | Land without scrub | 10-15,15-25, >25 | 6,7 | | |
| | Sw2 | Land without scrub | 15-25 | 6 | | |
| | Sw3 | Land without scrub | 10-15,15-25, >25 | 6, 7 | | |
| 7 | Sv1 | Land with scrub | 5-10, 10-15 | 2, 7 | Moderate | Fencing to control overgrazing |
| 8 | Fv1 | Open forest | 0-5, 5-10, 10-15 | 2, 6, 7 | Moderate | Forest Plantation |
| | Fv2 | Open forest | 0-5, 5-10, 10-15 | 2, 6, 7 | | |
| 9 | NAD | Permanent fallow | 5-10, 10-15 | 6, 7 | Moderate | Not suitable for agriculture/Pasture-grassland land development |
| 10 | NAD | Single crop | 0-5, 5-10 | 6, 7 | Moderate | Not suitable for agriculture/Soil conservation measures |
| | | Double crop | 0-5, 5-10 | 6, 7 | | |
| | | Tripple crop | 0-5, 5-10 | 6, 7 | | |
| 11 | NAD | Single crop | 10-15, 15-25, >25 | 6, 7 | Moderate | Over utilized/Soil conservation measures |
| | | Double crop | 10-15, 15-25, >25 | 6, 7 | | |
| | | Tripple crop | 10-15, 15-25, >25 | 6, 7 | | |
| | Sv1 | Land with scrub | 15-25, >25 | 6, 7 | | |
| | Sv2 | Land with scrub | 15-25, >25 | 6 | | |
| 13 | NAD | Plantations | 0-5, 5-10, 10-15, 15-25,>25 | 2, 6, 7 | Moderate | Forest Protection/Conservation |
| 14 | NAD | Dense forest | | | Moderate | Sandy area |
| 15 | NAD | Sandy area | | | Moderate | Settlements |
| 16 | S | Settlements | | | Moderate | Social Forestry development |
| 17 | Sw4 | Land without scrub | 05-Oct | 7 | Moderate | Soil conservation measures |
| | Iw3 | Single crop | 15-25 | 7 | | |
| | Iw4 | | >25 | 6 | | |
| | Iw1 | | >25 | 6 | | |
| | Iw2 | | >25 | 6 | | |
| | NAD | Triple crop | Oct-15 | 2 | | |
| | Iw2 | | >25 | 6 | | |
| | NAD | | >25 | 6 | | |
| NAD | >25 | | 6 | | | |
| 18 | Fv3 | Degraded forest | 0-5, 5-10, 10-15 | 2, 7 | Moderate | Strictly reclamation and conservation /protection |
| | Fv4 | Degraded forest | 5-10, 10-15, 15-25, >25 | 1,2, 6, 7 | | |
| 19 | NAD | Single crop | 0-5, 5-10 | 2 | Moderate | Sustainable maintenance |
| | | Double crop | 0-5, 5-10 | 2 | | |
| | | Triple crop | 0-5, 5-10 | 2 | | |
| 20 | NAD | Single crop | 10-15, 15-25 | 2 | Moderate | Terrace farming |
| | | Double crop | 10-15, 15-25, >25 | 1, 2 | | |
| | | Triple crop | 10-15, 15-25, >25 | 2 | | |
| 21 | NAD | Permanent fallow | 0-5 | 2 | Moderate | Underutilized/ can be used for agriculture |
| 22 | NAD | Permanent fallow | 10-15, 15-25 | 2 | Moderate | Underutilized/ can be used for Terrace farming in sustainable way |
| 23 | W | | | | | Water bodies |

*CLDU = Composite land development units, *LULC = Land use-Land Cover, *LCAP = Land Capability, *DSM = Desertification status map, *GWP= Ground water prospects.