EVALUATION OF THE EXTENT OF FLOODS IN 2022 IN DIFFERENT LAND COVER

AND LAND USE CLASSES IN SINDH PROVINCE, PAKISTAN USING

SENTINEL-1 AND SENTINEL-2 IMAGERIES

By

Alina Shrestha

Bachelor of Engineering in Civil Engineering
Tribhuvan University
2015

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Engineering - Civil and Environmental Engineering

Department of Civil and Environmental Engineering and Construction
Howard R. Hughes College of Engineering
The Graduate College

University of Nevada, Las Vegas
May 2024

Thesis Approval

The Graduate College
The University of Nevada, Las Vegas

April 8, 2024

This thesis prepared by

Alina Shrestha

entitled

Evaluation of the Extent of Floods in 2022 in Different Land Cover and Land Use Classes in Sindh Province, Pakistan Using Sentinel-1 and Sentinel-2 Imageries

is approved in partial fulfillment of the requirements for the degree of

Master of Science in Engineering - Civil and Environmental Engineering
Department of Civil and Environmental Engineering and Construction

Sajjad Ahmad, Ph.D.
Examination Committee Co-Chair

Haroon Stephen, Ph.D.
Examination Committee Co-Chair

Marie-Odile Fortier, Ph.D.
Examination Committee Member

Ashok Singh, Ph.D.
Graduate College Faculty Representative

Alyssa Crittenden, Ph.D.
Vice Provost for Graduate Education &
Dean of the Graduate College

Abstract

Floods are one of the most frequent and most devastating natural disasters, which cause widespread destruction and pose significant risks to human life, infrastructure, and the environment. Advancement in remote sensing technologies and methodologies have demonstrated their efficacy in disaster-related applications, such as the detection, monitoring, and analysis of floods. This study explores the utilization of Synthetic Aperture Radar (SAR) and optical imagery for flood extent mapping and studying the extent of flood over various land cover and land use classes in Pakistan's Sindh province, utilizing the cloud computing power of Google Earth Engine. The change detection method identified extensive flooding in Sindh province, covering an area of 25,229 km^2 in August and 19,181 km^2 in the first 19 days of September 2022. The Land Use/Land Cover dataset was developed for the pre-flood period. The study highlighted the effectiveness of Random Forest classification in distinguishing Land Use/Land Cover (LULC) types more accurately than K-means clustering. Additionally, the analysis provided insights into the spatial distribution of flood extent and vulnerability of land use/land cover classes such as urban areas, agricultural areas, and sparse natural vegetation as significant areas remained inundated in the province.

Acknowledgments

I sincerely appreciate the guidance and support provided by my advisors, Dr. Sajjad Ahmad and Dr. Haroon Sahotra, throughout my graduate studies. Their expert knowledge and practical advice were essential to the successful completion of my thesis. They not only helped me formulate essential questions that defined the objectives of my research, but also guided me through the necessary processes, motivated me to persevere at each step, and thoroughly analyze the outcomes. Furthermore, they equipped me with the skills needed to extract meaningful insights from my work and to communicate these findings clearly and concisely.

I am also grateful to Dr. Marie-Odile Fortier and Dr. Ashok Singh for their roles on my advisory committee. Their invaluable constructive feedback profoundly enriched my research and played a crucial role in its refinement.

I also want to extend my thanks to my husband, Ayush Shrestha, and my friend, Pooja Basnet, for their unwavering support during the strenuous period of completing my thesis. Their encouragement and companionship were essential in overcoming the challenges I faced. I also thank my roommate and family for their constant support and love, which have been fundamental to my success.

I extend my thanks to the friends in my research group, who faced similar challenges and grew alongside me throughout this journey. The camaraderie and shared experiences we had together were invaluable.

Table of Contents

vi

List of Tables

List of Figures

List of Abbreviations

SAR - Synthetic Aperture Radar

GIS - Geographic Information Systems

UAVs – Unmanned Aerial Vehicles

LULC - Land Use/Land Cover

LIDAR – Light Detection and Ranging

DII - Difference Image Index

DFI - Desert Flood Index

MNDWI - Modified Normalized Difference Water Index

NDVI - Normalized Difference Vegetation Index

GEE - Google Earth Engine

GRD - Ground Range Detected

VV – Vertical transmit -Vertical receive Polarization

HH – Horizontal transmit - Horizontal receive Polarization

VH – Vertical transmit - Horizontal receive Polarization

HV - Horizontal transmit - Vertical receive Polarization

DEM - Digital Elevation Model

RI - Ratio Image

NDFI - Normalized Difference Flood Index

IW - Interferometric Wide swath

dB - Decibels

ESA - European Space Agency

EC - European Commission

FAO - Food and Agriculture Organization

JRC - Joint Research Centre

NASA - National Aeronautics and Space Administration

NASADEM - NASA Digital Elevation Model

SRTM - Shuttle Radar Topography Mission

PCA - Principal Component Analysis

WCSS - Within-Cluster Sum of Squares

UA - User Accuracy

PA - Producer Accuracy

Chapter 1 - Introduction

1.1 Research Background

Climate change has significantly increased the severity and frequency of natural disasters worldwide, intensifying the risks associated with extreme weather events such as floods (Khokhar et al., 2021; Otto et al., 2023). Key climatic factors influencing flood risks such as atmospheric temperature, precipitation, humidity, and air circulation, have been altered, leading to more frequent and severe flooding events (Meresa et al., 2022). This trend is expected to continue, with a particular increase in high-magnitude floods. Notably, the impact of these changes varies by region (Hirabayashi et al., 2021; Meresa et al., 2022). Countries like Pakistan are particularly vulnerable, where heightened flood risks exacerbate existing challenges and compound the threats posed by global warming and climate-related disasters.

Despite contributing less than 1% to global greenhouse gas emissions, Pakistan is disproportionately affected by climate change and extreme weather events (Shehzad, 2023; Waqas, 2022; Waseem & Rana, 2023). Pakistan is ranked among the top eight countries most vulnerable to climate change impacts according to the Global Climate Risk Index by German Watch (Ali, 2013). Pakistan has been a victim of severe floods throughout history and faces an extreme flood almost every three years, recording up to 21 extreme floods between 1950 and 2019. These incidents have intensified recently, with more frequent droughts, extreme temperatures, heavy precipitation, and accelerated glacial melts. The nation's susceptibility to climate-induced events is compounded by its diverse topography and geographical characteristics, and the damaging impacts of such events are exacerbated by human interventions such as urbanization, and population growth (Kazi, 2014). The region's rising temperatures have

increased glacial melts during the summer, leading to excessive runoff into lowland areas and heightening the likelihood of flooding events. Heavy precipitation during the monsoon exacerbates the floods. The recent floods in 2010 and 2022 are stark examples of such extreme floods. The Flood of 2010 was considered the most devastating flood before 2022. The frequency and magnitude of the floods have been on the rise, underscoring the growing impact of extreme weather conditions in the region (Gupta et al., 2022; Yuan et al., 2023).

Pakistan exhibits varied topography and distinct climatic characteristics, encompassing the mountainous northern region, hilly and sloping areas, and low-lying plains. This diversity includes prominent features such as the Himalayan and Karakoram Mountain ranges, the Balochistan plateau, and the Indus River plains. River Indus is the major river in the region, the majority of which lies within Pakistan (Kazi, 2014). Such varied landscapes further influence the distribution and impact of floods across the country. The process of orographic lifting in mountainous areas plays a crucial role in generating the extreme precipitation characteristic of monsoon seasons (Houze et al., 2011; Yuan et al., 2023). Floods in Pakistan stem from various factors beyond heavy precipitation. Pakistan experiences a monsoon season from July to September, characterized by substantial rainfall. Melting glaciers in the northern regions, including the Himalayas and Karakoram Range, contribute to the swelling of the Indus River and its tributaries as well as major water bodies within the region during the pre-flood summer months. These natural causes of floods are compounded by human interventions and activities such as rapid urbanization, the construction of dams and reservoirs, and poor water management practices have worsened the severity (Shehzad, 2023; Waqas, 2022; Waseem & Rana, 2023).

Between 1950 and 2010, the extreme floods impacted around 446,000 km^2 of Pakistan, killed 8887 people and caused an economic loss of about $19.0 billion (Ali, 2013). The Super Flood of 2010 occurred during July and August 2010. The 2010 Floods directly caused the death of over 1900, affecting the lives of 20 billion people (Manzoor et al., 2022). An economic loss of $16.0 billion was reported by the Government of Pakistan. The floods of 2022 caused an economic loss of more than $10.0 billion, with over $2.3 billion loss in crops. The provinces of Sindh, Balochistan, and Punjab were hard hit by the floods (Shehzad, 2023).

The 2022 floods were triggered by a variety of causes, including prolonged heavy rains, the melting of glaciers, and strong low-pressure areas formed due to heat waves earlier in the summer. In the summer months of 2022, Pakistan experienced record-breaking temperatures, with readings surpassing 40°C in several regions. These intense heat conditions and heat waves during April and May contributed significantly to the increase in glacier melting (Nanditha et al., 2023; Shehzad, 2023). By mid-June, several intense episodes of monsoon rains commenced, exacerbating the higher river flows from the melting glaciers. This resulted in the Indus River basin enduring severe floods, which majorly affected the southern provinces of Balochistan, Sindh, and Khyber Pakhtunkhwa, from continuous heavy rainfall spanning May to August 2022. Additionally, in August 2022, there was a notable rise in the transport of water vapor, which led to elevated precipitation and subsequent widespread floods in the southern parts of Pakistan (Gupta et al., 2022; Nanditha et al., 2023).

Sindh province of Pakistan experienced a severe humanitarian crisis as extreme precipitation inundated large areas, with several regions covering over 117,850 km² enduring 15-day rainfall accumulations exceeding 400 mm. Particularly hard-hit were the communities along the Indus River, where many homes were destroyed, and families were displaced to overcrowded relief

camps lacking basic amenities. The floods caused approximately $2.3 billion in damage to agricultural lands across Sindh, Balochistan, and Punjab, destroying 1.9 million tons of rice, 10.5 million tons of sugarcane, and 3.1 million bales of cotton. These losses translated to significant economic setbacks in key commodities (Nanditha et al., 2023; Shehzad, 2023). Additionally, the floods decimated 42,273 heads of livestock worth $13 million and inflicted major losses on essential crops such as tomatoes, onions, and chilies.

Understanding the scope of flooding is crucial for supporting emergency responses and facilitating mitigation plans. Remote sensing methodologies, particularly through Earth observation satellites, offer critical data for gauging flood extents and managing, monitoring, and responding to flood events. This information is instrumental across various flood management domains, including rainfall-runoff analysis, flood prediction, and flood pattern analysis (Schumann, 2015). When integrated with Geographic Information Systems (GIS), remote sensing enhances the continuous collection of crucial flood-related data such as discharge rates, flooding depths, and inundation maps, significantly outperforming traditional methods. This continuous assimilation of information is essential not only for comprehending the extent and damage of floods but also for analyzing risks, assessing community vulnerabilities, and formulating effective management strategies to mitigate the impacts of future floods (Cohen et al., 2019; Nasr et al., 2022).

Flood extent mapping is a crucial tool in disaster management and urban planning, aiding in the identification of flood-prone areas and those affected by floods. This mapping is essential for enhancing disaster risk reduction strategies, improving preparedness, and increasing community resilience against floods (Ullah & Zhang, 2020). Utilizing advanced remote sensing technologies and high-resolution satellite imagery, accurate flood mapping provides critical insights for

4

sustainable development, emergency response, and effective flood management practices (Belabid et al., 2019). These technologies ensure the rapid dissemination of time-sensitive information, vital for informed decision-making and strategy formulation in disaster scenarios. Furthermore, historical flood data accessed through various platforms enhances the understanding of flood patterns, crucial for developing effective planning and adaptation strategies (Munawar et al., 2022). Overall, the continuous improvement in remote sensing data acquisition and mapping techniques significantly bolsters resilience and readiness amidst escalating flood threats, thereby supporting more resilient communities and infrastructure.

Damage assessments, alongside flood extent mapping, are vital for quantifying the economic impacts of floods, including agricultural losses, infrastructural damages, and changes in land use (Nie et al., 2017; Jin et al., 2013). These evaluations are crucial for estimating the extent of damage caused by disasters and play a significant role in the economic analysis of affected regions. Beyond immediate impacts, historic damage assessments are crucial in understanding climate change effects, aiding long-term planning and risk mitigation in flood-prone areas. Collaborative remote sensing methodologies significantly enhance these efforts by supporting the formulation of effective emergency responses and strategic long-term planning (Inoue et al., 2008; Wang et al., 2016).

In addition to flood extent mapping, damage assessments are crucial for quantifying the economic impacts of floods, including agricultural losses, infrastructural damages, and changes in land use (Nie et al., 2017). Damage assessments are crucial in the quantification of damage resulting from disasters such as floods, which impact the economies of the affected regions. The evaluation extends beyond agricultural losses to encompass infrastructural damages and changes in land use and cover categories (Jin et al., 2013).Historic damage assessments help in

5

understanding the effects of climate change in the region. These assessments are vital for understanding the effects of climate change and supporting long-term planning and risk mitigation strategies in flood-prone regions. Collaborative remote sensing methodologies play a significant role in advancing emergency response efforts, aiding in the formulation of effective relief initiatives and long-term planning strategies in flood-prone areas (Inoue et al., 2008; Wang et al., 2016).

Ongoing advancements in mapping technologies and the increased availability of high-resolution, remotely sensed satellite data are pivotal for improving flood risk management and enhancing community resilience. This detailed data supports the rapid dissemination of time-sensitive information, which is instrumental in making informed decisions about disaster management strategies and response efforts. Furthermore, the integration of advanced remote sensors, historical flood data accessed through various platforms, and innovative mapping techniques not only helps identify flood-prone areas but also aids in the development of effective planning and adaptation strategies. Such enhanced access to remote sensing data and historical analyses is essential for bolstering readiness and resilience, ultimately contributing to more robust infrastructure and safer communities in the face of escalating flood threats (Munawar et al., 2022).

1.2 Research Motivation

Flood incidents have always been significant globally due to their devastating consequences. Asia is home to some rivers most prone to flooding, accounting for nearly 38% of the flooding events. Regions in South Asia such as the Indo-Gangetic Plain, the Indus Delta, and the Ganges–Brahmaputra–Meghna basin are the key economic zones, whilst also being prone to flooding events and affecting the livelihood of more than one billion people. With dramatic changes in

weather patterns, climatic conditions, and summer monsoon rainfall, occurrences of severe droughts and floods are more frequent. A considerable segment of the population dwells in flood-prone lands that are exposed to floods repeatedly (Nanditha et al., 2023). With the increasing frequency and severity of flooding events globally, it is imperative to prioritize flood risk assessment and management in vulnerable regions like South Asia (Matheswaran et al., 2018). In South Asia, authorities and communities are frequently compelled to adapt and adjust to changing climatic conditions induced by disasters facing economic, human, and infrastructural losses consequently. Climate change plays a significant role in exacerbating flood risks, posing environmental threats and challenges to plant growth and development (Moortgat et al., 2022). The need for interdisciplinary approaches to understanding and addressing flood risks is highlighted to enhance preparedness and response strategies (Suarez & Meeroff, 2021).

Climate change and global warming have been linked to changes in weather patterns beyond prediction and have led to an increase in the frequency and severity of floods globally (Kapilaratne & Kaneta, 2020). This necessitates the development of accurate and real-time flood mapping technologies. Accurate and timely flood extent mapping is critical for effective flood management and mitigation efforts (Porter & Demeritt, 2012). Such mapping provides a comprehensive overview and improves upon the localized insights from traditional methods which usually depend solely on point-sourced data. Recent advancements in flood mapping technologies include the use of unmanned aerial vehicles (UAVs) equipped with high-resolution cameras for efficient and cost-effective data acquisition in hazardous areas (Gebrehiwot et al., 2019). The combination of remote sensing data and digital elevation models (DEMs) derived from Light Detection and Ranging (LiDAR) technology has proven effective in flood hazard mapping and assessment (Muhadi et al., 2020; Annis et al., 2020). This broader perspective

7

facilitates authorities to make informed decisions regarding evacuation efforts, resource

allocation, and emergency response strategies and enables targeted actions to mitigate economic

and human losses. Therefore, the progress in flood mapping methods is taken as an important

sector within climate adaptation strategies, highlighting the need for continued innovation and

investment in this domain.

Remote sensing technologies, including satellite imagery and aerial sensors, have aided in

significant advancements in flood detection, monitoring, and analysis capabilities. Utilizing

satellite data enables researchers and disaster management authorities to monitor flood dynamics

extensively, pinpointing inundated areas, observing water level changes, and forecasting flood

trends (Frappart et al., 2005). These techniques also facilitate damage assessment by allowing for

the comparison of pre- and post-flood images which also aids in evaluating the severity of

damage to infrastructure, agricultural land, and natural ecosystems. They are essential for

effective recovery planning and allocation of resources for rehabilitation efforts. Remote sensing

proves invaluable in delivering timely, precise, and exhaustive data regarding flood impacts.

Moreover, combining this information with GIS and other data sources, such as the Digital

Elevation Model (DEM), allows for a detailed analysis of floods' effects across various sectors.

Optical sensors, with their extensive spatial coverage and high temporal frequency ranging from

daily to bi-weekly, have been pivotal in flood analysis, operating across the visible to thermal

wavelengths as active sensors. Their effectiveness in water detection via the infrared spectrum is

well-documented, with notable examples including MODIS, AVHRR, ASTER, Landsat, and

Sentinel-2 (Zhang et al., 2014). Techniques leveraging the multispectral capabilities of these

sensors have been instrumental in identifying water bodies (Hamidi et al., 2023; Ticehurst et al.,

2009). However, their utility is constrained under conditions of cloud cover and dense foliage,

8

presenting challenges in flood monitoring (Ticehurst et al., 2009; Ward et al., 2014). Conversely, microwave remote sensing, particularly Synthetic Aperture Radar (SAR), operates at longer wavelengths and lower frequencies, enabling it to penetrate obstacles such as clouds and vegetation, thus ensuring all-weather capabilities for flood mapping (Tazmul Islam & Meng, 2022; Ticehurst et al., 2009). The introduction of Sentinel-1 has revolutionized SAR data usage, offering weekly access to SAR-C data, and facilitating high-resolution, near-real-time flood detection and monitoring under any weather condition (Hamidi et al., 2023). SAR's ability to detect changes in backscatter signals is crucial for identifying water bodies, using methods such as supervised classification, thresholding techniques, and multitemporal analysis (Amitrano et al., 2018; Long et al., 2014). Despite the valuable spectral information provided by optical sensors for water detection, SAR's all-weather operational capability offers an indispensable perspective for accurate and reliable flood mapping, underscoring the complementary roles of SAR and optical imagery in comprehensive flood extent mapping.

The impacts of floods in the Sindh region have been highlighted in various studies. Challenges are faced in South Asia due to floods, which emphasize the need for climate mitigation strategies (Nanditha et al., 2023). The devastating consequences of floods in Pakistan in 2022 have been reported by various sources (FAO, 2022; Manzoor & Adesola, 2022; Qamer et al., 2023), suggesting significant disruption to the agriculture sector, alteration in cropping patterns and losses in agriculture. Additionally, the floods destroyed irrigation structures and other farm facilities, further compounding the agricultural challenges (Ali et al., 2017; Manzoor et al., 2022). Changing weather patterns due to climate change have also reduced crop yields and livestock productivity, adversely affecting the dairy and poultry sectors, vital to Pakistan's

economy (Ishaque et al., 2022). Given these impacts, there is a clear need for targeted interventions to support farmers and enhance sector performance (Rajpar et al., 2019).

The Modified Normalized Difference Water Index (MNDWI) has been established as an effective tool for the detection of open water bodies, gaining recognition for its utility in numerous studies (Baig et al., 2013; Hamidi et al., 2023; Hidayah et al., 2022; Markert et al., 2018; Sajjad et al., 2020; Sajjad et al., 2023; Ward et al., 2014). This index is particularly valued for its ability to minimize errors associated with the presence of built-up areas, vegetation, and soil, making it highly suitable for surface water detection in flood-related research. Despite its advantages, MNDWI differentiates water at the pixel scale, rather than by the area's proportion of water, limiting its precision in accurately extracting water information. Furthermore, its sensitivity towards vegetation over soil compromises water mapping accuracy (Baig et al., 2013).

In the context of Sindh Province, Pakistan, which experiences a climate ranging from arid to semi-arid, with hot summers, mild winters, low rainfall, and high monsoon humidity, the limitations of MNDWI become apparent. The index's inability to effectively distinguish between water bodies and hill shadows, or to accurately identify built-up areas in lowlands, poses challenges. Moreover, MNDWI is vulnerable to signal noise from land cover in built-up and disturbed areas that may undermine the reliability of water body extraction due to its sensitivity to green and shortwave infrared parts of the spectrum (Sajjad et al., 2020; Sajjad et al., 2023). Additionally, the presence of vegetation impacts the green band reflectance, potentially leading to erroneous water body detection in areas with sparse vegetation.

This study highlights a significant research gap in employing Sentinel-1 SAR data for flood detection and damage assessment within Pakistan's Sindh region. Despite the existing literature on flood mapping using various remote sensing techniques and data sources (Notti et al., 2018; Qamer et al., 2023; Rahman & Di, 2020; Sajjad et al.,2020), there is a notable absence of a comprehensive change detection method for detecting floods in Sindh using Sentinel-1 SAR imagery. Moreover, the comparison of flood extent mapping obtained from Sentinel-1 SAR data with other data sources has not been adequately explored, particularly in the context of the Sindh region, where direct comparisons remain unaddressed in the current literature. This lack of standardized methodologies for Sentinel-1 SAR-based flood extent mapping studies leads to issues with the reliability and comparability of flood mapping results across different studies and regions. Although there is a common understanding that the fusion of multi-temporal and multi-spatial datasets can provide an accurate representation of flooding events, this approach remains largely unexplored, especially for Pakistan. This underscores the necessity for developing methodologies that not only address the limitations of using Sentinel-1 SAR imagery but also enhance the accuracy and comparability of flood mapping results.

Research in damage assessment using Land use/Land cover (LULC) classification images and flood extent mapping has seen advancements. Publicly available global remotely sensed land use land cover (LULC) datasets are crucial for damage assessment in flood studies. Researchers generally make use of publicly available datasets such as Dynamic World V1, ESA WorldCover 10m V100, ESA WorldCover 10m, V200, and Copernicus Global Land Cover Layers. By integrating the latest LULC data into flood damage assessments, researchers can gain a comprehensive understanding of landscape changes and flood vulnerability, facilitating more informed decision-making in disaster management and mitigation efforts. However, such

datasets are updated annually to maintain accuracy and relevancy. This necessitates the exploration of machine learning and artificial intelligence techniques to improve the accuracy of classification images and damage assessment practices. While there are studies that have utilized machine learning approaches, there is a lack of in-depth understanding of leveraging these advanced techniques.

The integration of Sentinel-1 SAR imagery for flood extent mapping and the application of land use and land cover (LULC) classification for damage assessment has significant implications for disaster management and policy making. The use of an index-based approach on Sentinel-1 SAR images enables the rapid, accurate identification of flooded areas, even under challenging conditions such as cloud cover or night-time (Anusha & Bharathi, 2020). This real-time flood mapping capability allows for timely and effective emergency response, resource allocation, and evacuation planning, thereby minimizing human and material losses. Furthermore, the combination with LULC classification data enhances the capacity for detailed damage assessment post-disaster (Benoudjit & Guida, 2019). By identifying the specific types of land use and cover affected by flooding, authorities can better understand the impact on critical infrastructure, agriculture, and residential areas. This detailed damage assessment informs recovery and rehabilitation efforts, ensuring that they are well-targeted and efficient. For policymaking, these technologies offer a solid evidence base for the development of more resilient infrastructure and land use planning strategies. By analyzing flood extent and damage patterns over time, policymakers can identify vulnerable areas and prioritize interventions to reduce flood risk and enhance community resilience. Additionally, this data-driven approach supports the implementation of adaptive policies that can respond to the changing dynamics of

flood risk under climate change scenarios, ensuring sustainable development and the protection of vulnerable populations.

1.3 Research Objectives

The overarching goal of the study is to understand the impacts associated with flood events by utilizing open-access remote sensing datasets. Optical and microwave sensors have been used in flood studies due to their larger spatial coverage and high temporal resolution at frequencies ranging from daily to every two weeks. Optical sensors operate in the visible portion of the wavelength spectrum but also include the infrared and thermal regions and fall under active sensors. However, due to their limitations in detecting through clouds and tree canopies, flood progress is difficult to monitor in many circumstances where cloud and smoke concentrations are significant. Microwave remote sensors, on the contrary, operate on longer wavelengths than optical remote sensors, the operating wavelengths ranging from 1 cm to 1m which facilitates flood inundation mapping in all-weather circumstances as microwaves can penetrate cloud, aerosol, haze, and tree canopy. The change detection method, however, must be adjusted to be applicable to the Sindh province. Utilizing remote sensors for flood detection proves valuable in damage assessments and recovery efforts. These aspects will be investigated using the Google Earth Engine. The thesis has been organized into two objectives, and each objective attempts to answer a set of research questions. The objectives of the study are:

1. Assessing the performance of three optical and microwave remote sensing data-based flooding indices to map the flooding extent in semi-arid regions.

    a. How can the change detection method be adopted for the detection of floods in Sindh with Sentinel-1 SAR images?

b. What were the spatial extent and total area affected by the flooding in the Sindh region?

c. How can the flood extent mapping derived from Sentinel-1 SAR data be compared to flood extent mapping obtained from other data sources?

2. Evaluating the extent of flood on various land cover classes using remote sensing-based flood maps and Land Use/Land Cover (LULC) dataset.

a. What is the spatial distribution of flood-affected areas across different land classes within the districts of Sindh?

b. How does the flood extent vary among different land classes (e.g., agricultural fields, urban areas, forests) in the administrative districts of Sindh during the flood event?

1.4 Thesis Organization

The research conducted in this project has been documented in the form of a manuscript. The document comprises three chapters, each serving a specific purpose. The initial chapter, titled "Introduction" lays the groundwork by presenting the problem statement and the imperative need for this study. A summary of the literature work previously done on the subject has been included in this chapter. The research objectives are laid out alongside the crucial research questions which aid in realizing the objectives of the study.

The succeeding chapter, Chapter 2 titled "Assessing the effectiveness of Sentinel-1 SAR backscatter-based Difference Image Index and Sentinel-2 reflectance-based Desert Flood Index to map the flood extent." is dedicated to the first objective and answers the first set of research

questions under the same objective. The chapter is dedicated to developing the change detection algorithm for mapping the extent of floods in Sindh province using Sentinel-1 images. The chapter further describes how the change detection method is developed and optimized with the use of Sentinel-2 images with less cloud cover. The results of fine-tuning the hyperparameters and the results of the change detection algorithm are discussed in this chapter. The results are also compared with other sources such as published journal articles, and newspaper articles.

The third chapter titled "Evaluating the extent of flood impact on various land cover classes using remote sensing-based flood maps and Land Use/Land Cover (LULC) dataset." is dedicated to the second objective of the thesis study. This chapter addresses the objective of assessing the damage caused by the floods of 2022 in the Sindh province in the overall region of Sindh Province. In this chapter, the methodology adopted to derive the Land Use/Land Cover (LULC) dataset is discussed. The Land Use/Land Cover dataset has been prepared using supervised random forest classification and unsupervised clustering method with the help of the publicly available Land Cover Dataset by the Food and Agricultural Organization (FAO). The methodology and results of the spatial coverage of flood extent over various land use classes have been discussed. Chapter 2 and Chapter 3 are dedicated to two objectives of the study which make use of publicly available datasets in Google Earth Engine and highlight the importance of high-resolution, and frequent imaging datasets in flood studies. The fourth chapter presents the summary of the research work, its contributions, limitations, and recommendations for water managers and future research.

Chapter 2 - Assessing the Performance of Three Optical and Microwave Remote Sensing Data-Based Flooding Indices to Map the Flooding Extent in Semi-Arid Regions.

Abstract

Mapping of flood extents is essential for effective disaster management and mitigation efforts as they are one of the most widespread natural disasters. Remote sensing technologies and techniques have proven to be effective and efficient in various applications related to disasters such as flood detection, monitoring, and analysis capabilities. With advancements in technologies, remotely sensed high spatial resolution and high temporal frequency are available. The effectiveness of the Synthetic Aperture Radar (SAR) datasets for near real-time flood mapping has been studied. As a microwave sensor, SAR is superior to optical sensors as it overcomes limitations such as cloud cover and adverse weather conditions. All the computations have been performed in Google Earth Engine, which is a cloud-based platform powered by Google which allows users to access multiple geospatial datasets and utilize the cloud computing capacity and storage facilities. A change detection algorithm has been developed and optimized using the Difference Image Index (DII) based on Sentinel-1 SAR data. Desert Flood Index (DFI) imagery based on Sentinel-2 data has been used to refine the threshold value of DII to produce accurate flood extent mapping for the region of Sindh province. The study identified extensive flooding that covered an area of 25,229 square kilometers in August and 19,181 square kilometers in the first 19 days of September 2022 in Sindh province.

Keywords: flood extent mapping, change detection, thresholding, Sentinel-1 flooding, Sentinel-2 flooding, Google Earth Engine

2.1 Introduction

One of the most destructive natural disasters in the world is flooding, which is defined as water overflowing into typically dry ground. As the impacts of climate change and global warming become apparent, the occurrences of extreme flood events are bound to increase. There has been a significant increase in the number of flood-related research in recent decades which highlights the importance of understanding floods (Zhang et al., 2023). Disasters such as floods very often cause considerable damage to the economy, infrastructures, and livelihoods of the people. They uproot communities and force people to relocate, frequently surpassing the capacity of the existing disaster management systems. Flood disaster mitigation initiatives are crucial for lessening the effects of floods on ecosystems and populations, highlighting the significance of taking preventative action in disaster management.

Understanding the nature and extent of increasing extreme flooding events requires the exploitation of existing data resources to maximum potential. Geo-information technologies, particularly remote sensing, and GIS, serve as effective tools for understanding these geo-events. The advent of advanced technologies enables a comprehensive overview and faster analysis compared to traditional surveying methods (Liang & Liu, 2020; Sunar et al., 2019). Remote sensing has emerged as cost-effective and one of the popular choices among water resources and environmental researchers due to its high spatial coverage and temporal resolution. Numerous studies have utilized both active and passive remotely sensed data sources to assess their effectiveness in reliably capturing events like floods, droughts, etc. without requiring in-situ data collection (Gupta et al., 2022; F. Zhang et al., 2014). Integrating remote sensors and satellites in research offers long-term time series data, allowing researchers to map and understand changes in aspects due to climate changes and human advancements, land use change, increase in global

temperature, flooding patterns in different regions, change in forests and vegetation, etc. since

the launch of ERTS-1 (Earth Resources Technology Satellite), also commonly known as

Landsat-1 in 1972 (Shastry et al., 2023). Remote sensing data is preferred compared to site-based

data such as weather station records, and gauge data. By nature of remote sensing, researchers

have easier access to otherwise tedious data collection. The rising use of these techniques

enhances our understanding of flood dynamics and supports the study of hydrology and ecology

in flood-prone areas.

Microwave remote sensors, on contrary to optical sensors, operate on longer wavelengths than

optical remote sensors, ranging from 1 cm to 1m and at frequencies from 89 GHz to 0.3 GHz,

respectively. These properties allow them to penetrate cloud cover, aerosols, haze, and tree

canopies, facilitating all-weather flood inundation mapping (Tazmul Islam & Meng, 2022;

Ticehurst et al., 2009). Despite the popularity of microwave remote sensing in flood mapping,

most of the data, especially Synthetic Aperture Radar (SAR) was not freely available before

Sentinel-1, which often made extensive flood monitoring ineffective. The deployment of

Sentinel-1 has enabled the liberated use of SAR data and enhanced accessibility with its short

revisit times, providing SAR-C data weekly. SAR satellites have gained popularity in flood

detection and monitoring because of their ability to deliver high-resolution surface data

promptly, regardless of weather conditions (Hamidi et al., 2023). Flood detection with SAR

involves detecting changes in the backscattered signals, such as identifying dark regions on water

bodies that correspond to low backscatter signals compared to other land-cover categories.

Multiple techniques, including supervised classification and threshold algorithms, are used to

map floods using SAR data. Additionally, multitemporal approaches that utilize both pre- and

post-flood SAR data have proven to be highly reliable in generating accurate flood maps (Amitrano et al., 2018; Long et al., 2014).

SAR change detection, which involves generating different images and applying classification methods, has found widespread application in urban planning, flood mapping, agricultural monitoring, and other domains. Unsupervised methods are often preferred for SAR image change detection due to the complexity involved in classifying SAR images. Statistical techniques, including thresholding algorithms, have been successfully employed to extract flood extent information. Challenges encountered in flood detection using SAR data include the presence of radar shadows, layover effects, and the influence of wind and heavy rainfall on flood delineation (Amitrano et al., 2018). Despite these challenges, the benefits offered by SAR data outweigh the limitations, thereby establishing SAR as an invaluable tool for flood detection.

Google Earth Engine (GEE) is a cloud-based platform powered by Google and allows remote processing of geospatial data in the cloud without requiring the users to download the data to their devices. GEE facilitates accessing a comprehensive repository of multiple datasets and processing them without the need to download them. Allowing researchers to utilize the computational resources of Google has revolutionized Earth science research. (Kumar et al., 2022). GEE allows researchers to access remote sensing datasets seamlessly (Tazmul Islam & Meng, 2022). Long-term data are available in GEE which facilitates trend analyses of climate change, drought and flood conditions, crop yields, and many more (Venkatappa et al., 2021).

The escalating frequency and severity of flooding events globally, particularly in vulnerable regions like Pakistan, underscore the urgent need for advanced flood management techniques. Despite experiencing numerous extreme flood incidents, Sindh, a province in Pakistan remains

an understudied area in terms of comprehensive flood extent mapping using remote sensing techniques. The integration of Synthetic Aperture Radar (SAR) from Sentinel-1 and optical sensors from Sentinel-2 is significant in overcoming the limitations posed by traditional monitoring methods, especially under adverse weather conditions and cloud cover during floods. These technologies offer high spatial coverage and temporal resolution data, essential for accurate mapping. Particularly, the Modified Normalized Difference Water Index (MNDWI) has proven effective in enhancing water detection and has been a popular choice among researchers. MNDWI effectively reduces the misclassification errors often associated with the presence of built-up areas, vegetation, and soil and calculates water presence at the pixel level. It can lead to inaccuracies when assessing water bodies that do not occupy entire pixels. In contrast, the Desert Flood Index (DFI) improves upon MNDWI by introducing a correction factor that reduces the influence of vegetation on the water signal, making it more effective in arid and semi-arid regions where desert landscapes predominate alongside vegetative areas.

The objective of the study is to assess the effectiveness of the Sentinel-1 SAR backscatter-based Difference Image Index and Sentinel-2 reflectance-based Desert Flood Index to map the flood extent for the Sindh province of Pakistan. The objective of the study is achieved by utilizing the multispectral satellite images from Sentinel-1 and Sentinel-2. These satellites have frequent revisit times and high spatial resolution making them suitable for near-real time studies. A change detection algorithm based on the respective indices for these datasets was developed and optimized so that there is maximum agreement between the results in pixel levels. The change detection algorithm once tuned provides fast and reliable results in assessing flooding extents using near real-time SAR datasets.

2.2 Study Area and Data

2.2.1   Study Area

Sindh province lies in the southern part of Pakistan and is bordered by the Arabian Sea in the south, Balochistan in its north and west, Punjab in the northeast, and India in the east. Sindh covers approximately 141,000 km^2 and 17.7% of Pakistan. The Sindh province is situated in a tropical to subtropical zone, resulting in extreme summers with temperatures exceeding 46°C from May to August, and brief winters with average temperatures dropping to around 2°C from December to January. The region receives most of its rainfall between July and September. The major river in the province is Sindh whose tributaries such as the Jhelum, Chenab, Ravi, and Sutlej, also flow in the region. However, these rivers flow through other provinces before joining the Indus River in Sindh. Figure 2-1 shows the study area, Sindh province.

Figure 2-1. Map of the Study area: Sindh province of Pakistan.

2.2.2 Datasets

Sentinel-1 is a mission undertaken as a joint initiative of the European Commission (EC) and the

European Space Agency (ESA) and uses Synthetic Aperture Radar (SAR) technology. C-band

SAR sensors of Sentinel-1 operate under four modes, with short revisit times and dual-polarization capability. Long-term time series data collected by the Sentinel-1 mission has been available since April 2014. A satellite covers the Earth in 12 days. However, with a two-satellite constellation, a region of interest is mapped every six days. SAR tools can capture valuable data regardless of weather conditions, including cloudy situations, and operate both day and night. The mission provides imagery at three different levels of detail, with spatial resolutions of 10, 25, and 40 meters. Four combinations of polarization: single polarizations, vertical-vertical (VV) or horizontal-horizontal (HH) bands, as well as dual-band options such as VV + VH and HH + HV, are available (Kaplan & Avdan, 2018). These features make the Sentinel-1 images informative for various applications such as sea-ice mapping, humanitarian aid, crisis response, and forest management. Sentinel-1 data Ground Range Detected (GRD) imagery can be accessed in Google Earth Engine (GEE) as 'COPERNICUS/S1_GRD'.

The study employs a comprehensive approach integrating multisource and multitemporal remote sensing datasets to accurately extract floodwater information. The Sentinel-1 GRD dataset is used for flood inundation mapping. Sentinel-1 SAR GRD data, available in Google Earth Engine as 'COPERNICUS/S1_GRD'. The spatial resolution for this dataset is 10 m. The revisit time is 12 days. The image collection is filtered to access images with 'VV' polarization, instrument mode 'IW,' and orbital pass 'Ascending' for obtaining flood extent maps in the region of interest. JRC Global Surface Water Mapping Layers, v1.4 can be accessed from the Google Earth Engine as 'JRC/GSW1_4/GlobalSurfaceWater' and is used to obtain permanent water bodies based on seasonality. The water bodies are generated from long-term continuous images between 16 March 1984 and 31 December 2021 from Landsat 5, 7, and 8. NASADEM, 'NASA/NASADEM_HGT/001' is a modernized Digital Elevation Model developed from the

Shuttle Radar Topography Mission (SRTM). The resolution of this dataset is 30 m data used to filter out the pixels whose slope is greater than 5°. These datasets are available in the GEE platform and are used in conjunction to determine the flooded regions.

The sentinel-2 dataset is used to validate the flood extent result obtained from Sentinel-1 imageries. The dataset is available in GEE as 'COPERNICUS/S2'. The revisit frequency of satellites of Sentinel-2 missions is 10 days, and the combined frequency of the constellation is 5 days. The spatial resolution of the S2 dataset in GEE ranges from 10-60 m depending on the bands. The bands suitable for water mask detection have been used in this study. The datasets used in this study are given in Table 2-1.

Table 2-1. List of Data in the study.

Dataset	Spatial Resolution	Revisit Days	Corresponding Date	Source
Sentinel-1 SAR (S1)	10 m	12 days	-	ESA Copernicus
Sentinel-2 (S2)	B3, B4, B8: 10 m and B11: 20 m	5 days	-	ESA Copernicus
JRC Global Surface Water Mapping	30 m	-	March 1984 - December 2021	Global Surface Water Explorer
NASADEM	30 m	-	February 2000	NASA

2.3 Methodology

This section describes the methodology implemented to achieve the objectives of investigating and optimizing the change detection-based thresholding algorithm for mapping the flooding extents in Pakistan. The methodology is presented in Figure 2-2. The methodology consists of the following steps:

- Data acquisition and preprocessing,

- Change detection algorithm,

- Sentinel-1 flooding extent,

- Sentinel-2 flooding extent,

- Threshold optimization,

- Results and Analyses.

Sentinel-1
Metadata Filtering

Speckle filtering

Pre flood images

Post flood images

Sentinel-2
Metadata Filtering

Post flood image

SAR Image Index DII

S2 Indices, DFI/MNDWI

Scale DII (0-1)

Scale Index (0-1)

Set DII threshold
Thr = mean(DII)+k_r.stdev(DII)
k_r(0:0.05:2)

Index threshold
Thr (0:0.05:1)

Binary Flood Mask

Binary flood mask

Remove steep slopes
and permanent water

Remove clouds and
permanent water

DII Flood extent
map (S1 FEM)

S2 Indices
Flood extent map
(S2 FEM)

Agreement between
S1 FEM and S2 FEM

Is DFI Thr
> Thr
max?   No

Yes

Is k_r >
k_r max?   No

Yes

Select optimal
threshold

Threshold fine-tuning

Final DII Flood
extent map

Figure 2-2. Flowchart diagram of flood extent mapping using Sentinel-1 and Sentinel-2 images.

2.3.1 Data Acquisition and Preprocessing

The datasets used in this study are available in Google Earth Engine and all the processing and computations have been carried out in GEE. Sentinel-1 and Sentinel-2 datasets are used in this study. Sentinel-1 data is available In GEE as 'COPERNICUS/S1_GRD'. For the flood inundation mapping study, the Sentinel-1 images acquired between June to September have been used; the details are presented in Table 2.2. Sentinel-1 GRD is processed to backscatter coefficient ($\sigma°$) in decibels (dB) in GEE. The scenes are corrected for thermal noise removal, radiometric calibration, and terrain correction. However, SAR images are affected by interference between returns from backscatters of the surface, and patterns like salt and pepper can be observed, which are also called speckles. Therefore, speckle filtering needs to be applied during the preprocessing of Sentinel-1 SAR images in GEE. (Dasari et al., 2016; Tazmul Islam & Meng, 2022). Sentinel-2 data have been used for optimizing the threshold value of the Sentinel-1 image index. The cloud coverage during the period of the study was high due to which, a rectangular region in Dadu district was selected where cloud cover was almost insignificant. Table 2-2 shows the data obtained for each objective of the study.

Table 2-2. Data acquisition for flood extent mapping.

Objective	Dataset	Description	Acquisition date
Thresholding (A rectangular region in Dadu district)	Sentinel-1	Before flood	'2022-05-01' - '2022-05-14'
		After flood	'2022-08-18' - '2022-08-30'
	Sentinel-2	Before flood	'2022-05-09' - '2022-05-15'
		After flood	'2022-08-29' - '2022-08-30'
Flood Inundation Mapping	Sentinel-1	Before flood	'2022-05-01' - '2022-05-14'
		After flood, August	'2022-08-01' - '2022-08-31'
		After flood, September	'2022-09-01' - '2022-09-19'

2.3.2   Change Detection Algorithm

The change detection method is a well-established method for detecting changes due to floods and has been explored by several researchers (Amitrano et al., 2018; Cian et al., 2018; Clement et al., 2018; Giustarini et al., 2013; Hamidi et al., 2023; Kumar et al., 2022; Long et al., 2014;

Nasirzadehdizaji et al., 2019). The change detection method compares two scenes of the same location to quantify the changes that occurred in the interval. In this study, a change detection-based thresholding method has been implemented.

2.3.3   Sentinel-1 Flooding Extent

In this study, for Sentinel-1 flooding extent, two stacks of images were retrieved for pre-flood and post-flood conditions in Google Earth Engine. Speckles filters are applied. For pre-flood or reference and post-flood Sentinel-1 images, the mean, minimum, and maximum of 'VV' values are computed. For the compound reference image, the mean values of the backscatter of each pixel are obtained, and this represents the average behavior of the land covers before flooding. The backscatter values of water bodies are lower than other land classes. This is because of the smoothness of the water surfaces. However, these values can be medium-high due to wind (Cian et al., 2018). For non-smooth classes of land use, the mean backscatter values are high. This distinction between spectrally smooth and non-smooth land-use classes helps in distinguishing land-use classes impacted by flood. The backscatter values of land-use classes are reduced when the region is under water or impacted by flood which is why the minimum values of the stack of post-flood images are obtained. The change detection method distinguishes the difference between these values to identify the floods.

A SAR index called the Difference Image Index is used in the study. DII is quantified as the difference between absolute values of minimum backscatter values of post-flood images stack and absolute values of mean or average of the reference images stack. The SAR image index DII, is given by Hamidi et al. (2023) and is calculated as,

$$DII = \left|min\{VV_{after}\}\right| - \left|mean\{VV_{before}\}\right| \tag{1}$$

where $min\{VV_{after}\}$ represents the minimum backscatter values of Sentinel-1 images of the pre-flood stack and $mean\{VV_{before}\}$ represents the mean of backscatter values of Sentinel-1 images of the post-flood stack.

There is a substantial change to the backscatter values of the pixels when the flooding occurs. Therefore, thresholding is required to distinguish between flooded and non-flooded regions. Thresholding approaches commonly used in flood studies are based on the distribution of histograms. Hamidi et al. (2023) defined the threshold value for flooding as,

$$Thr = mean\{DII\} + k_{DII} \times stdev\{DII\} \tag{2}$$

The thresholding values are obtained for areas where cloud cover was insignificant. Regions where the DII values of the pixels were less than the threshold were considered flooded.

2.3.4 Sentinel-2 Flooding Extent

The flooding extent observed from DII is compared with a flooding extent observed from a widely used optical sensor, Sentinel-2 data. Sentinel-2 satellites are equipped with high-resolution optical sensors, making them particularly useful for observing and analyzing changes in land cover and vegetation before and after flood events. In this study, Sentinel-2 images were acquired for pre-flood and post-flood conditions to identify changes in vegetation health, soil saturation, and water coverage. The analysis focused on the utilization of specific spectral bands and vegetation index, Normalized Difference Vegetation Index (NDVI) to get the flooding index

called the Desert flood index (DFI). DFI has shown better results in desert areas with vegetation compared to the Modified Normalized Difference Water Index (MNDWI). The DFI approach has also been explored in the Lower Indus Basin of Pakistan for mapping 2010 floods (Baig et al., 2013). MNDWI is a widely accepted index for flooded pixel detection. The comparison of pre-flood and post-flood indices allowed for the identification of areas affected by flooding, quantifying the extent of inundation, and evaluating the impact on different land cover classes.

The DFI was obtained from the following equations,

$$DFI = \frac{\rho_{green} - \rho_{swir} + 0.1}{(\rho_{green} - \rho_{swir})(NDVI + 0.5)} \tag{3}$$

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \tag{4}$$

$$MNDWI = \frac{\rho_{green} - \rho_{swir}}{\rho_{green} + \rho_{swir}} \tag{5}$$

where $\rho green$, ρred, ρnir, and $\rho swir$ are the reflectance values of B3, B4, B8, and B11 bands of Sentinel-2, respectively. Since clouds and their shadows can obscure features and distort information, the clouded pixels were masked out from the computation. Permanent water bodies were removed from the flood extent image. Regions where the DFI and MNDWI values were higher than the threshold values are taken as flooded for both cases.

2.3.5 Threshold Fine-Tuning

In this study, threshold optimization is done using equations from Hamidi et al. (2023). In evaluating the accuracy of SAR-based flood maps, agreement percentage ($Agr\%$) between SAR indices and results from optical imagery has been determined. This percentage is calculated using the formula,

$$Agr\% = \frac{B}{O} \times 100\% \tag{6}$$

where B represents the overlapping area of flood detections from both SAR and optical sources, and O represents the area of flood detections in the optical imagery. Essentially, this equation quantifies the proportion of SAR flood detection that coincides with optical water detection.

The total agreement percentage ($ToAgr\%$) is calculated to estimate flood extents from both SAR and optical sources. This equation serves for sensitivity analysis of flood extent maps derived from SAR- and Optical-based indices. $ToAgr\%$ is calculated as,

$$ToAgr\% = \frac{B}{O + S_{rc} - B} \times 100\% \tag{6}$$

where S_{rc} represents the area of SAR flood detections excluding regions covered by clouds in the optical image. Cloud-covered areas are excluded from the SAR flood extent map to ensure comparability with the optical image flood extent. Subtracting B from the denominator avoids counting the intersection area twice. This equation effectively illustrates the overall disparities between SAR and optical flood detection areas in regions unaffected by clouds. This takes into

account disagreements, capturing instances where SAR might inaccurately classify areas as flooded.

The optimal values of k_{DII} and DFI, and k_{DII} and MNDWI have been determined by the exhaustive grid search, whereby k_{DII} values between 0 to 2 (Hamidi et al., 2023; Cian et al.,2018), and DFI and MNDWI values between 0 to 1 were tested at a constant interval of 0.05. The agreement percentage and total agreement percentages are obtained for each set of k_{DII} and DFI and k_{DII} and MNDWI. The optimal values of k_{DII} are obtained for the values of DFI and MNDWI whose agreement percentage and total agreement percentages were highest. Higher percentages of these agreements define the sensitivity of Sentinel-1-derived flood maps to Sentinel-2-derived flood maps.

2.4 Results

2.4.1 Thresholding

Sentinel-1 SAR images and Sentinel-2 images were used simultaneously to obtain the optimum threshold values for Sentinel-1-based DII images and Sentinel-2-based DFI and MNDWI images. A rectangular region, approximately 3,220 km^2 in Dadu district with minimum cloud cover was chosen for threshold calibration. Pre-flood and post-flood images were obtained for the rectangular region and their composite images were prepared by computing the mean and minimum of the backscatter values from the image stacks, respectively. Figure 2-3 shows the mean of backscatter values from May 1 to May 14 as reference image composite or pre-flood images composite, and the minimum of backscatter values for after flood period from August 18 to August 30, 2023. A comparison of these pre- and post-flood image composites shows an increase in the darker regions in the post-flood images due to a reduction in backscatter signals.
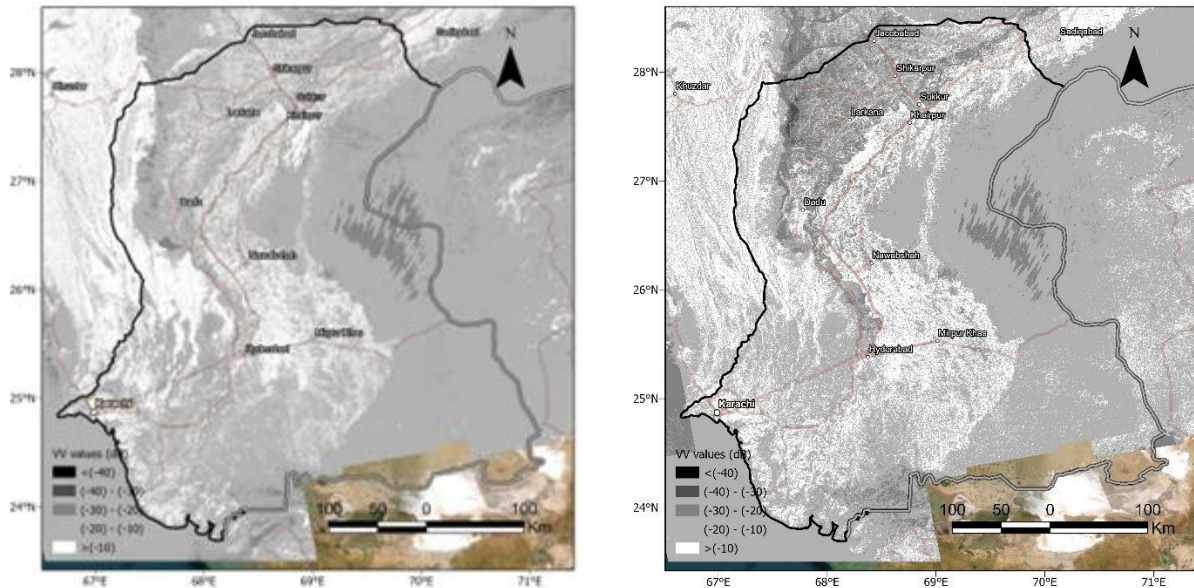
Figure 2-3. Sentinel-1 images backscatter values in dB before and after Pakistan floods 2022.

35

Flood extent mappings were computed for Sentinel-1 and Sentinel-2 images. The Difference

Image Index (DII) image was computed for the region of interest using equation (1), the Desert

Flood Index (DFI) image using equation (3), and the Modified Normalized Difference Water

Index (MNDWI) using equation (5).

2.4.2   Case I: Thresholding for DII and DFI

In the comparative analysis of flood extents between DII and DFI flood extent mappings, notable

agreement and total agreement percentages were observed. The total agreement has been

preferred over the agreement since it depicts the percentage of intersection of floods in Sentinel-

1 and Sentinel-2 images. Figure 2-4 shows a 3D surface plot generated to visualize the region of

maximum agreement and total agreement region. The values of total agreement and agreement

percentages for values of DFI and k_{DII} from 0.2 to 0.45 have been provided in Table 2-3 and

Table 2-4 respectively. At a DFI value of 0.25, the comparison between DFI and DII flood

extents achieved a total agreement of 72.08% with a k_{DII} value set at 0.35. Additionally, the

agreement between them stood at 79.31%. The optimum threshold value for DII was computed

as 0.414.

Table 2-3. Total Agreement percentage of DII and DFI Flood extent maps at different thresholds.

		DFI					
		0.2	0.25	0.3	0.35	0.40	0.45
k_{DII}	0.2	70.97%	70.78%	69.71%	67.93%	65.52%	62.64%
	0.25	71.40%	71.53%	70.73%	69.19%	66.94%	64.15%
	0.3	71.54%	71.99%	71.50%	70.21%	68.15%	65.47%
	0.35	71.31%	72.08%	71.90%	70.89%	69.06%	66.52%
	0.4	70.72%	71.79%	71.91%	71.19%	69.60%	67.24%
	0.45	69.84%	71.17%	71.60%	71.16%	69.84%	67.68%

Table 2-4. Agreement percentage of DII and DFI Flood extent maps at different thresholds.

		DFI					
		0.2	0.25	0.3	0.35	0.4	0.45
k_{DII}	0.2	81.28%	83.38%	85.02%	86.28%	87.15%	87.77%
	0.25	79.86%	82.08%	83.86%	85.24%	86.22%	86.92%
	0.3	78.39%	80.74%	82.65%	84.16%	85.25%	86.03%
	0.35	76.84%	79.31%	81.34%	82.98%	84.20%	85.06%
	0.4	75.24%	77.81%	79.97%	81.73%	83.06%	84.02%
	0.45	73.58%	76.24%	78.51%	80.39%	81.84%	82.90%
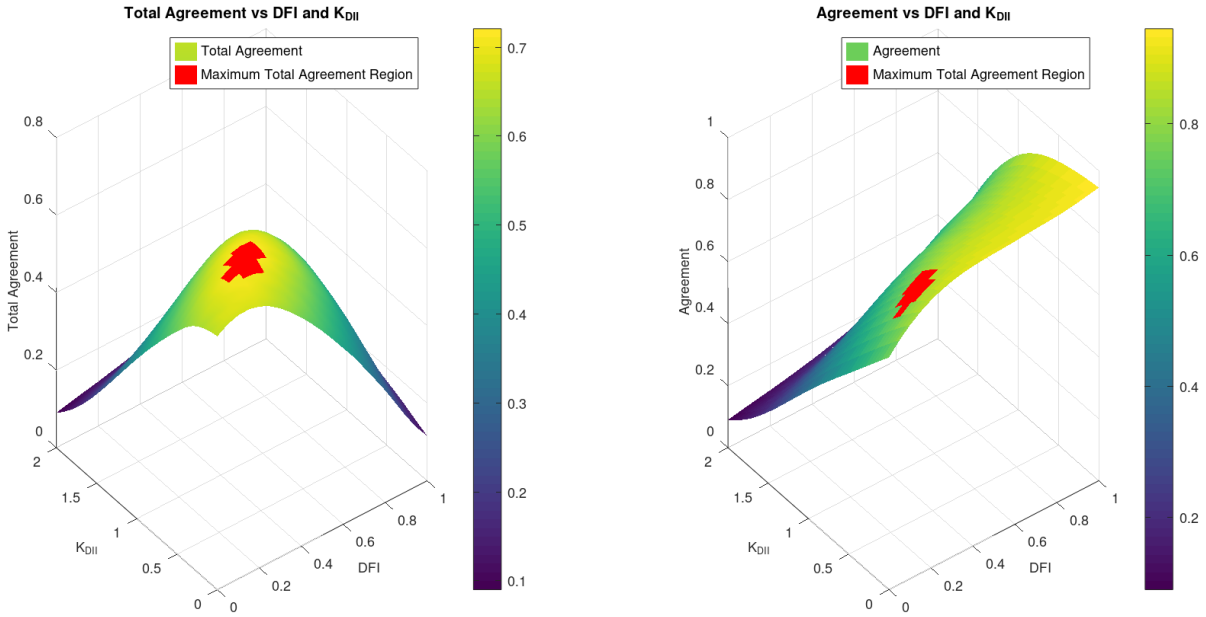
Figure 2-4. 3D-plot of agreement and total agreement at different values of DFI and k_{DII}.

The Figure 2-5, Figure 2-6, and Figure 2-7 depict the utilization of DII and DFI indices for flood detection in a specific rectangular area within the Dadu district, with Figure 2-5 and Figure 2-6 showcasing the individual application of DII and DFI indices, respectively, for threshold-based flood identification, and Figure 2-7 illustrating the agreement between the flood detection outcomes of both indices within the same region.

Figure 2-5. Flood detection in the rectangular region in Dadu district using the DII index for thresholding.

Figure 2-6. Flood detection in the rectangular region in Dadu district using the DFI index for thresholding.

Figure 2-7. Agreement visualization of DII and DFI Flood detection in the rectangular region of Dadu district.

Flooded regions identified solely through the Difference Image Index (DII) data were labeled with the value 1, encompassing an area of 119.65 km^2. Likewise, areas flooded exclusively according to the Desert Flood Index (DFI) were designated with the value 3, amounting to 223.77 km^2. Regions concurrently identified as flooded by both DII and DFI indices were assigned the value 4, covering a significant area of 877.57 km^2. The agreement percentage of 79.31% and total agreement of 72.08% indicate a high level of reliability in the coherence between the flood detection capabilities of Sentinel-1 and Sentinel-2 datasets.

2.4.3 Case II: Thresholding for DII and MNDWI

The methodology was repeated for the comparative analysis of flood extents between DII and MNDWI flood extent mappings. Figure 2-8 shows a 3D surface plot generated to visualize the region of maximum agreement and total agreement region. The values of total agreement and agreement percentages for values of MNDWI and k_{DII} ranging from 0.1 to 0.35 have been provided in Table 2-5 and Table 2-6 respectively. At an MNDWI value of 0.15, the comparison between MNDWI and DII flood extents achieved a total agreement of 70.06% with a k_{DII} value set at 0.25. Additionally, the agreement between them stood at 83.86%. The optimum threshold value for DII was computed as 0.427.
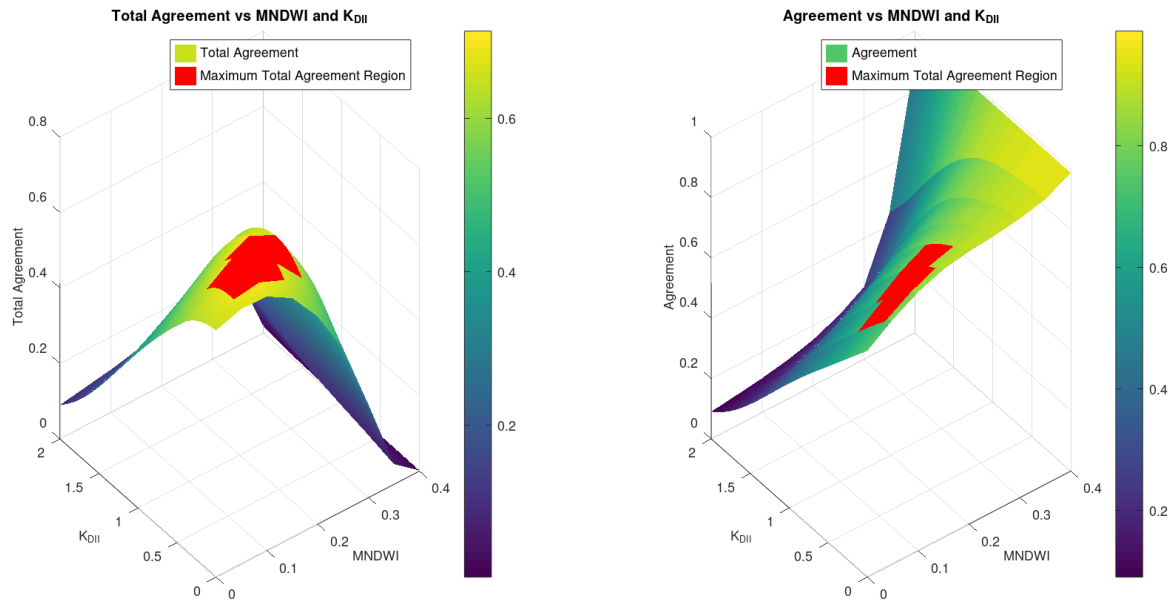
Figure 2-8. 3D-plot of agreement and total agreement at different values of MNDWI and k_{DII}.

Table 2-5. Total Agreement percentage of DII and MNDWI Flood extents at different thresholds.

		MNDWI					
		0.1	0.15	0.2	0.25	0.3	0.35
k_{DII}	0.1	69.26%	66.40%	58.32%	45.81%	26.47%	5.81%
	0.15	70.00%	67.78%	59.95%	47.32%	27.47%	6.06%
	0.2	70.56%	69.01%	61.52%	48.81%	28.49%	6.32%
	0.25	70.91%	70.06%	62.99%	50.27%	29.51%	6.59%
	0.3	70.98%	70.85%	64.27%	51.60%	30.49%	6.85%
	0.35	70.69%	71.28%	65.28%	52.76%	31.40%	7.11%

Table 2-6. Agreement percentage of DII and MNDWI Flood extents at different thresholds.

| | | MNDWI | | | | | |
		0.1	0.15	0.2	0.25	0.3	0.35
k_{DII}	0.1	83.25%	87.05%	89.02%	90.12%	91.69%	94.28%
	0.15	81.94%	86.07%	88.29%	89.55%	91.27%	94.15%
	0.2	80.55%	85.00%	87.49%	88.90%	90.81%	94.02%
	0.25	79.10%	83.86%	86.62%	88.22%	90.31%	93.84%
	0.3	77.61%	82.66%	85.71%	87.49%	89.77%	93.67%
	0.35	76.05%	81.37%	84.73%	86.71%	89.21%	93.45%

Obtained flood extent maps from the MNDWI and DII index were examined at the pixel level. A comparison map was prepared by mapping the study region into 4 distinct classes assigning numeric values: 0, 1, 3, and 4, each representing specific flood conditions. Pixels assigned the value 0 indicated non-flooded regions, signifying non-flooded by both Sentinel-1 and Sentinel-2 images. The regions inundated solely based on the Difference Image Index (DII) data were categorized with the numerical value 1, encompassing an area of 209.45 km^2. Similarly, areas

submerged exclusively as per the Modified Normalized Difference Water Index (MNDWI) were denoted by the value 3, totaling 165.88 km^2. Regions concurrently identified as flooded by both DII and MNDWI indices were assigned the value 4, covering a substantial area of 865.58 km^2. The agreement percentage of 70.06% and the total agreement of 83.86% demonstrate a reliable coherence in the flood detection capabilities between Sentinel-1 and Sentinel-2 datasets.

Figure 2-9 and Figure 2-10 illustrate the application of the MNDWI and DII indices, respectively, for flood detection in a specific rectangular region within Dadu district, while Figure 2-11 presents a comparative visualization, showcasing the agreement between the flood detection results obtained from both the DII and MNDWI methods in the same area.
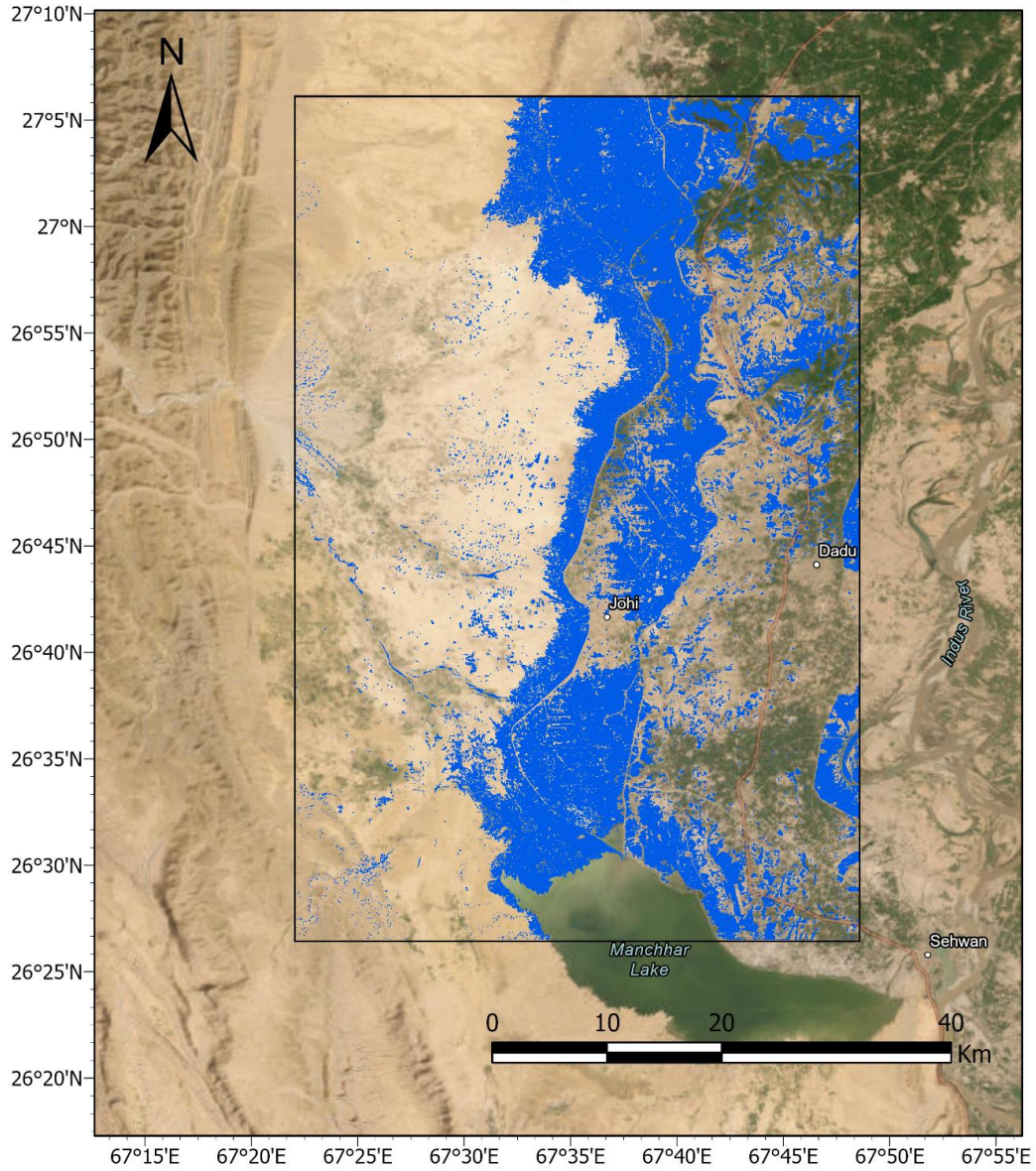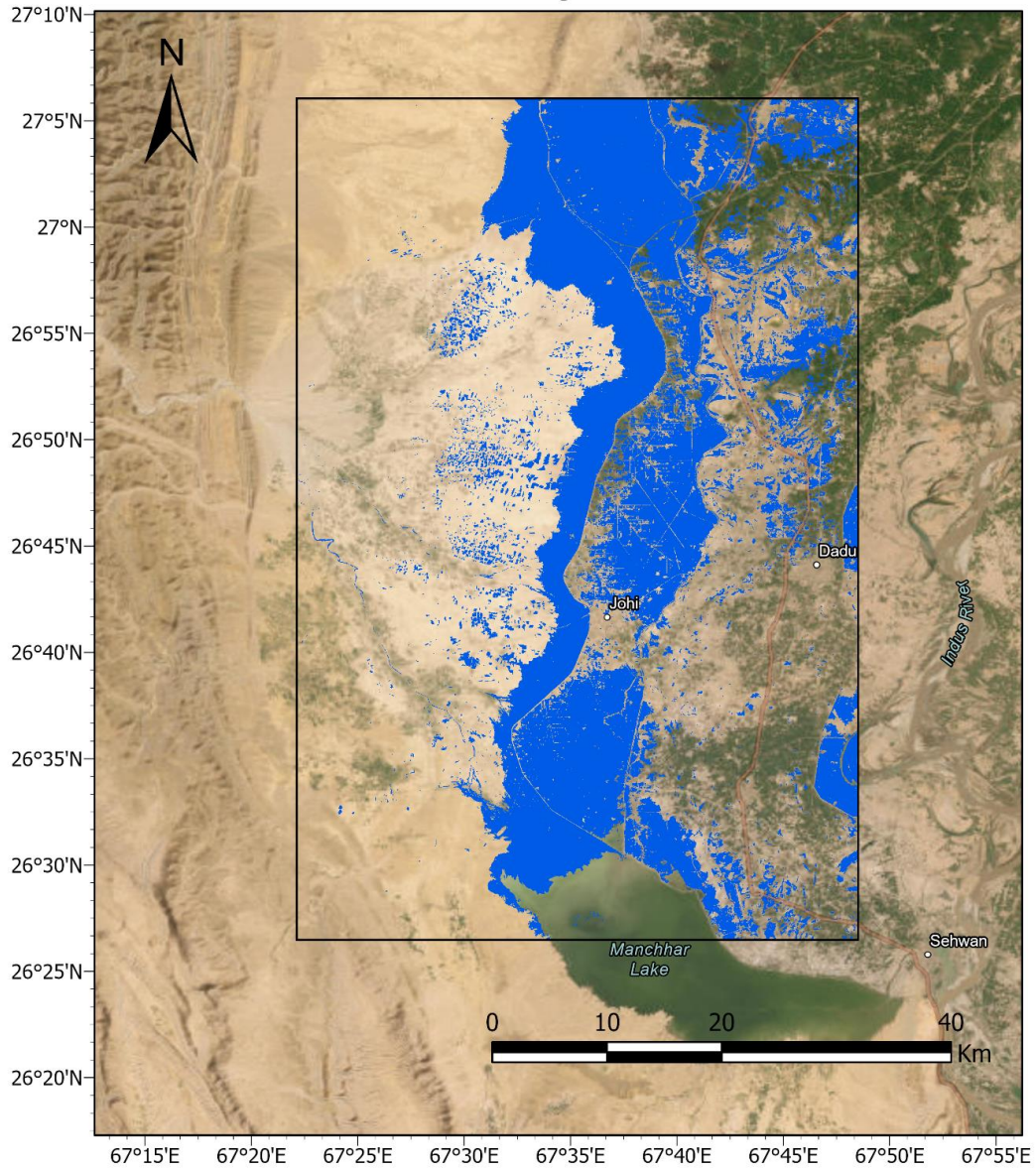
Figure 2-9. Flood detection in the rectangular region in Dadu district using the MNDWI index for thresholding.
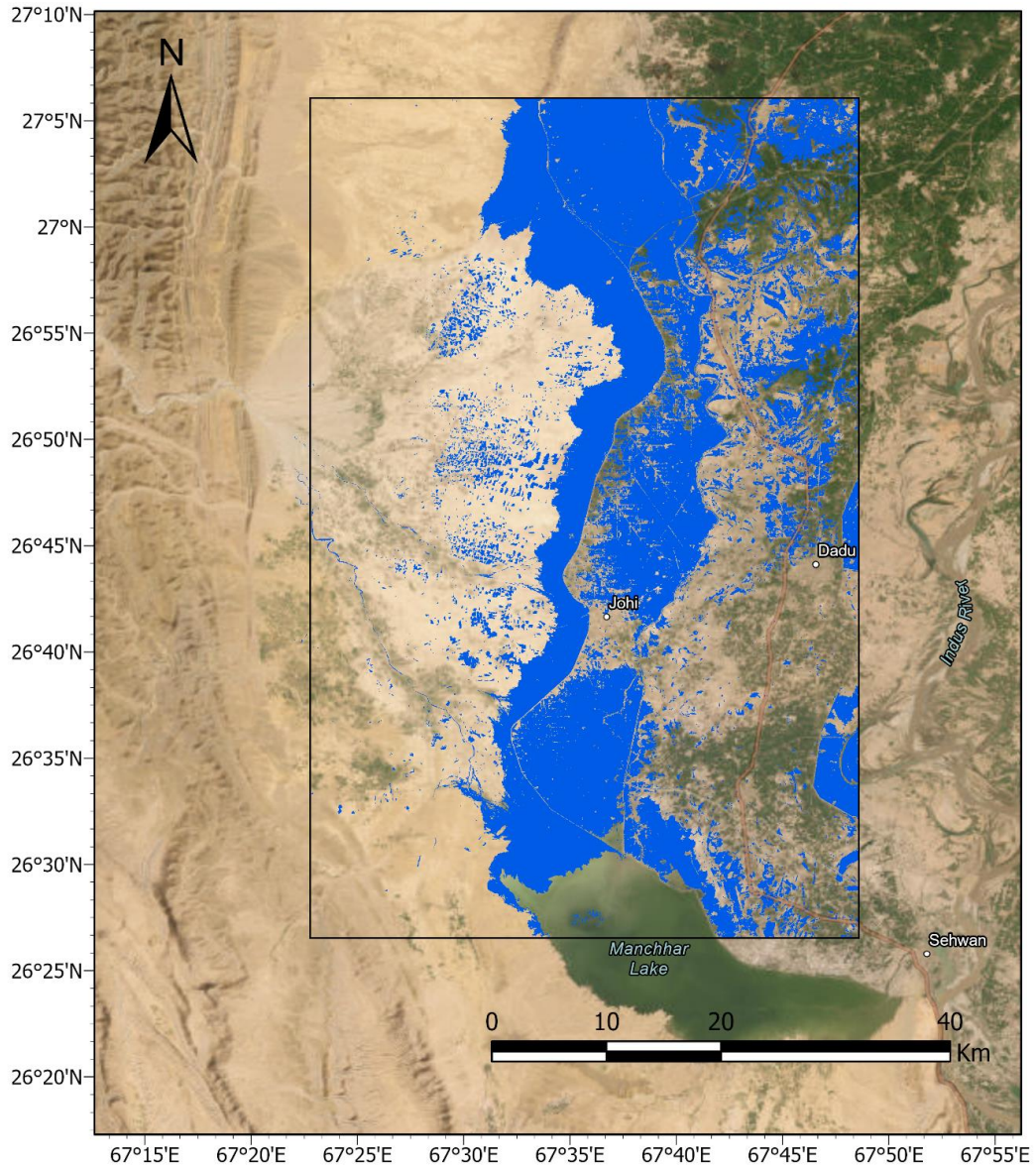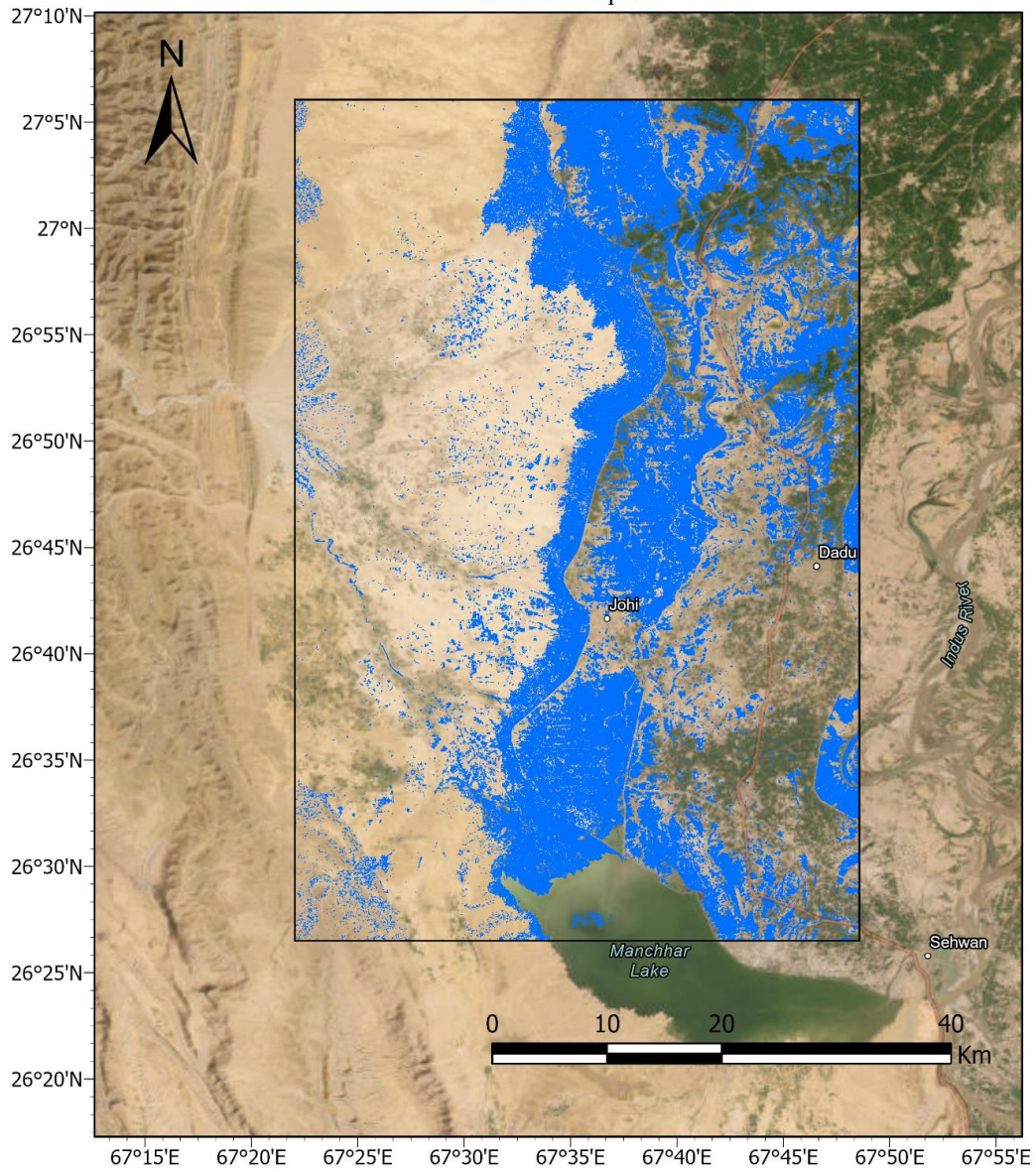
Figure 2-10. Flood detection in the rectangular region in Dadu district using the DII index for thresholding.
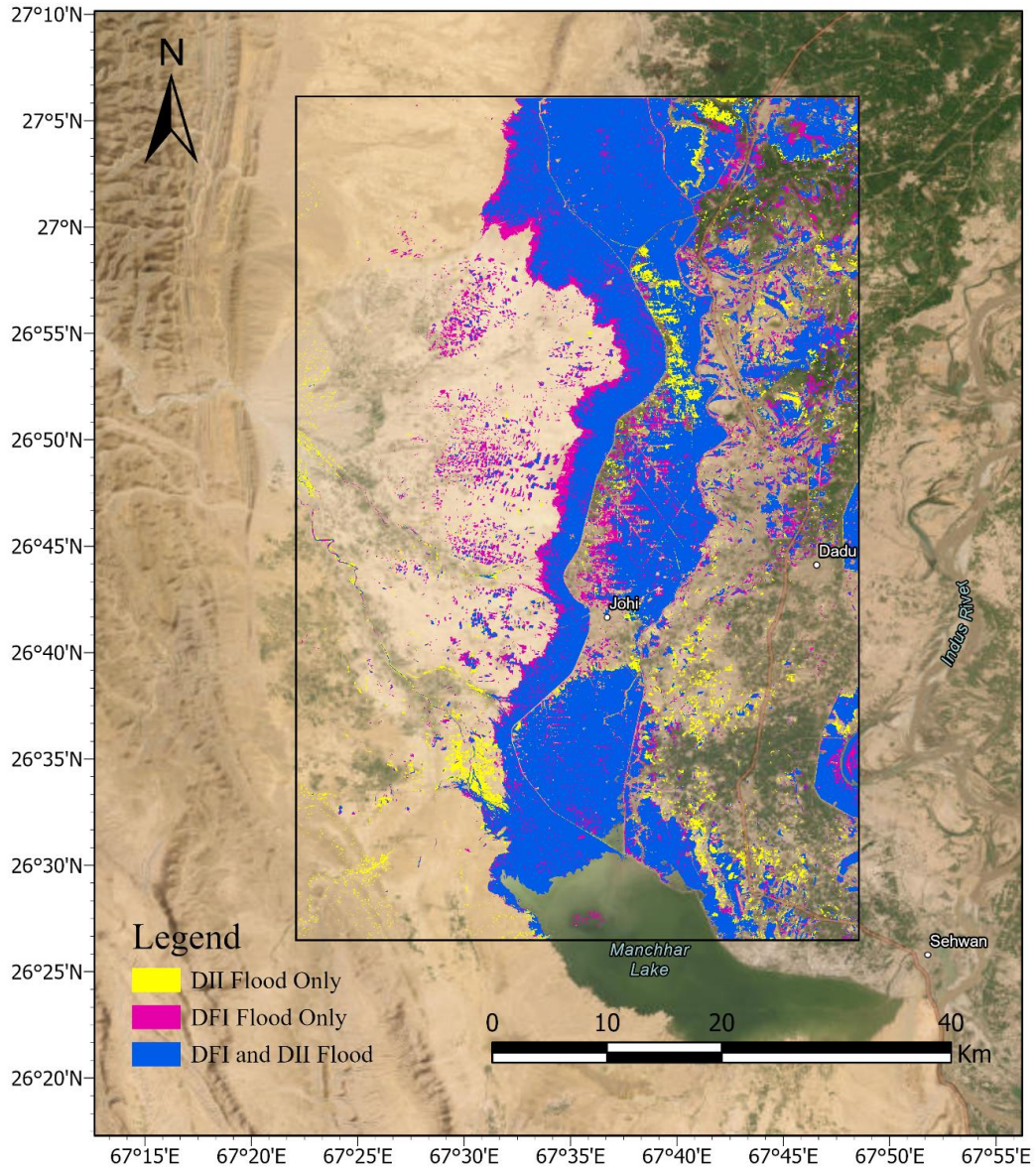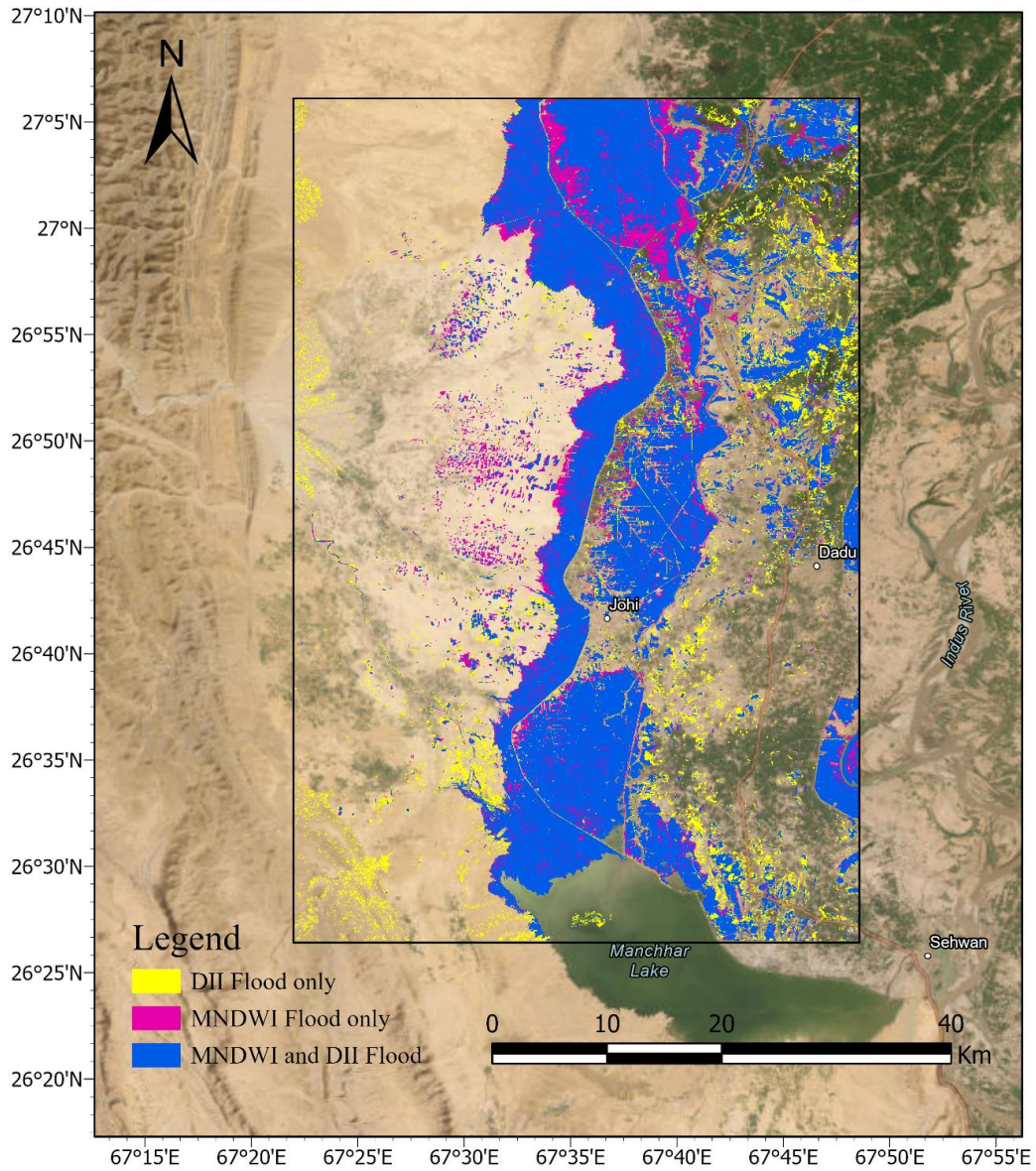
Figure 2-11. Agreement visualization of DII and MNDWI Flood detection in rectangular region of Dadu district.

2.4.4   Flood Extent Mapping of Sindh

Utilizing the refined threshold value obtained for the Difference Image Index (DII) through the integration of both the Desert Flood Index (DFI) and the Modified Normalized Difference Water Index (MNDWI), detailed flood extent maps of Sindh were generated for the flooded months of August and September 1 to September 19. The resulting inundation maps delineated the areas affected by floods within the Sindh province are shown in Figure 2-12, and Figure 2-13 respectively for August and September 1-19 using threshold values derived from analyses of DFI and DII values. It was found that 25,229 km^2 was inundated in August, and 19,181 km^2 in the first 19 days of September. Similarly, the flood extent maps derived from threshold analysis between MNDWI and DII indices (Figure 2-14, Figure 2-15) show that the total flooded area was 30,582 km^2 for August and 21,851 km^2 in the first 19 days of September respectively.
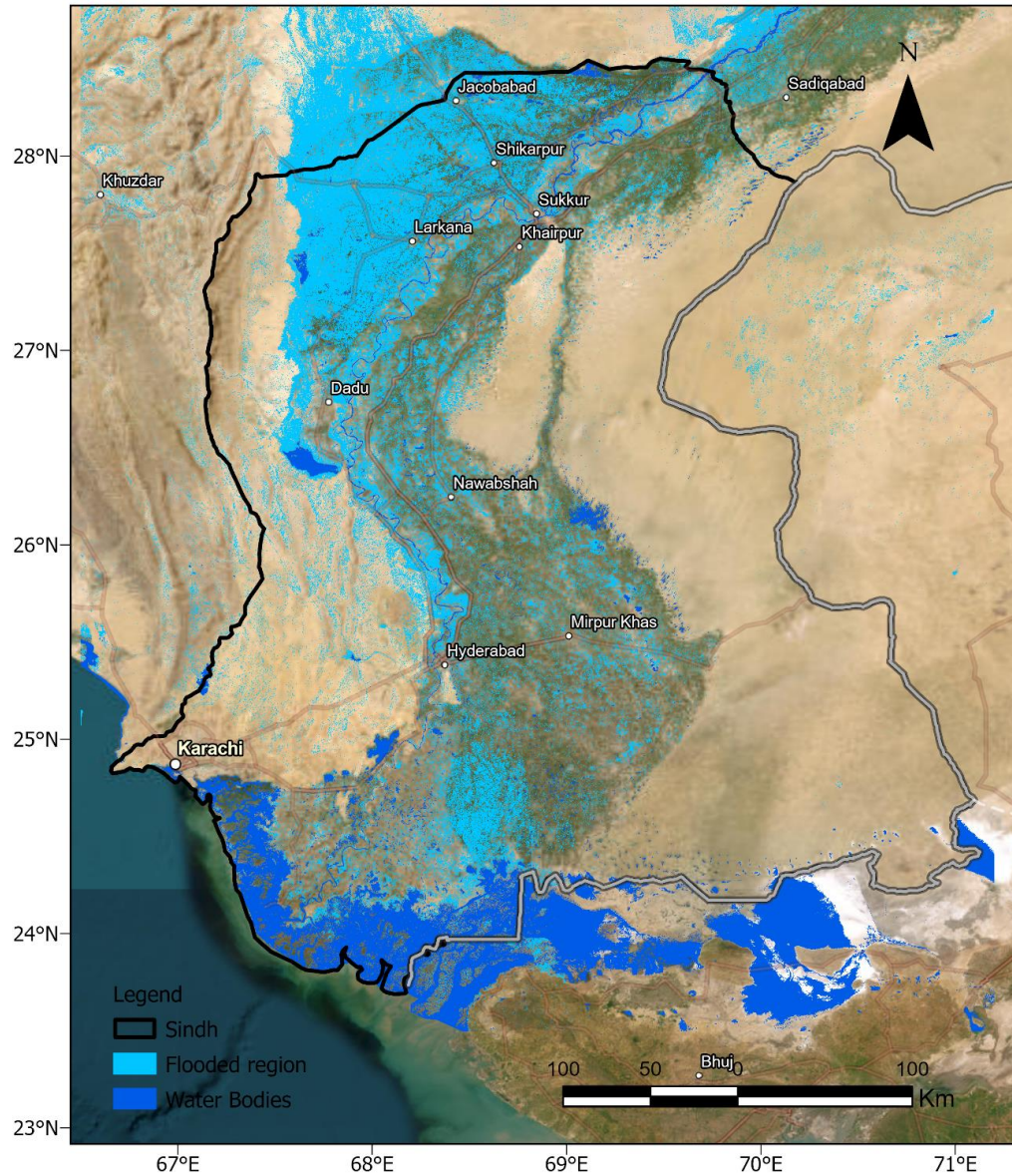
Figure 2-12. Flood Extent Map of Sindh province for August using DII threshold defined by agreement assessment between DFI and DII.
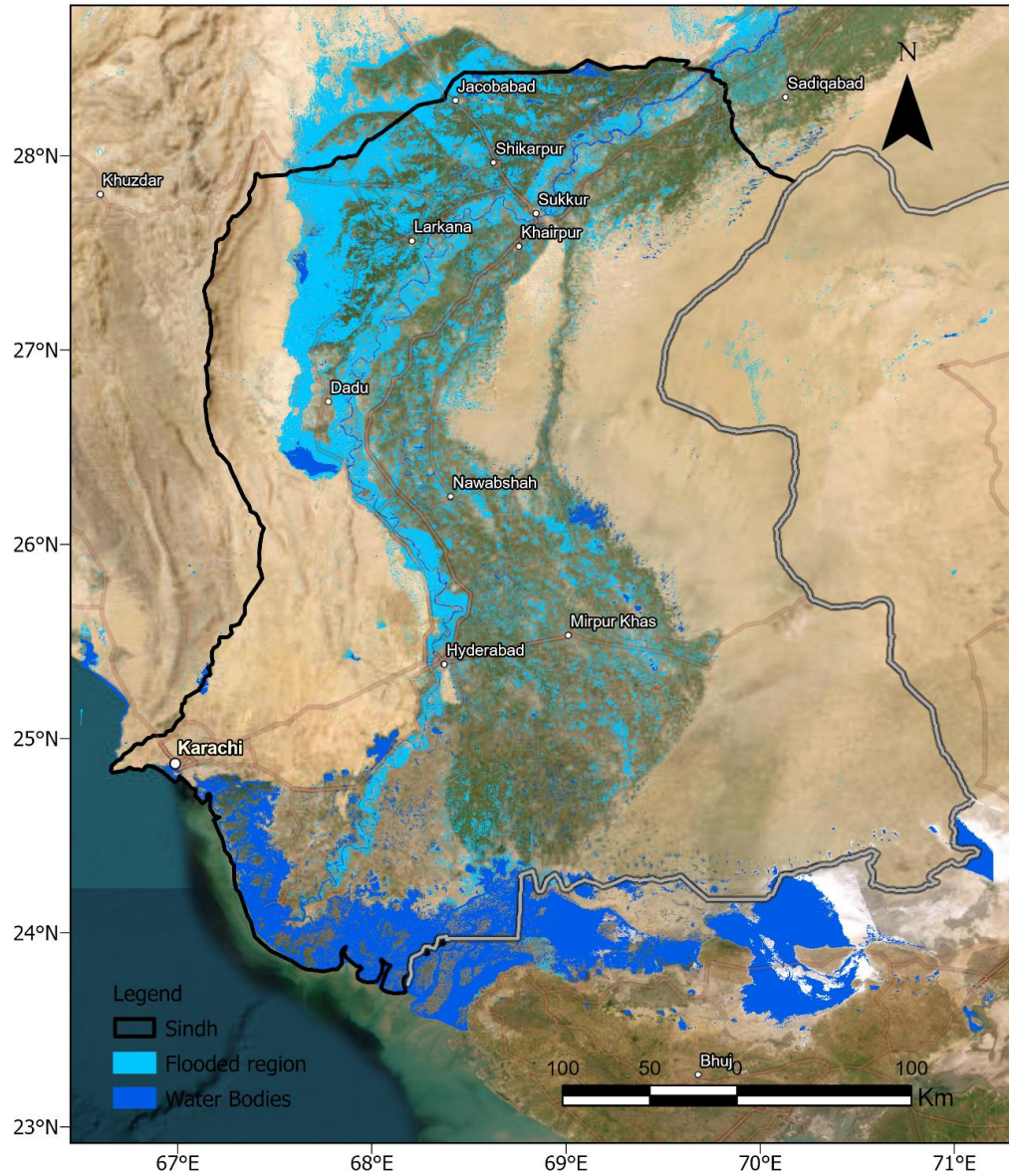
Figure 2-13. Flood Extent Map of Sindh province for September 1-19 using DII threshold defined by agreement assessment between DFI and DII.
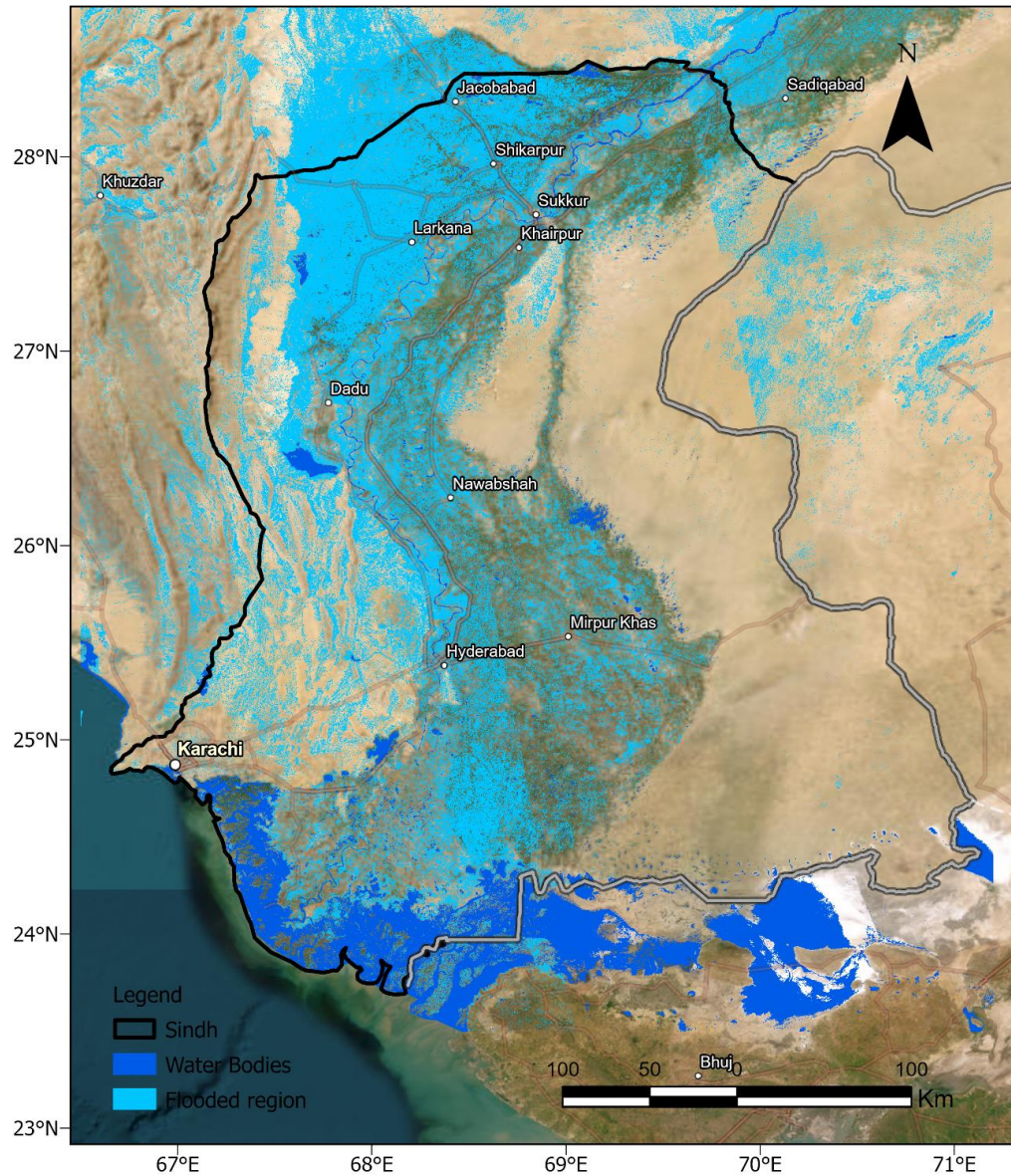
Figure 2-14. Flood Extent Map of Sindh province for August using DII threshold defined by agreement assessment between MNDWI and DII.

.
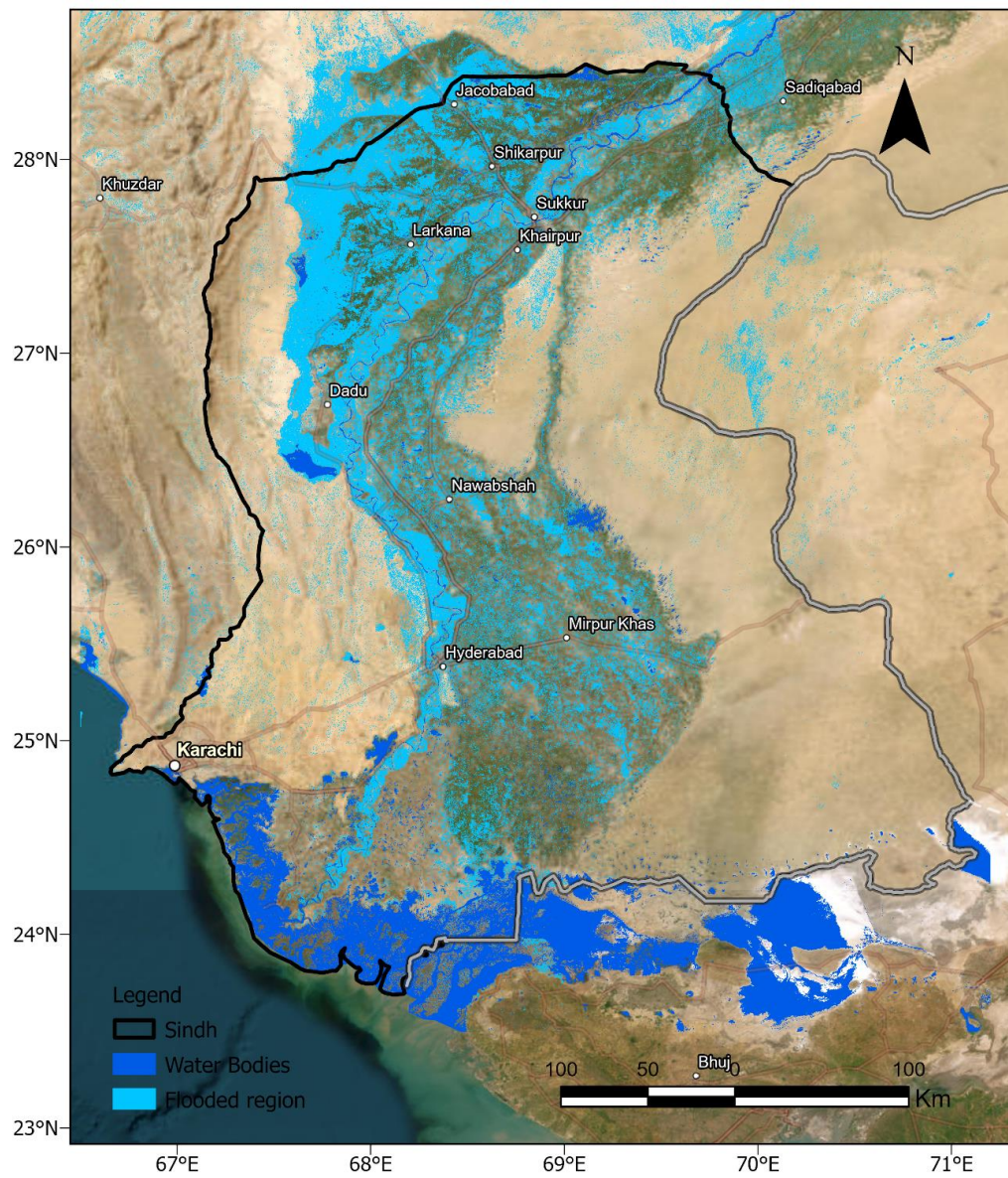
Figure 2-15. Flood Extent Map of Sindh province for September 1-19 using DII threshold defined by agreement assessment between MNDWI and DII.

2.5 Discussion

The study's results underscore the utility of Sentinel-1 SAR and Sentinel-2 optical data in mapping flood extents, especially in regions like the Sindh province, where adverse weather conditions and cloud cover can impede the use of optical data alone. The integration of SAR with optical data improves the accuracy and reliability of flood mapping efforts through the optimization of their respective indices' thresholds. The methodology has been applied through the Google Earth Engine (GEE) platform which allows researchers to access and analyze a wide array of satellite images and geospatial datasets. GEE allows the users to compute complex geospatial arrays leveraging Google's cloud infrastructures without having to download them into the user's computer.

Pakistan experienced severe flooding, intensified by a sequence of extreme weather phenomena. Notably, between the 16th and 25th of August, extreme rainfall events were recorded, with the most intense two-day precipitation observed on the 17th and 18th of August, followed by another notable rainfall event on the 24th and 25th of August. This period was marked by anomalously high precipitation levels. The heavy rainfall persisted into mid-September, contributing to the widespread and devastating impact of the floods across the region.

The flood extent maps have been obtained by classifying the DII index images with DFI index images and MNDWI index images by thresholding using an exhaustive grid search method. The threshold value of DII (0.414227) was obtained at k_{DII} value at 0.35 during the thresholding with DFI index (0.25) with the agreement and total agreement observed at 79.31% and 72.08%, respectively. The threshold value of DII (0.42718) was obtained during thresholding with MNDWI Index (0.15), whereby k_{DII} was obtained at 0.25, with the agreement and total

agreement observed at 83.86% and 70.06%, respectively. A similar study was conducted by

Hamidi et al. (2023) to assess flood impacts from Hurricane Ida in 2021 near New Orleans,

Louisiana, and Hurricane Harvey in 2017 near Houston, Texas. The study used SAR-based

indices Normalized Difference Flood Index (NDFI), Ratio Image (RI), and Difference Image

Index (DII) to delineate flood extents. The results were compared against MNDWI-based flood

extents. For Hurricane Ida, the optimum threshold value for DII was obtained at 0.6003 at k_{DII}

1.5, yielding agreement and total agreement percentages of 80% and 66%, respectively.

Similarly, for Hurricane Harvey, the optimum threshold value for DII was obtained at 0.6480 at

k_{DII} 0.8, with the agreement and total agreement also reaching 80% and 63%, respectively.

Based on the study area and environmental characteristics, the value of k_{DII} can vary compared

to the value of 1.5 suggested often in literature (Cian et al., 2018; Hamidi et al., 2023). A study

by Baig et al. (2013) showed that the reported value of the threshold for DFI and MNDWI was

0.3098, while 0.41 was chosen as the threshold for effective classification for the Lower Basin of

the Indus River. In a study conducted in the Barotse floodplain in Zambezi River by Zimba et al.

(2018), DFI thresholds between 0.3 and 0.7 have been suggested.

The flood mapping results indicate that an area of 25,229 km^2 was submerged during August,

and 19,181 km^2 during the initial 19 days of September as determined by the DFI Index.

Meanwhile, the MNDWI Index identified 30,582 km^2 of inundation in August and 21,851 km^2 in

the same period of September. These findings were then validated against external sources,

including reports from the Food and Agriculture Organization (FAO, 2022) and research by

Qamer et al. (2023), which documented an inundation of approximately 25,000 km^2,

approximately 18% of the total area of Sindh. Similarly, Wang et al. (2023) noted 20,912 km^2

flooded in their study of Sindh. They are presented in Table 2-7Table 2-7. Notably, variations in

the data range considered in these studies from August to September did not lead to differences in the total reported inundation area.

Table 2-7. Results comparison with other sources

Sources	Flooded area (km²)	Data acquisition dates
This study (Case 1)	25,229	Aug 1-31
This study (Case2)	30,582	Aug 1-31
Qamer et al. (2023)	25,000	Aug 22 - Sept 3
FAO (2023)	25,440	Aug 1-31
Wang et al. (2023)	20,912	Aug 25 - Sept 3

This study contributes to flood mapping efforts by highlighting the advantages of integrating SAR and optical data available publicly. The effectiveness of using Sentinel-2 optical data in flood extent mapping is significantly compromised by persistent cloud cover during flooding periods. This limitation necessitated the exclusion of cloud-obscured areas, restricting the analysis to smaller, clear regions which may not accurately reflect wider flood impacts. Additionally, the fixed revisit times of both Sentinel-1 and Sentinel-2 satellites introduce further

complications; the intervals between image acquisitions can miss critical flood development

stages, potentially leading to under or overestimations of flood extents. This temporal gap affects

the consistency and reliability of flood mapping, particularly at the edges of images where flood

conditions are most likely to change between passes. Using an exhaustive grid method allows for

the adjustment of the MNDWI value to be more suitable for regional specificities, thereby

enhancing the precision of flood extent maps. A universal threshold for flood detection is not

suitable globally as it fails to account for geographical and environmental variability, often

leading to inaccurate flood boundary delineations and severity assessments.

2.6 Conclusion

The accurate delineation of flood extents plays a crucial role in the efficient management of

disasters and the implementation of mitigation strategies. Advances in remote sensing

technology and methodologies have demonstrated their value in disaster-related applications,

including the identification, monitoring, and assessment of flood events. Due to advancements in

technologies, high spatial resolution, and frequent temporal coverage remote sensing data are

available. The utility of Synthetic Aperture Radar (SAR) data for timely flood mapping has been

explored, highlighting its advantage over optical sensors by providing reliable data even under

conditions of cloud cover or inclement weather. The processing and analysis of these datasets

have been facilitated by Google Earth Engine, a cloud-based platform. This study underscores

the critical role of integrating Synthetic Aperture Radar (SAR) and optical satellite data in

enhancing flood extent mapping, particularly in the Sindh province of Pakistan. Through the

optimization of thresholds for the Difference Image Index (DII), Desert Flood Index (DFI), and

Modified Normalized Difference Water Index (MNDWI), the accuracy and reliability of flood

mapping efforts were assessed. Based on the total agreement percentages observed between the

sets of indices, DII performed better than MNDWI for the Sindh province. The DFI index yielded a closer result to previously published studies compared to the MNDWI Index.

The study's methodology employed an exhaustive grid search for threshold optimization for flood detection and analysis. The findings reveal the variability of optimal threshold values depending on specific environmental conditions and study area characteristics, thereby emphasizing the importance of a nuanced approach to flood classification. The comparative analysis of SAR-based flood extents identified by optical-based DFI and MNDWI Indices with external sources confirms the validity of the study's results and particularly the efficacy of the DFI index in open water detection in arid regions like Pakistan.

This study contributes valuable insights into the integration of SAR and optical data for flood mapping, highlighting the potential of cloud-based platforms like GEE for processing and analyzing large geospatial datasets efficiently. The study reaffirms the critical need for accurate flood mapping in regions prone to extreme weather events and demonstrates the potential for methodological innovations that can enhance the precision and reliability of such efforts. As climate change continues to exacerbate the frequency and severity of flooding events, the findings of this study underscore the importance of continuous advancements in remote sensing technologies and methodologies. Such advancements are vital for informing and guiding disaster management strategies, aiding in the mitigation of flood impacts on vulnerable communities and ecosystems.

2.7 References

Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., & Ruello, G. (2018). Unsupervised
    Rapid Flood Mapping Using Sentinel-1 GRD SAR Images. *IEEE Transactions on*

Geoscience and Remote Sensing, 56(6). https://doi.org/10.1109/TGRS.2018.2797536

Baig, M. H. A., Zhang, L., Wang, S., Jiang, G., Lu, S., & Tong, Q. (2013). COmparison of
MNDWI and DFI for water mapping in flooding season. *International Geoscience and
Remote Sensing Symposium (IGARSS).* https://doi.org/10.1109/IGARSS.2013.6723425

Cian, F., Marconcini, M., & Ceccato, P. (2018). Normalized Difference Flood Index for
rapid flood mapping: Taking advantage of EO big data. *Remote Sensing of
Environment, 209.* https://doi.org/10.1016/j.rse.2018.03.006

Clement, M. A., Kilsby, C. G., & Moore, P. (2018). Multi-temporal synthetic aperture radar
flood mapping using change detection. *Journal of Flood Risk Management, 11(2), 152–
168.* https://doi.org/10.1111/jfr3.12303

Dasari, K., Anjaneyulu, L., Jayasri, P. V., & Prasad, A. V. V. (2016). Importance of speckle
filtering in image classification of SAR data. *2015 International Conference on
Microwave, Optical and Communication Engineering, ICMOCE 2015.*
https://doi.org/10.1109/ICMOCE.2015.7489764

Giustarini, L., Hostache, R., Matgen, P., Schumann, G. J. P., Bates, P. D., & Mason, D. C.
(2013). A change detection approach to flood mapping in Urban areas using TerraSAR-
X. *IEEE Transactions on Geoscience and Remote Sensing, 51(4).*
https://doi.org/10.1109/TGRS.2012.2210901

Gupta, P. K., Dubey, A. K., & Pradhan, R. (2022). Pakistan Flood of 2022: Assessment
using suite of satellite sensors and hydrological modelling.
https://doi.org/10.21203/rs.3.rs-2257969/v1

Hamidi, E., Peter, B. G., Munoz, D. F., Moftakhari, H., & Moradkhani, H. (2023). Fast
Flood Extent Monitoring With SAR Change Detection Using Google Earth Engine.
IEEE Transactions on Geoscience and Remote Sensing, 61.
https://doi.org/10.1109/TGRS.2023.3240097

Kumar, H., Karwariya, S. K., & Kumar, R. (2022). Google Earth Engine-Based
Identification of Flood Extent and Flood-Affected Paddy Rice Fields Using Sentinel-2
MSI and Sentinel-1 SAR Data in Bihar State, India. *Journal of the Indian Society of
Remote Sensing, 50(5).* https://doi.org/10.1007/s12524-021-01487-3

Liang, J., & Liu, D. (2020). A local thresholding approach to flood water delineation using
Sentinel-1 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 159.*
https://doi.org/10.1016/j.isprsjprs.2019.10.017

Long, S., Fatoyinbo, T. E., & Policelli, F. (2014). Flood extent mapping for Namibia using
change detection and thresholding with SAR. *Environmental Research Letters, 9(3),
035002.* https://doi.org/10.1088/1748-9326/9/3/035002

Nasirzadehdizaji, R., Akyuz, D. E., & Cakir, Z. (2019). Flood mapping and permanent water
bodies change detection using sentinel sar data. International Archives of the
Photogrammetry, *Remote Sensing and Spatial Information Sciences - ISPRS Archives,
42(4/W18), 797–801.* https://doi.org/10.5194/isprs-archives-XLII-4-W18-797-2019

Qamer, F. M., Ahmad, B., Hussain, A., Salman, A., Muhammad, S., Nawaz, M., Shrestha,
S., Iqbal, B., & Thapa, S. (2022). The 2022 Pakistan floods: Assessment of crop losses
in Sindh Province using satellite data. https://doi.org/10.53055/ICIMOD.1015

Shastry, A., Carter, E., Coltin, B., Sleeter, R., McMichael, S., & Eggleston, J. (2023).
Mapping floods from remote sensing data and quantifying the effects of surface
obstruction by clouds and vegetation. *Remote Sensing of Environment, 291*.
https://doi.org/10.1016/j.rse.2023.113556

Sunar, A. F., Yagmur, N., & Dervisoglu, A. (2019). Flood analysis with remote sensing data
– A case study: Maritsa River, Edirne. International Archives of the
Photogrammetry, *Remote Sensing and Spatial Information Sciences - ISPRS Archives,
42(3/W8).* https://doi.org/10.5194/isprs-archives-XLII-3-W8-497-2019

Tazmul Islam, M., & Meng, Q. (2022). An exploratory study of Sentinel-1 SAR for rapid
urban flood mapping on Google Earth Engine. *International Journal of Applied Earth
Observation and Geoinformation, 113, 103002.*
https://doi.org/10.1016/J.JAG.2022.103002

Ticehurst, C. J., Dyce, P., & Guerschman, J. P. (2009). Using passive microwave and
optical remote sensing to monitor flood inundation in support of hydrologic modelling.
*18th World IMACS Congress and MODSIM 2009 - International Congress on
Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical
and Computational Sciences, Proceedings*.

Venkatappa, M., Sasaki, N., Han, P., & Abe, I. (2021). Impacts of droughts and floods on
croplands and crop production in Southeast Asia – An application of Google Earth
Engine. *Science of the Total Environment, 795.*
https://doi.org/10.1016/j.scitotenv.2021.148829

Wang, J., Li, K., Hao, L., Xu, C., Liu, J., Qu, Z., Yan, X., Sajjad, M. M., & Sun, Y. (2023).

    Disaster mapping and assessment of Pakistan's 2022 mega-flood based on multi-source

    data-driven approach. Natural Hazards. https://doi.org/10.1007/s11069-023-06337-8

Zhang, F., Zhu, X., & Liu, D. (2014). Blending MODIS and Landsat images for urban flood

    mapping. *International Journal of Remote Sensing, 35(9).*

    https://doi.org/10.1080/01431161.2014.903351

Zhang, J., Liu, K., & Wang, M. (2023). Flood detection using Gravity Recovery and

    Climate Experiment (GRACE) terrestrial water storage and extreme precipitation data.

    Earth System Science Data, 15(2). https://doi.org/10.5194/essd-15-521-2023

Chapter 3 - Evaluating the Extent of Flood on Various Land Cover Classes Using Remote Sensing-Based Flood Maps and Land Use/Land Cover (LULC) Dataset.

Abstract

Disaster damage assessment is a crucial phase in emergency management, involving activities to evaluate the impact, scope, and specific needs following a disaster. It is more relevant to countries with regions like Sindh which are frequently troubled by floods, due to its geographic and climatic conditions. Assessment of damage can help authorities and relief organizations to make informed decisions regarding efficient resource allocation and distribution and make policies regarding focused mitigation practices. This study investigates the flooding impacts on different land cover classes in Pakistan's Sindh province, using remote sensing flood maps and Land Use/Land Cover (LULC) datasets. Unsupervised K-means clustering and supervised Random Forest classification to Sentinel-2 imagery are implemented to analyze flood-affected spatial distributions across croplands, urban areas, and forests. The study demonstrates the efficacy of Random Forest classification over K-means clustering in accurately delineating LULC classes. The study further revealed the vulnerability of built-up areas, irrigated crops, and sparse shrubs to flooding, especially in the Kambar Shahdad Kot, Jacobabad, and Larkana districts.

Keywords: Flooding extent, Land use/Land cover, LULC, K-means clustering, Random Forest

3.1 Introduction

Remote sensing is a critical tool in flood studies, providing essential data for flood inventory, monitoring, mapping, and disaster response. High-resolution satellite imagery is employed to assess the extent of flooding, identify affected regions, and comprehend the dynamics of

65

floodwaters over time (Bioresita et al., 2018). Remote sensing technologies such as Synthetic

Aperture Radar (SAR) offer all-weather, day-and-night imaging capabilities, enabling

continuous flood monitoring even in adverse weather conditions, and are dependable for flood-

related studies (Irwin et al., 2018; Kitajima et al., 2021; White et al., 2015). Numerous studies

have emphasized the significance of remote sensing in flood analysis, particularly in mapping

and monitoring floods to support disaster response efforts (Cian et al., 2018; Hamidi et al., 2023;

Schumann, 2015). The comparative analysis of satellite images captured before and after a flood

event facilitates the examination of alterations in land cover, identification of inundated areas,

and the quantification of flood severity. Such maps can be integrated with other data such as the

Digital Elevation Model (DEM) and GIS techniques to improve the understanding of flood

dynamics (Martinis et al., 2013). Additionally, remote sensing technology aids in analyzing

changes in vegetation cover, influencing water dynamics and runoff patterns, thereby

contributing to assessing the impacts of floods (Bolanos et al., 2016). Remote sensing-based

investigations support urban planning, vulnerability assessment, the development of flood

mitigation strategies, and support development of resilient communities (Sun et al., 2022).

Assessing damage is crucial for understanding the impacts of disasters on public health and the

economy, playing a pivotal role in disaster management and policy formulation (Ghaffarian et

al., 2018). This process not only aids in coordinating efforts among various stakeholders but also

involves communities and offers invaluable insights for the development of response strategies.

Moreover, damage assessment is instrumental in identifying vulnerabilities and risks, enabling

targeted mitigation measures such as resilient infrastructure development and informed land-use

planning to reduce the likelihood and severity of future disasters (Chen et al., 2021). Damage

assessment provides vital information for resource allocation, planning, and improving the

efficiency of response efforts by strategic deployment based on identified needs (Nex et al., 2019). Damage assessment also plays a key role in identifying vulnerabilities and risks and facilitates targeted mitigation measures such as resilient infrastructure development and informed land-use planning to decrease the likelihood and severity of future disasters (Ghaffarian et al., 2018).

Land use and land cover (LULC) encapsulate the physical state of the land and the impact of human activities on these landscapes. They are key elements in driving environmental transformations at both local and global scales, impacting ecosystems, land stewardship, and water management strategies. Remote sensing, GIS, and field observations are integrated to generate LULC maps, aiding in understanding human-nature interactions. Machine learning algorithms, utilizing labeled training data and high-resolution satellite imagery, enhance land cover classification accuracy (Xie et al., 2019). The alterations in land cover are central to understanding climate change dynamics, highlighting the significant role of human interventions (Morin et al., 2021). Understanding LULC dynamics is, therefore, essential for developing comprehensive approaches to managing environmental resources, mitigating the effects of climate change, and ensuring sustainable development. Through the integration of remote sensing, geographic information systems (GIS), and field observations, LULC maps can be generated, facilitating an understanding of the intricate interactions between human actions and the natural environment. Moreover, machine learning algorithms, leveraging labeled training data that represent different land cover classes and employing multi-temporal, high-resolution satellite imagery, enhance the accuracy of land cover classification (Bui & Mucsi, 2021). Land use and land cover (LULC) are subject to dynamic changes influenced by both natural events and human activities. These alterations, stemming from both natural causes and various human

67

interventions, can provide significant environmental challenges and hazards. Monitoring transformations in land use and land cover (LULC) is essential for understanding potential concerns. This surveillance enables researchers and policymakers to understand the intricate relationships between human actions and the natural world, tracking these changes over time to facilitate better-informed decisions (Tzepkenlis et al., 2023).

Unsupervised classification in remote sensing is a powerful technique for systematically categorizing and grouping pixels into distinct classes or clusters based on similar spectral characteristics without prior knowledge of the land classes. Detailed class information is not a necessity in unsupervised classification which is advantageous in exploratory studies. This feature is particularly beneficial in remote sensing applications where the area of interest is inaccessible or has not been extensively studied. Machine learning techniques play a vital role in enhancing the efficiency and accuracy of unsupervised classification in remote sensing. Using algorithms such as k-means clustering, hierarchical clustering, and Principal Component Analysis (PCA), machine learning enables the segmentation of the image into meaningful clusters without manual intervention (Hu et al., 2015). Advanced machine learning and deep learning models can handle high-dimensional data leading to more accurate classification outcomes.

While unsupervised classification offers considerable advantages, it also encounters obstacles, including the potential for subjective interpretation of classes. This issue is particularly evident in algorithms such as K-Means, where the results can be significantly influenced by how the algorithm is initialized, thereby affecting the overall classification accuracy (Sonobe, 2019). Advanced machine learning models have been applied to remote sensing tasks, such as scene classification and crop mapping, demonstrating the effectiveness of transfer learning and self-

supervised learning approaches (Berg et al., 2022; de Lima & Marfurt, 2020). Another challenge

arises in dealing with mixed pixels, which occur when a single spatial unit displays

characteristics of multiple land cover types, complicating the classification process for

unsupervised methods. Hence, knowledge of the field is extremely important to get accurate

results in such techniques.

One of the most used unsupervised machine learning algorithms is K-means clustering. In

remote sensing, this algorithm categorizes the pixels based on their spectral similarities by

partitioning the dataset into K number of pre-determined non-overlapping clusters by minimizing

within-cluster variance while maximizing variance between clusters (Li et al., 2014). Pixels are

allocated to the closest cluster centroid by measuring the Euclidean distance, followed by a

recalculation of the centroids. The solution is produced when the reassignment of data does not

change the result with iterations. In remote sensing, unsupervised clustering algorithms like K-

Means are essential for processing and interpreting remote sensing data efficiently and accurately

and facilitate tasks such as land cover classification, flood mapping, and vegetation monitoring

(Sublime & Kalinicheva, 2019; Weinstein et al., 2019). The performance of K-means clustering

depends heavily on the initial seed values chosen and specifying the number of cluster centers

before cluster identification.

With unsupervised classification techniques like K-means clustering, the optimal number of

clusters is not predetermined and can significantly impact the results and interpretability of the

analysis (Langford et al., 2019). Tools like the elbow method are helpful in this context which

calculates the within-cluster sum of squares (WCSS) for each configuration and an optimal

number of clusters is determined where the rate of WCSS decreases sharply. Selecting a cluster

number in the optimal range effectively balances between too few clusters (underfitting) and too

many clusters (overfitting), ensuring the most efficient representation of the data's structure without unnecessary complexity (Li et al., 2016). This ensures the balance between fitting the data well and avoiding overfitting (Amatya et al., 2019). The elbow method has been applied in remote sensing and environmental science studies to estimate the number of clusters (Huang et al., 2016; Karim & Zhang, 2021; Santaga et al., 2021; Scarth et al., 2019; Viana-Soto et al., 2020).

Supervised classification in remote sensing is essential for various applications such as creating land cover maps, monitoring vegetation, and detecting land use changes. Supervised classification relies on labeled datasets to accurately categorize land cover and land use. This method involves collecting representative samples from different land cover classes, extracting features like spectral and spatial attributes, and training machine learning algorithms (Yang et al., 2021). Supervised classification algorithms utilize a dataset with predefined labels to understand the correlation between features and their corresponding class labels during the training phase. After the model has been adequately trained, it is then employed to categorize pixels throughout the image, attributing a class label to each pixel. The accuracy of supervised classification is crucial and is typically assessed by comparing the results with independent validation data or ground truth information (Yang et al., 2021).

Random Forest (RF) is an ensemble classifier particularly helpful for Land use/ land cover (LULC) mapping. The effectiveness of RF classifiers has been well-studied in vegetation mapping and land cover classification (Saini & Ghosh, 2018; Wulder et al., 2018). RF classifier has demonstrated robust and superior performance in comparison to other machine learning classifiers such as Support Vector Machine (SVM), k-nearest Neighbors (KNN), and Decision Trees (DT) for remote sensing image classification. RF is an ensemble classifier that generates

multiple decision trees constructed from random segments of the training data and its attributes avert overfitting and enhance variability among the trees to collectively make more accurate predictions. RF can work with complex datasets and extract significant insights for precise land cover classification (Mellor et al., 2013). Random Forest has been successfully applied in various remote sensing scenarios, including crop classification using optical and radar images (Hütt et al., 2016).

Accurately mapping flood extents and assessing damage using remote sensing is crucial, especially in areas like the Sindh province of Pakistan, which are prone to severe and recurrent floods. Land Use/Land Cover (LULC) classification images combined with flood maps from Sentinel-1 SAR imagery improve disaster management and policymaking. The integration of publicly accessible global remotely sensed LULC datasets, such as Dynamic World V1, ESA WorldCover 10m V100, and the Copernicus Global Land Cover Layers, is instrumental for accurately assessing flood impacts. By detailing the specific types of land use and cover impacted by floods, authorities can better assess the effects on critical infrastructure, agriculture, and residential areas. This detailed damage assessment informs targeted recovery and rehabilitation efforts and supports the development of resilient infrastructure and land use planning strategies. However, the need to update these datasets annually to maintain accuracy and relevancy necessitates further exploration into advanced techniques to enhance the precision of classification images and overall damage assessment practices. Despite the existence of studies utilizing these advanced techniques, there is a significant gap in fully leveraging them in flood studies.

The province of Sindh, located in the southeastern part of Pakistan, is particularly susceptible to floods due to its geographical location and climatic conditions. Due to high intraseasonal

variability in the region, Sindh receives significant rainfall between July and September each year, making it vulnerable to calamities such as floods and droughts. This study aims to evaluate the impact of flooding on various land cover classes within Sindh, utilizing remote sensing-based flood maps and Land Use/Land Cover (LULC) datasets. By analyzing the spatial distribution of flood-affected areas across different land classes detected by Sentinel-1 Synthetic Aperture Radar (SAR) imagery, this study seeks to understand how the extent of flooding varies among different land classes, such as croplands, urban areas, and forests, during a flood event. The objective of the study is to evaluate the extent of flood impact on various land cover classes using remote sensing-based flood maps and the Land Use/Land Cover (LULC) dataset in the GIS environment.

3.2 Study Area and Data

Located in the southern part of Pakistan, Sindh province borders the Arabian Sea to the south, Balochistan to the north and west, Punjab to the northeast, and India to the east as shown in Figure 3-1. Spanning about 141,000 km^2, Sindh constitutes about 17.7% of Pakistan's total area. Positioned within a zone that transitions from tropical to subtropical, Sindh experiences intensely hot summers, with temperatures often soaring above 46°C from May through August, while the winters are relatively short, with temperatures averaging around 2°C in December and January. The province's primary watercourse is the Indus River, fed by its tributaries like the Jhelum, Chenab, Ravi, and Sutlej, which traverse through other provinces before converging in Sindh. Sindh province has a semi-arid climate, but it features diverse natural features such as coastal and riverine forests, vast freshwater lakes, and varied landscapes of mountains and deserts. The province is predominantly arid and has sparse vegetation, except the areas along the Indus Valley that are irrigated.
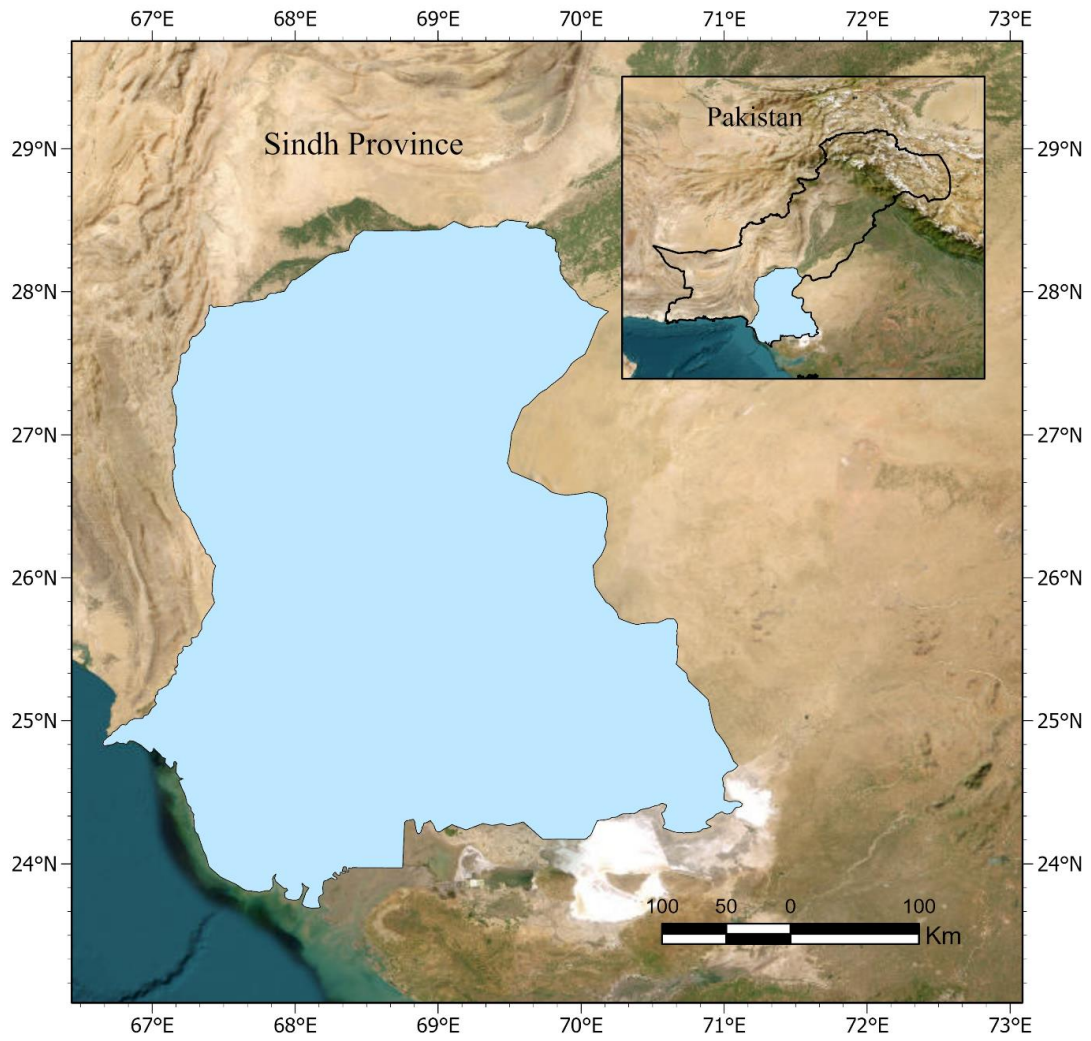
Figure 3-1. Map of the Study area: Sindh province of Pakistan.

The European Union's Copernicus Sentinel constellation delivers data through both optical (such as Sentinel-2) and synthetic aperture radar (SAR, like Sentinel-1) technologies. Sentinel-2 mission has a pair of Satellites, Sentinel-2A, and Sentinel-2B with Multispectral Instrument (MSI) capable of capturing imagery over 13 spectral bands spanning from visible to shortwave infrared spectrum (Helber et al., 2019). The dataset 'COPERNICUS/S2_SR' from the Sentinel-2 mission, a Level-2A product, is accessible through the Google Earth Engine (GEE) platform and has been used in this study for land use and land cover (LULC) analysis. The dataset offers high spatial resolution, ranging from 10 to 60 meters, and frequent observation intervals of every 5 days (Silva-Cardoza et al., 2022). Land cover (Sindh Province, Pakistan - 10m) is the dataset released by the FAO Geospatial Team NSL and the FAO GIS team in Pakistan and has been used as ground data in this study. Land cover data offers a detailed account of the biological and non-biological resources within a province, covering a wide range of classifications such as farmed areas, natural plant life, barren landscapes, and urban territories. It primarily focuses on the geographical characteristics of a province. Google Earth Engine has been utilized to carry out both supervised and unsupervised classification of Sentinel-2 data for the province of Sindh, to acquire the Land Use and Land Cover (LULC) dataset. The post-processing of the results and statistical analyses are carried out in ArcGIS Pro.

3.3 Methodology

This section of the thesis describes the methodological framework utilized to evaluate the spatial distribution and extent of flooding on various land cover classes in the administrative districts of Sindh province. The study leverages remote sensing-based Land Use/Land Cover (LULC) datasets and flood maps derived from Sentinel-1 Synthetic Aperture Radar (SAR) imagery. The approach involves creating LULC maps via both unsupervised and supervised classification

methods and comparative analysis to select the most accurate LULC map. This map is then

integrated with flood extent and administrative district maps to assess the flood impact. The

flowchart of the methodology is shown in Figure 3-2.
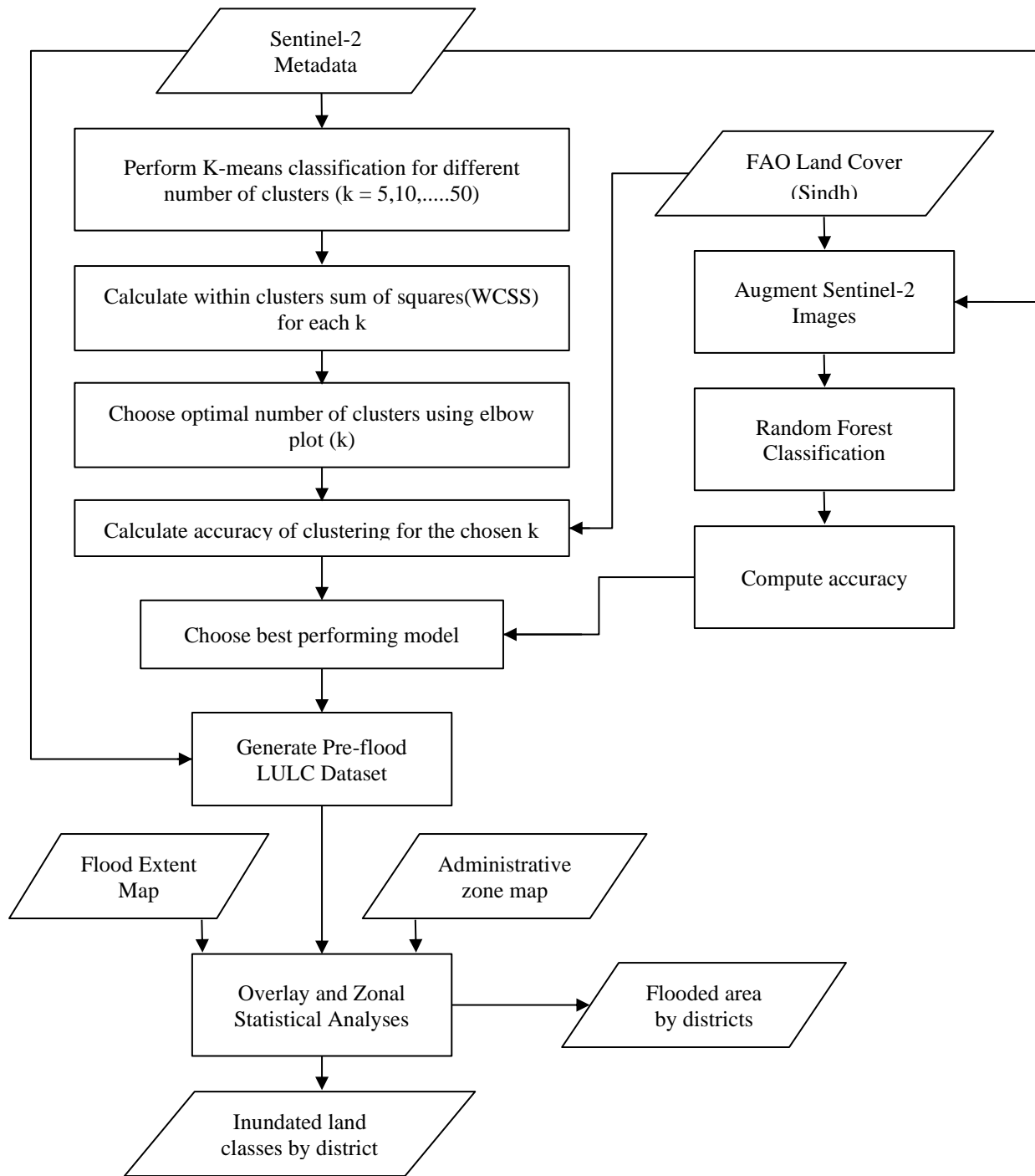
Figure 3-2. Flowchart diagram of the methodology.

3.3.1   Data Acquisition

To develop a Land Use and Land Cover (LULC) Dataset, the Sentinel-2 Level-2A dataset 'COPERNICUS/S2_SR' was accessed through the Google Earth Engine. Sentinel-2 images collected in the initial half of May 2022, directly preceding the flooding period were filtered. For the unsupervised classification, the bands B1-B12 were taken in the median composite. The filtered images do not require further processing in Google Earth Engine. The study uses the Land Use/Land Cover (LULC) dataset for the whole period of the year 2021, obtained from the Food and Agriculture Organization (FAO) as the ground data, shown in Figure 3-3. The LULC dataset was ingested into the GEE as assets. Overlay and zonal statistical analyses were performed in ArcGIS Pro software. Table 3-1 presents the land cover classes identified within the FAO land cover classification - Sindh dataset.

Table 3-1. Labels of classes from Land Cover Map (FAO, 2023)

Class Value	Label
C_1	Built-up (BUP)
C_2	Roads (ROA)
C_3	Tree Orchards (ORC)
C_4	Herbaceous crops irrigated (HCI)
C_5	Herbaceous crops rainfed (HCR)
C_6	Herbaceous crops in flood plain (HCF)
C_7	Herbaceous natural vegetation (HER)
C_8	Shrubs sparse natural vegetation (SOP)
C_9	Shrubs dense natural vegetation (SCL)
C_10	Shrubs in temporary wet soil (SWE)
C_11	Trees sparse natural vegetation (TOP)
C_12	Tree forest plantations (TFP)
C_13	Trees dense natural vegetation (TCL)

C_14	Sand dunes (SDU)
C_15	Bare soil (BSO)
C_16	Bare soil in temporary wet (BTW)
C_17	Water bodies (WBO)
C_18	Wetlands (WET)
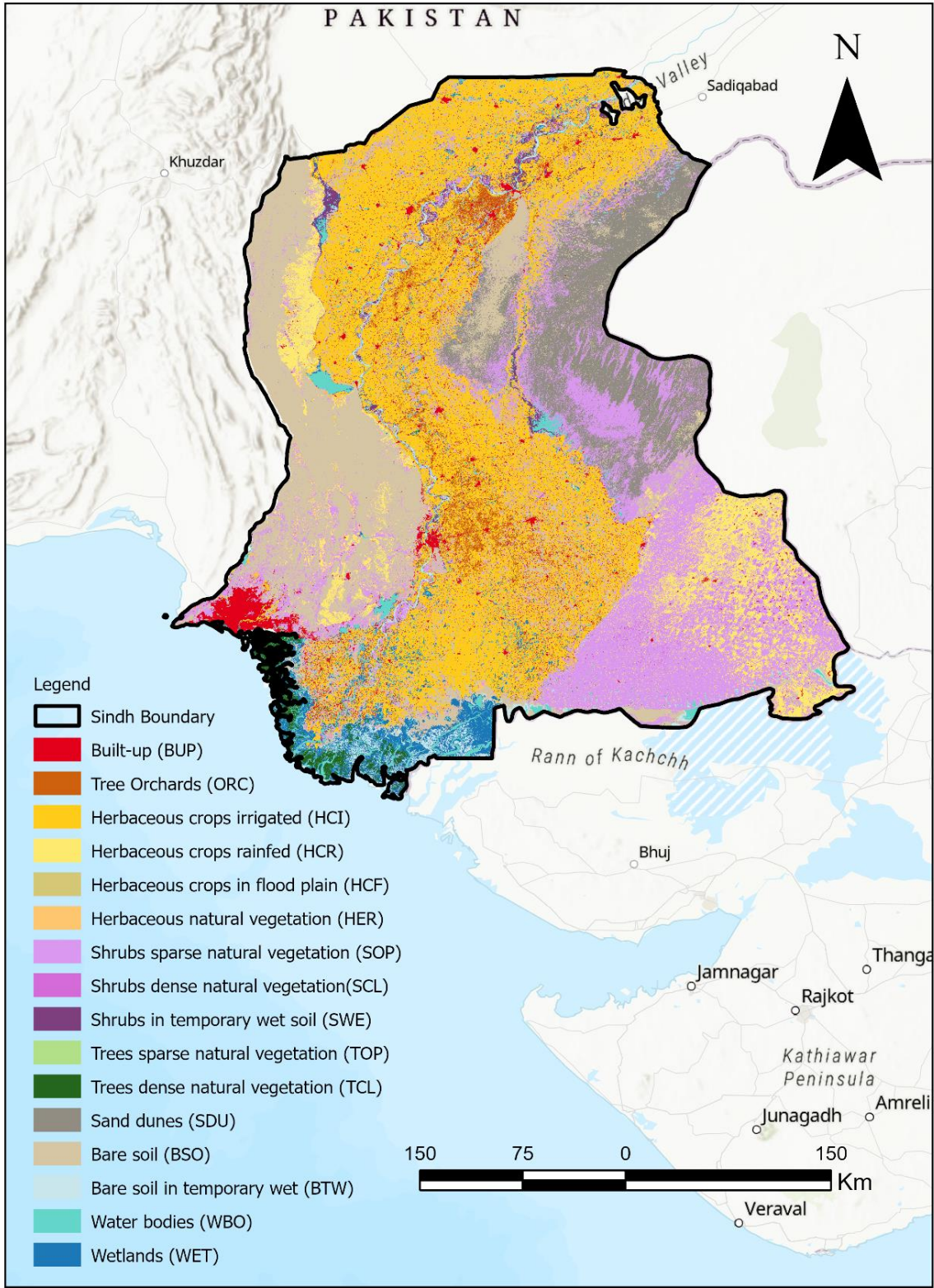
Figure 3-3. Land Cover Map for Sindh (FAO, 2023).

3.3.2 LULC Dataset Development

LULC Dataset has been developed by supervised and unsupervised classification methods. In the unsupervised classification, the K-means clustering algorithm was performed on the mosaic composite of Sentinel-2 images. In unsupervised classification approaches, determining the optimal number of clusters is crucial as it directly influences the outcomes and the clarity of the analysis. The Elbow plot method was implemented which computes the within-cluster sum of squares (WCSS) across various configurations, identifying the optimal cluster count. Supervised classification was conducted using the Random Forest (RF) method, known for its high accuracy and robustness in dealing with non-linear data. This approach involved training the model on a subset of the data with known classifications and then applying it to classify the entire dataset into distinct land cover classes. The choice of initial seeds can markedly influence the choice of training samples as well as the results by determining the starting initial centroids of clusters, resulting in slightly different final cluster arrangements. Hence, multiple iterations were performed until the results. Hence, multiple iterations were performed until the change in seed value no longer significantly influenced the results in both Random Forest and K-means classification, thereby ensuring the stability and robustness of the models.

To evaluate the accuracy of the Land Use and Land Cover (LULC) maps produced by unsupervised (K-means) and supervised (Random Forest) classification techniques, a range of metrics are utilized. The Confusion Matrix plays a pivotal role as it illustrates the performance of a classification model against a dataset with known outcomes. It aligns the model's predictions against the actual classifications on orthogonal axes, where each entry represents the count of predictions for a predicted class against the actual class. This aids in the detailed assessment of the model's predictive accuracy and error rates. Accuracy serves as another fundamental metric,

defining the overall correctness of the model by the proportion of correct predictions made from total predictions, thus offering a direct measure of performance efficiency.

The kappa statistic, k enhances this analysis by evaluating the congruence between predictions and actual classifications, adjusted for random chance agreement. A k value of 1 denotes perfect congruence, whereas a value of 0 indicates no improvement over random chance, providing deeper insights into the model's predictive precision. Furthermore, user accuracy (UA) and producer accuracy (PA) assess the model's predictive reliability and completeness, respectively. UA, also known as precision, is the ratio of correct positive predictions to all positive predictions made. In contrast, PA, or recall, measures the proportion of correct positive predictions out of the actual positive cases.

3.3.3   Flood Extent Mapping

The flood extent mapping was developed using the Sentinel-1 SAR dataset using the Difference Image Index (DII) to define the flooded areas. A change detection algorithm was used to analyze the indices derived from the SAR dataset with the Sentinel-2 dataset. A sensitivity analysis was conducted to evaluate the threshold values for each index so that there is maximum agreement between the results in pixel levels. This analysis determined the optimal threshold values for each index, crucial for achieving maximum alignment of the results. Utilizing these threshold values, a comprehensive flood extent map was generated, delineating the areas affected by flooding in the Sindh province during August.

3.3.4   Overlay and Zonal Statistical Analyses

The LULC Dataset generated through unsupervised classification methods, is chosen based on

its high accuracy in comparison with the FAO land cover dataset. The LULC dataset is overlaid

with the Flood extent map and administrative district maps of the Sindh province. The overlay

analyses are carried out in the GIS environment which enabled the identification and

quantification of flood-affected areas within each land cover class across different administrative

districts. This analysis provided insights into the spatial distribution and magnitude of the

severity and distribution of flood impacts.

3.4 Results

3.4.1   K-means Clustering

The optimal number of clusters is not predetermined in unsupervised classification techniques

like K-means clustering. This optimal cluster number forms the foundation for subsequent

analyses, including the detailed examination of flood extents across different land cover classes

within the administrative districts of the study area. To select the optimal number of clusters, the

elbow method was used where the number of clusters was plotted against WCSS values with K-

means clustering. At the Elbow point, change in WCSS becomes negligible with an increase in

clusters, indicating that adding more clusters does not significantly improve the model's

performance. Initially, the WCSS values were plotted against the number of clusters ranging

from 5 to 50. The analysis revealed a gradual decline in WCSS as the number of clusters

increased, with notable inflection points observed at specific cluster values, as shown in Figure

3-4.

The WCSS was initially high at 1,270,852 for 5 clusters and decreased as more clusters were

added, reflecting improvements in cluster homogeneity. However, as the number of clusters

increased further to 15, 20, 25, and beyond, the rate of decrease in WCSS began to diminish. The

plot visually supports the principle, with the "elbow"; the point of inflection on the curve situated

between 10 and 20 clusters.
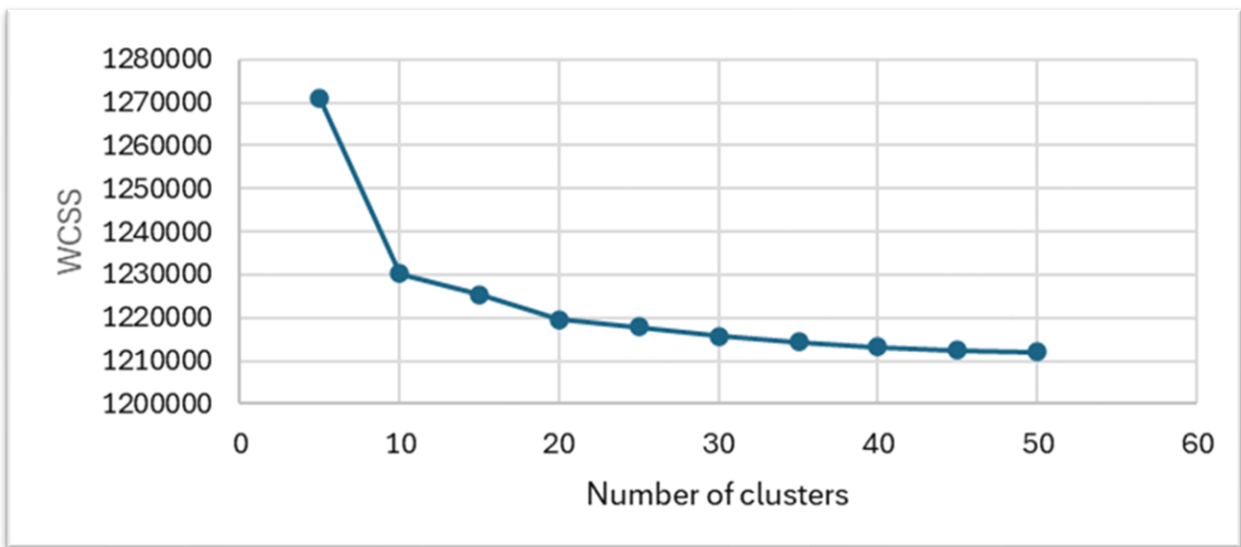
Figure 3-4. Elbow Plot Showing the Within-Cluster Sum of Squares (WCSS) against the Number

of Clusters (1).

To determine the optimal number of clusters, an elbow plot was refined to include additional

data points, representing the WCSS values for each cluster count ranging from 10 to 20, as

shown in Figure 3-5. The curve began to flatten after 10 clusters and continued to decrease at a

diminishing rate up to 20 clusters. Beyond 20 clusters, the reduction in WCSS became less

prominent, suggesting that the additional complexity of more clusters did not improve

homogeneity within clusters. The curve starts to level off after 10 clusters. Hence, clusters

numbers 10, 14, and 18 were chosen to proceed with the unsupervised clustering.
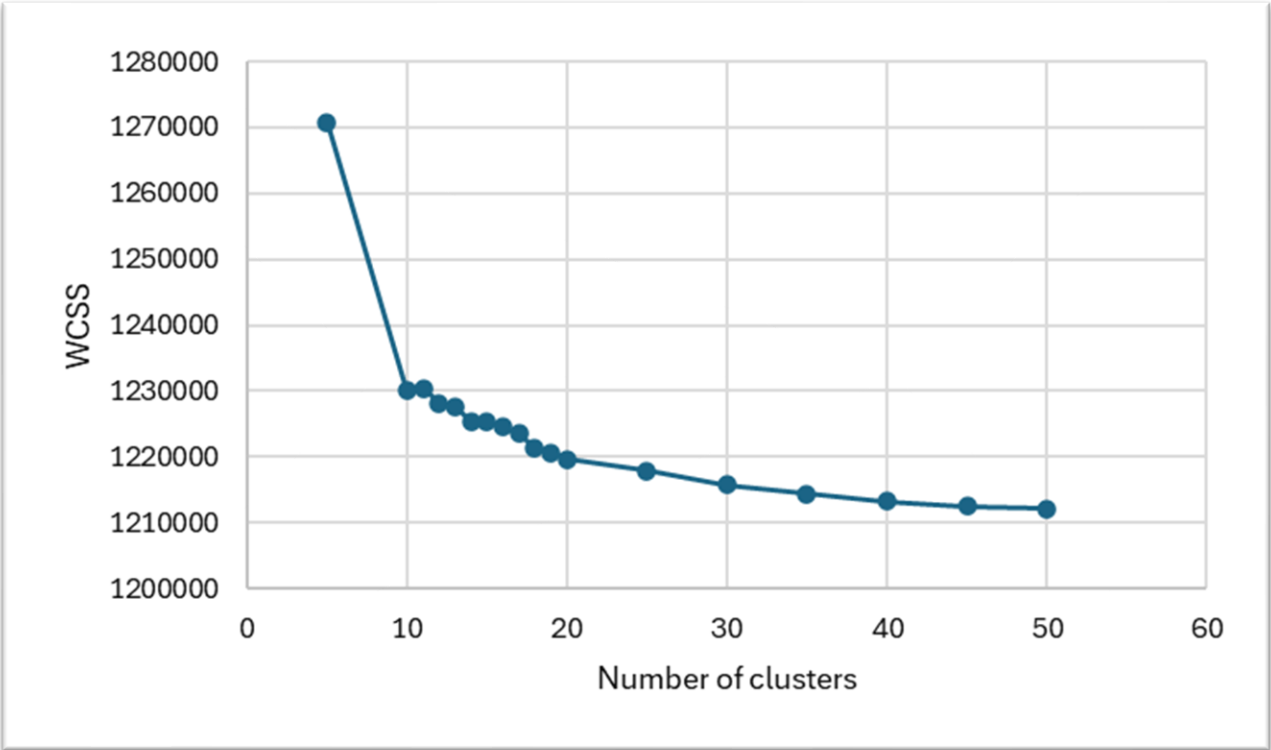
Figure 3-5. Elbow Plot Showing the Within-Cluster Sum of Squares (WCSS) against the Number

of Clusters (2).

3.4.1.1 Results of K-means Clustering with k=10

The K-means clustering algorithm was applied to the median composite of Sentinel-2 satellite images in GEE for each predetermined cluster count (k=10, 14, and 18) to identify the land use and land cover (LULC) patterns within the study area. Each of the clusters identified represents distinct land cover types. For each predetermined cluster count (k=10, k=14, and k=18), the k-means algorithm was iteratively executed with varying initial seed values. The derived clusters were labeled and reclassified based on the land use and land cover characteristics provided by the FAO-Pakistan. The accuracy of the land cover classifications for each cluster count was accessed through comparison with the FAO-Pakistan LULC dataset. The validity and accuracy of the clustered LULC classes for each k value (10, 14, and 18) were quantitatively assessed using confusion matrices. These matrices quantify the performance of clustering algorithms for selected cluster numbers.

Figure 3-6 presents the results obtained from applying the k-means clustering algorithm to Sentinel-2 satellite composite imagery, where the initial number of clusters (k) was set to 10. This figure illustrates the segmentation of the landscape into 10 distinct clusters, each representing a potential unique land use and land cover (LULC) category identified from the Sentinel-2 imagery. Subsequently, these clusters were labeled to correspond with specific LULC types for interpreting the segmented imagery meaningfully.
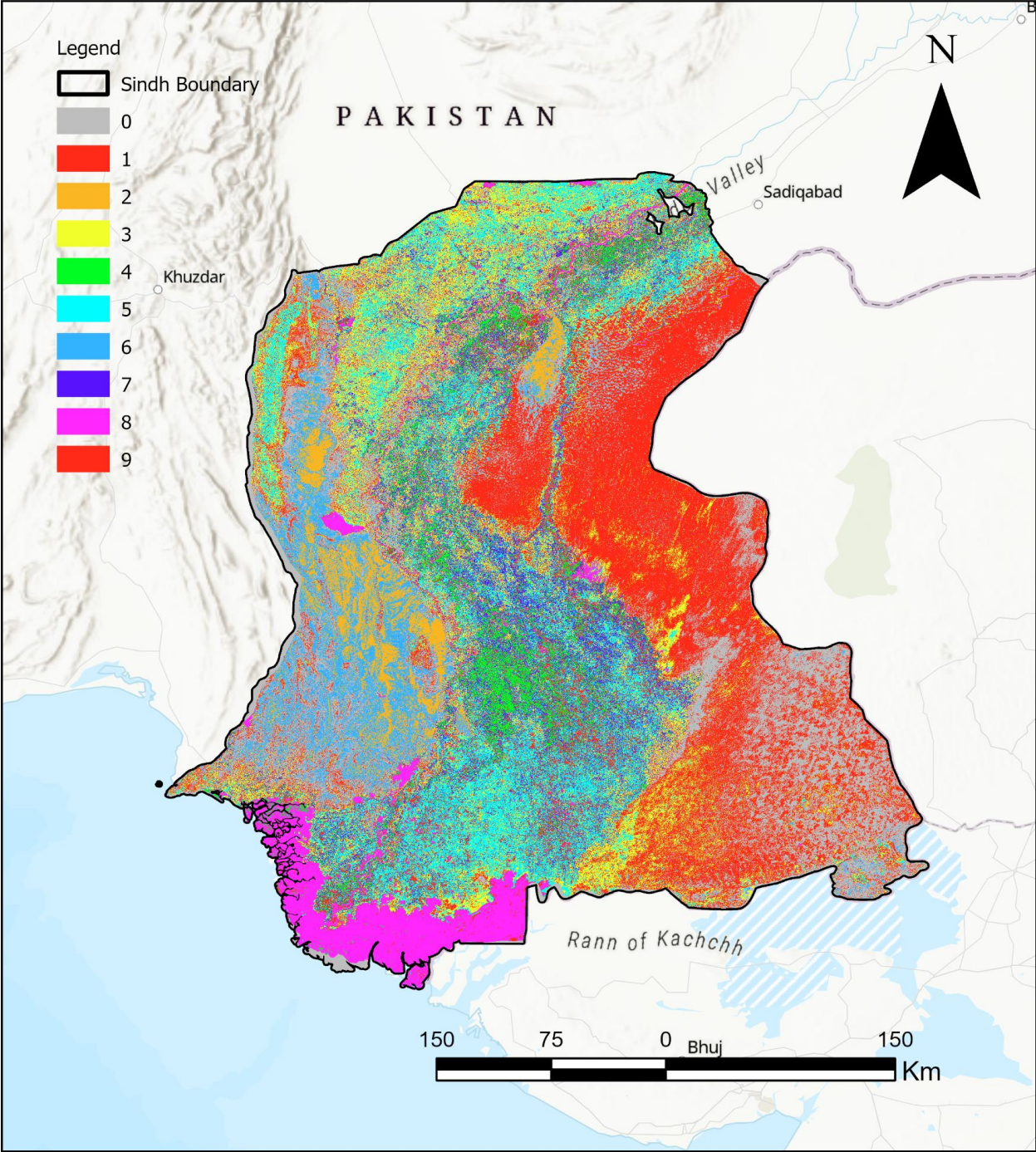
Figure 3-6. Output of K-means clustering using 10 clusters.

During the labeling stage, reclassification was needed for an accurate representation of the area's Land Use and Land Cover (LULC) distribution. This led to the merging of the original 10 clusters into 6 clusters. Land cover class C_1 which represented built-up areas was removed from the final classified map and replaced with city masks due to the classification's inability to adequately capture this land cover type and the resulting map has been presented in Figure 3-7. The performance of classification was then accessed through confusion matrices. Table 3-2 presents the outcomes derived from the confusion matrix analysis.

Table 3-2. Confusion matrix between K-means (k=10) derived LULC map with FAO LULC map

Class	C_4	C_5	C_14	C_15	C_17	Total	U Acc (%)	Kappa
C_4	256324	1881	67852	39525	4842	370424	69.2	0
C_5	16648	40495	36345	33895	736	128119	31.61	0
C_14	11779	13465	92103	39653	631	157631	58.43	0
C_15	3580	13598	9988	67717	286	95169	71.15	0
C_17	922	14	511	1198	16497	19142	86.18	0
Total	289253	69453	206799	181988	22992	770485	0	0
P Acc (%)	88.62	58.31	44.54	37.21	71.75	0	61.41	0
Kappa	0	0	0	0	0	0	0	0.4581

The confusion matrices revealed varying levels of accuracy across different LULC classes. Class C_4 had a user accuracy of 69.2%, while Classes C_5 and C_14 experienced significantly lower user accuracies of 31.61% and 58.43%, respectively. Conversely, Class C_15 displayed a relatively high user accuracy of 71.15%, indicating a better alignment with the ground truth. In contrast, Class C_17 exhibited the highest user accuracy of 86.18%, suggesting an excellent

agreement with the reference LULC dataset. Producer accuracy also varied significantly across classes, with Class C_4 achieving an accuracy of 88.62%, whereas Class C_14 had a lower producer accuracy of 44.54%. The kappa statistic of 0.4581 indicates a moderate agreement between the k-means classified output and the reference LULC dataset, after accounting for the removal of built-up areas and their replacement with city masks to address the classification's limitations. The overall accuracy of the classification was observed to be approximately 61.41%.
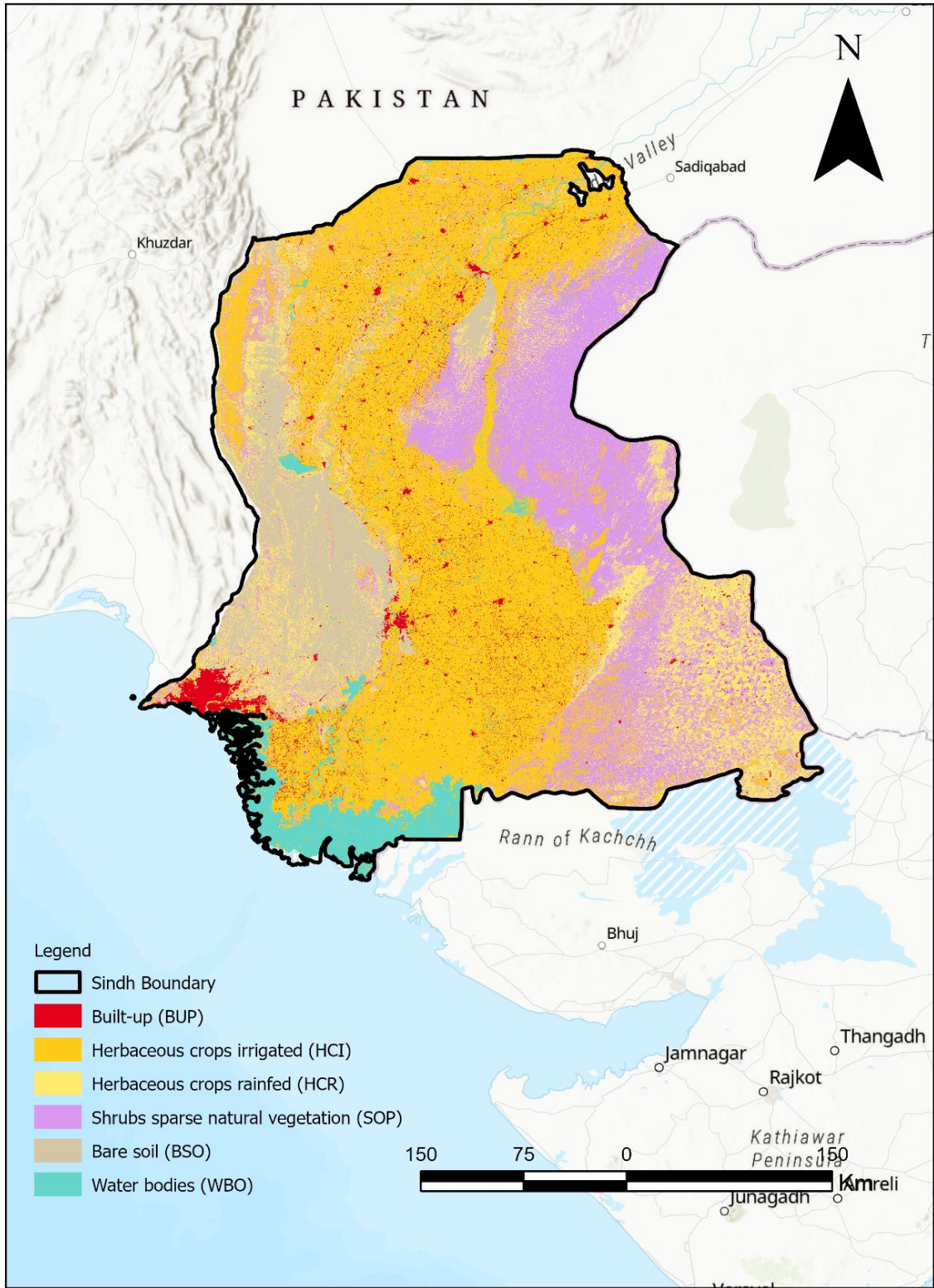
Figure 3-7. LULC map generated from K-means clustering using 10 clusters.

3.4.1.2 Results of K-means Clustering with k=14

The k-means clustering algorithm was initiated with 14 clusters on Sentinel-2 satellite composite imagery and initializing it with 14 clusters. Each cluster is suggested to correspond to a distinct category of land use and land cover (LULC), as identified from the Sentinel-2 dataset and presented in Figure 3-8. The clusters were labeled to match specific LULC types. During labeling the LULC distribution consolidated from 14 clusters to 8 meaningful labeled clusters. The land cover class C_1, indicative of urban or built-up regions, was substituted with city masks due to the classifier's inability to accurately depict this particular type of land cover. Consequently, class C_1 was omitted from the confusion matrix analysis. The final classification underwent an evaluation using confusion matrices, the results of which are detailed in Table 3-3. Figure 3-9 represents the revised labeled LULC map.
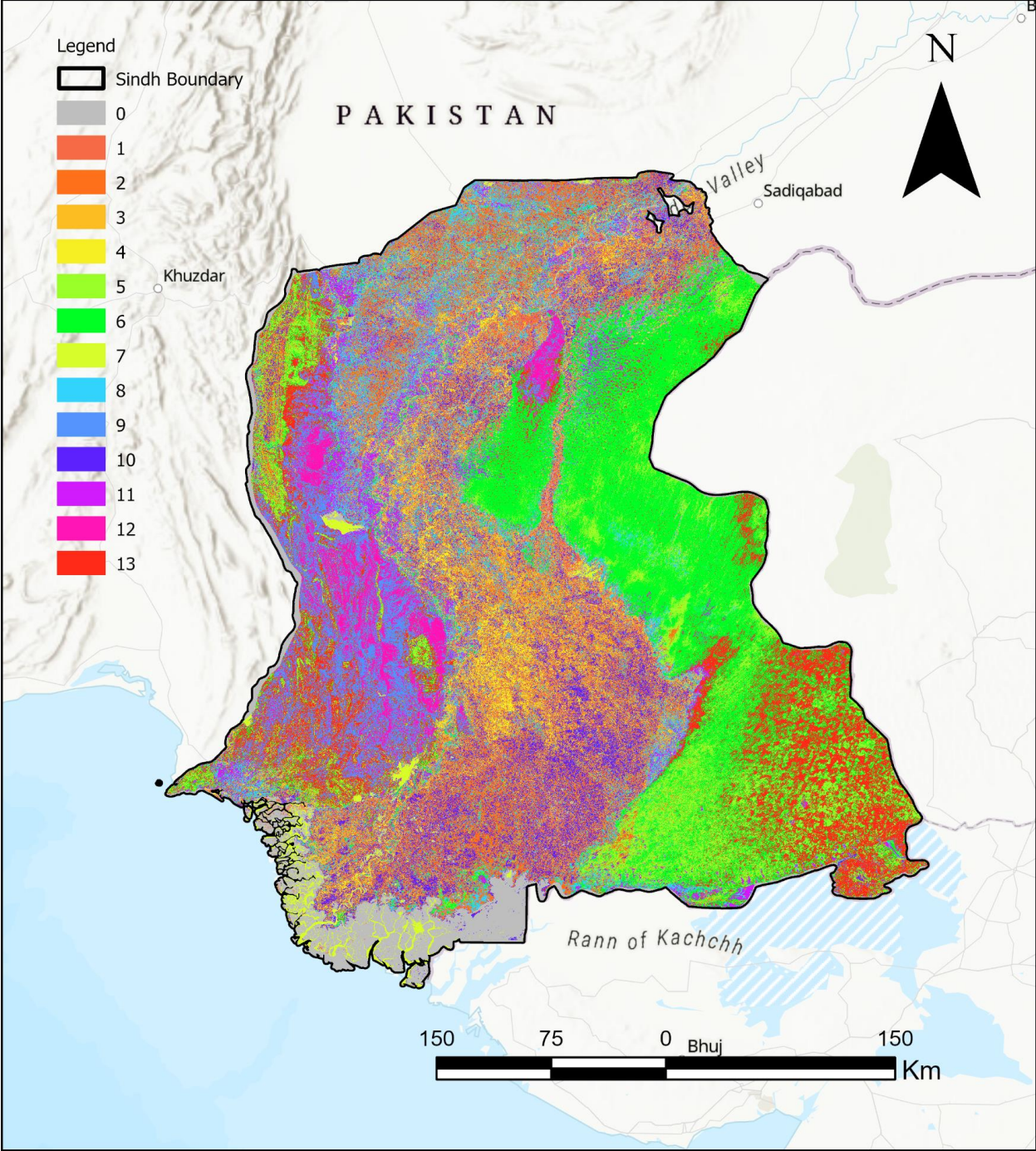
Figure 3-8. Output of K-means clustering using 14 clusters.

The analysis conducted via the confusion matrices showed varied accuracy levels across the different LULC classes. Class C_4, for example, had a user accuracy of 71.43%, while classes such as C_5 and C_14 displayed considerably lower user accuracies of 37.88% and 38.52%, respectively. On the other hand, Class C_15 and C_17 exhibited a notably higher user accuracy of 73.68% and 80.67%, suggesting a strong alignment with the reference LULC dataset. The producer accuracy across the classes displayed significant variability. Class C_4 achieved a high producer accuracy of 96.69%, which contrasts sharply with the much lower accuracies seen in classes like C_5 and C_14, with producer accuracies of 54.14% and 82.71%, respectively. The kappa statistic of 0.4878 indicates a moderate agreement between the k-means classified output and the reference LULC dataset. The classification model's performance was measured, revealing an overall accuracy rate of 59.95%.

Table 3-3. Confusion matrix between K-means (k=14) derived LULC map with FAO LULC map

Class	C_4	C_5	C_8	C_14	C_15	C_17	C_18	Total	U Acc (%)	Kappa
C_4	279456	2389	55131	1151	39952	5476	7669	391224	71.43	0
C_5	944	37593	28320	3344	28979	13	54	99247	37.88	0
C_8	2083	3782	50878	9922	16072	22	36	82795	61.45	0
C_14	2909	13729	63151	70087	31626	223	225	181950	38.52	0
C_15	2073	11928	8347	227	63769	98	102	86544	73.68	0
C_17	122	12	54	0	59	10401	2246	12894	80.67	0
C_18	1445	6	818	4	1488	6758	16090	26609	60.47	0
Total	289032	69439	206699	84735	181945	22991	26422	881263	0	0
P Acc (%)	96.69	54.14	24.61	82.71	35.05	45.24	60.9	0	0	0
Kappa	0	0	0	0	0	0	0	0	0	0.4878
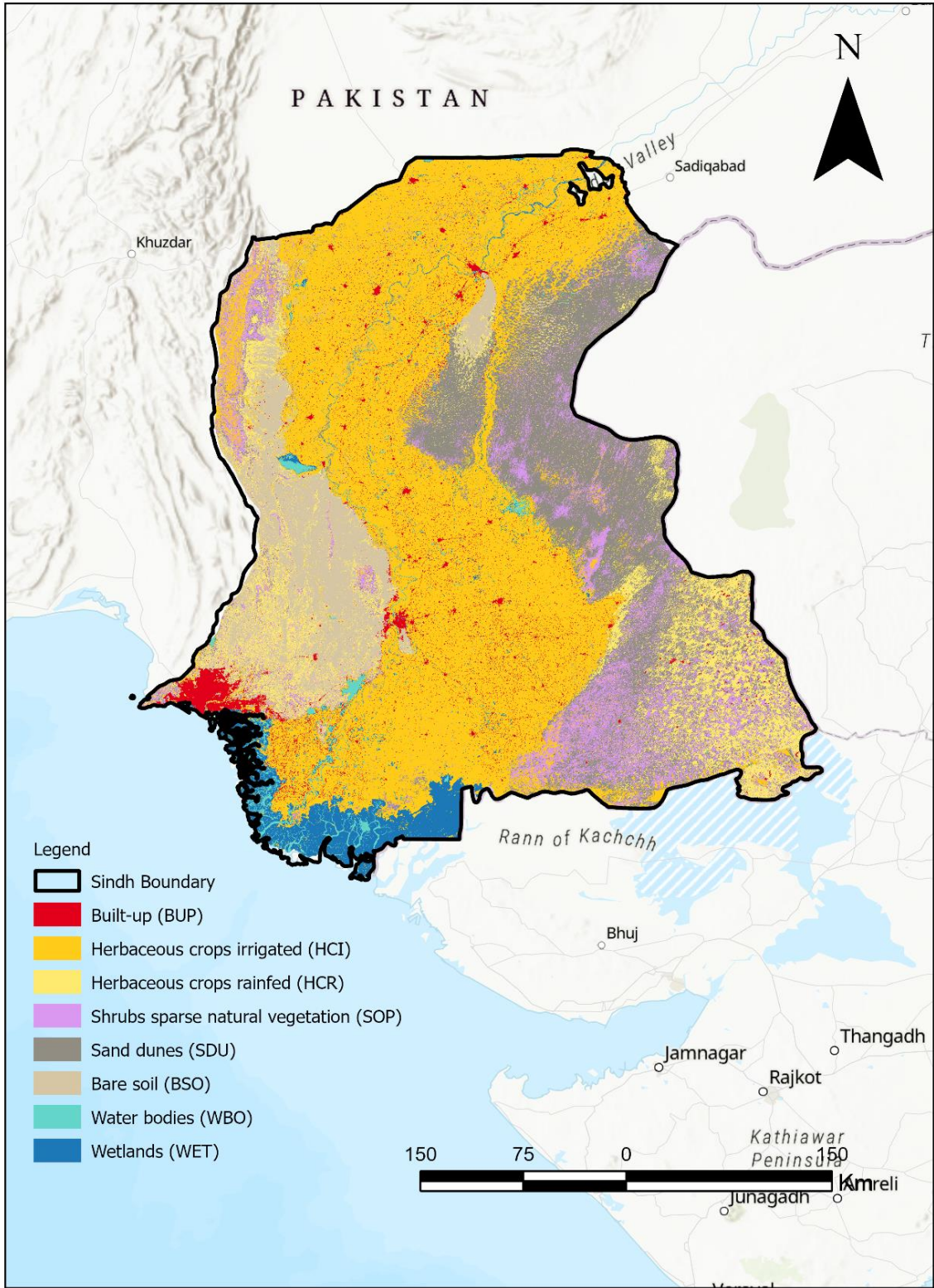
Figure 3-9. LULC map generated from K-means clustering using 14 clusters.

3.4.1.3 Results of K-means Clustering with k=18

Figure 3-10 illustrates the application of the k-means clustering algorithm to Sentinel-2 satellite composite imagery, initially configured with 18 clusters. This process divided the landscape into 18 distinct clusters, each class representing a unique land use and land cover (LULC) category as identified from the Sentinel-2 data. These clusters were subsequently assigned labels corresponding to specific LULC types which also required a reclassification to ensure an accurate LULC distribution. Class C_1 was retained in the LULC map despite being excluded from the confusion matrix analysis as the classifier did not identify this class accurately. The adjusted classification map is demonstrated in Figure 3-11, where the performance of this classification is evaluated using confusion matrices detailed in Table 3-4.
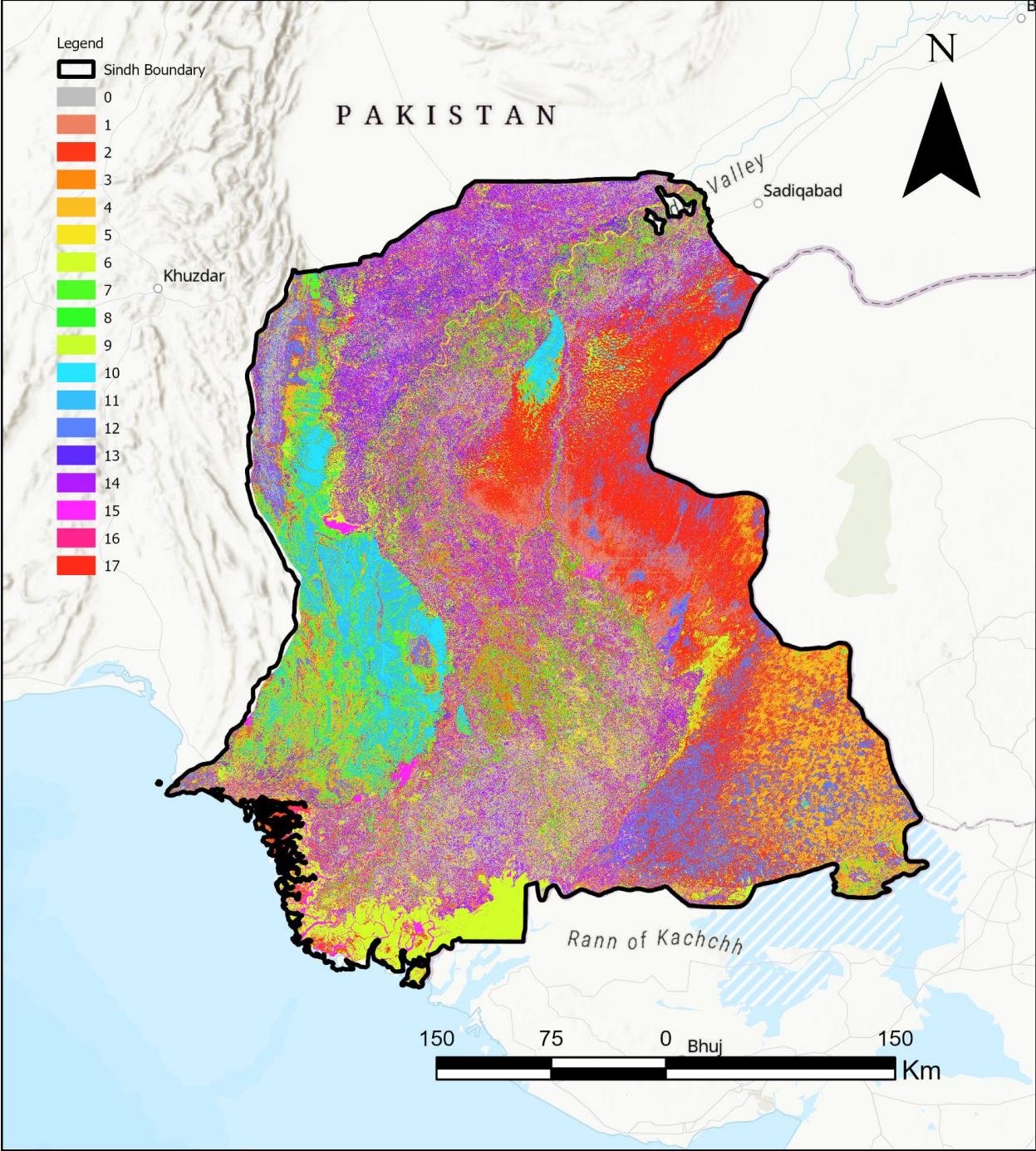
Figure 3-10. Output of K-means clustering using 18 clusters.

The confusion matrix analysis, adapted for 18 clusters, reveals varied accuracy levels across the LULC classes. Class C_4, for instance, displayed a user accuracy of 62.53%, while Classes C_5 and C_14 had lower user accuracies of 38.95% and 40.22%, respectively. Class C_15 showed a higher user accuracy of 69.65%. Class C_17 exhibited the highest user accuracy among the assessed classes at 76.12%. The producer accuracy presented a broad spectrum, with Class C_4 achieving 90.21% and C_14 showing 73.11%. The kappa statistic of 0.4703 indicates a moderate agreement between the k-means classified output and the reference LULC dataset, while the overall accuracy at 58.51% signifies that around 42% of the predictions were inaccurately classified.

Table 3-4. Confusion matrix between K-means (k=18) derived LULC map with FAO LULC map

Class	C_4	C_5	C_8	C_13	C_14	C_15	C_17	C_18	Total	U Acc (%)	Kappa
C_4	261039	3690	52509	2867	3007	35821	4515	6171	369619	62.53	0
C_5	3175	34590	25364	2	2842	22389	11	39	88412	38.95	0
C_8	16576	4080	65262	16	16375	28453	1224	1681	133667	47.07	0
C_13	1514	3	501	4581	0	110	404	415	7528	44.06	0
C_14	1915	12659	52280	0	62277	25097	26	29	154283	40.22	0
C_15	4277	14415	10401	1	232	68726	139	129	98320	69.65	0
C_17	114	10	61	132	0	69	9826	1868	12080	76.12	0
C_18	643	6	421	3556	3	1323	6847	16096	28895	43.45	0
Total	289253	69453	206799	11155	84736	181988	22992	26428	892804	0	0
P Acc (%)	90.21	49.72	31.49	40.12	73.11	37.44	41.74	60.32	0	58.51	0
Kappa	0	0	0	0	0	0	0	0	0	0	0.4703
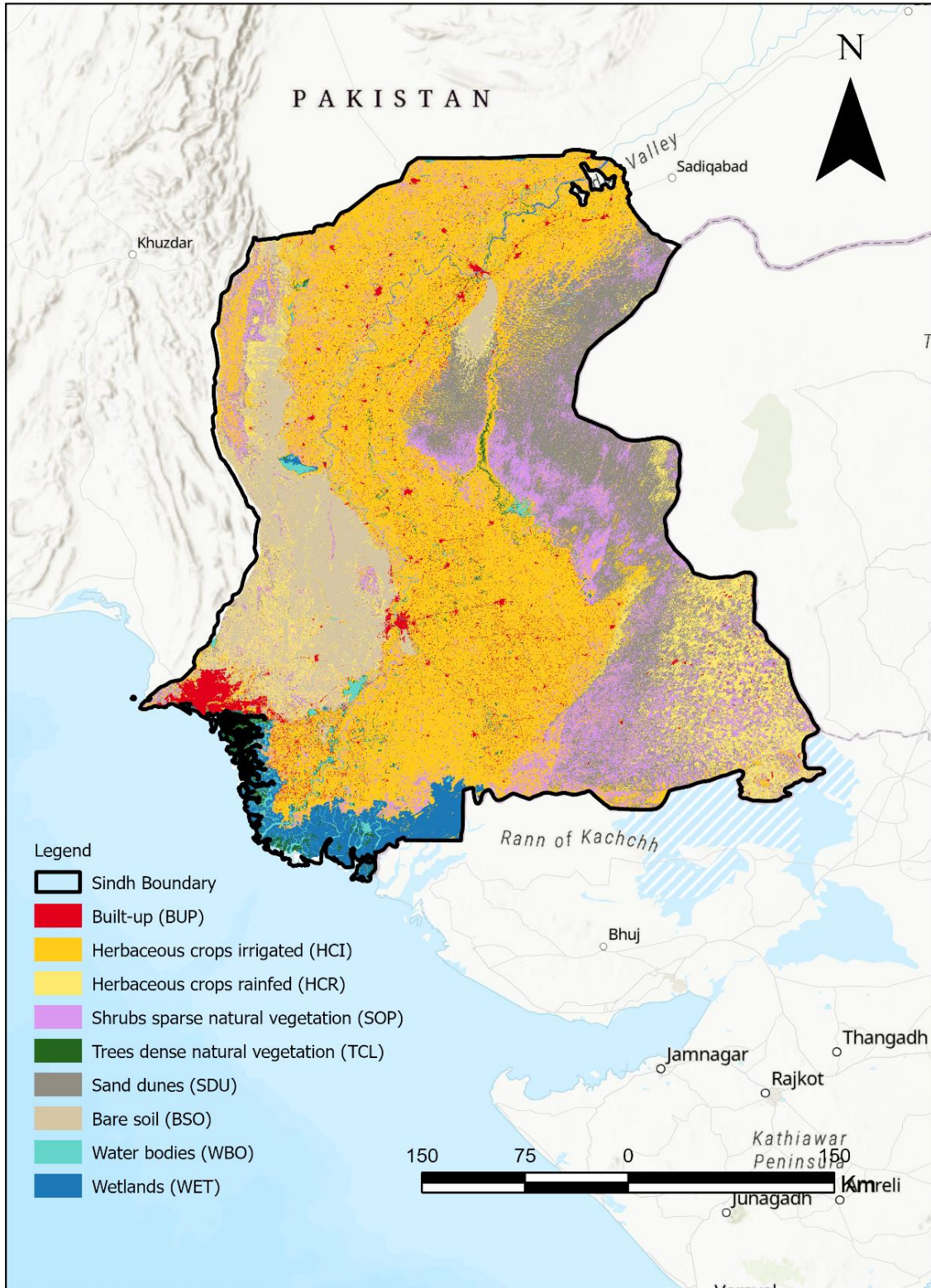
Figure 3-11. LULC map from K-means clustering using 18 clusters.

3.4.2 Random Forest Classification

The Random Forest classification algorithm was implemented on Sentinel-2 satellite composite imagery and the region of interest, Sindh province, was divided into distinct land use and land cover (LULC) categories. The resulting map, which segments the landscape into various LULC classes based on the Sentinel-2 data, is detailed in Figure 3-12. The effectiveness of this classification was evaluated using confusion matrices, with the results tabulated in Table 3-5.

The confusion matrix for the Random Forest classification reveals the algorithm's varied performance across LULC classes. Irrigated herbaceous crops represented by Class C_4 was accurately identified, as evidenced by a user accuracy of approximately 68.92%. In contrast, certain classes such as C_6 and C_7 demonstrated much lower user accuracies of 0.1% and nearly 0.05%, respectively. The algorithm effectively classified Class C_15 with a user accuracy of 74.18%. The producer accuracies across classes also varied, highlighting the algorithm's differing effectiveness in identifying specific LULC types. Class C_4, for example, achieved a high producer accuracy of 89.69%. Similarly, Class C_14, which indicates Sand dunes, showed a significant producer accuracy of 77.27%. The kappa statistic, standing at 0.5637, indicates a meaningful level of agreement between the Random Forest classified output and the reference LULC dataset, underscoring the model's overall predictive accuracy.

Table 3-5. Confusion matrix between RF-derived LULC map with FAO LULC map.

Class	C_1	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10	C_11	C_13	C_14	C_15	C_16	C_17	C_18	Total	U Acc (%)	Kappa
C_1	2774	39	1979	87	20	0	1635	2	189	167	15	14	1309	381	288	348	9247	30	0
C_3	363	5683	2520	6	3	0	401	76	606	474	323	0	35	13	68	30	10601	53.61	0
C_4	18412	8474	259100	1460	557	82	35837	317	8361	12827	2009	85	19056	1865	3107	4377	375926	68.92	0
C_5	1638	400	6052	35393	20	6	15214	68	566	800	132	357	6299	78	135	180	67338	52.56	0
C_6	593	146	2097	2218	9	0	2526	29	211	324	57	87	621	36	72	74	9100	0.1	0
C_7	370	49	1112	1951	2	4	2812	41	149	181	31	70	603	21	39	52	7487	0.05	0
C_8	4196	122	5959	18576	77	43	109963	1812	596	979	106	7818	12687	188	94	174	163390	67.3	0
C_9	339	55	980	366	6	3	3142	172	229	250	57	766	1024	37	90	58	7574	2.27	0
C_10	227	250	1589	241	10	0	1735	161	1109	518	311	572	645	41	147	177	7733	14.34	0

C_11	77	31	392	142	1	1	1122	10	128	83	44	606	531	27	163	70	3428	2.42	0
C_13	63	41	487	133	3	0	1368	5	311	55	6091	1402	841	85	411	614	11910	51.14	0
C_14	177	3	343	1406	7	2	13837	8	131	20	256	64923	12883	153	188	248	94585	68.64	0
C_15	1918	6	5044	7300	26	35	15341	147	280	132	183	7294	122397	1592	1308	2002	165005	74.18	0
C_16	270	1	433	11	6	0	236	2	143	11	271	13	1150	3809	1025	2498	9879	38.56	0
C_17	163	4	524	12	7	1	369	8	422	27	414	4	538	1651	13936	3051	21131	65.95	0
C_18	93	2	280	4	1	0	129	1	138	6	807	6	715	4102	1815	12412	20511	60.51	0
Total	31673	15306	288891	69306	755	177	205667	2859	13569	16854	11107	84017	181334	14079	22886	26365	984845	0	0
P Acc (%)	8.76	37.13	89.69	51.07	1.19	2.26	53.47	6.02	8.17	0.49	54.84	77.27	67.5	27.05	60.89	47.08	0	64.77	0
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5637
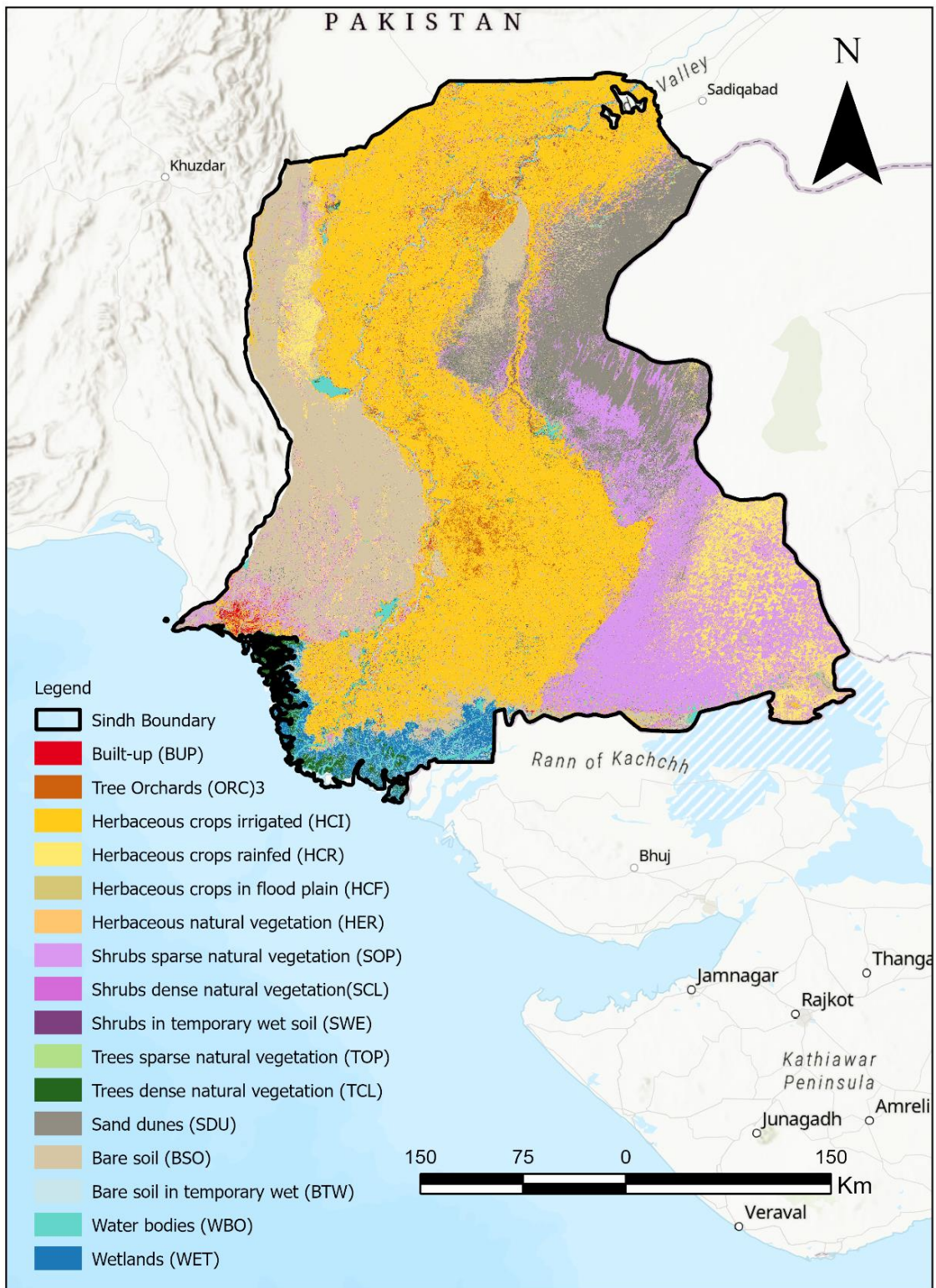
Figure 3-12. LULC map generated by Random Forest Algorithm.

3.4.3   LULC Dataset Selection

Following a detailed analysis of land use and land cover (LULC) patterns using k-means clustering and Random Forest classification algorithms on Sentinel-2 satellite imagery, two distinctive methodologies were identified. Each result provides perspectives on the distribution of LULC across the area studied. The k-means clustering algorithm was explored with varying cluster counts of 10, 14, and 18. In this study, K-means clustering, and Random Forest classification techniques were implemented to analyze and classify land use and land cover (LULC) patterns. K-means clustering was initiated with k=10, 14, and 18 clusters. In the comparative analysis between K-means clustering and Random Forest classification for mapping land use and land cover (LULC) within the study area, the Random Forest classification showed better results – presented in Table 3-6. The Random Forest method's higher Kappa statistic of 0.5637, suggests a more robust agreement with the reference LULC dataset compared to K-means clustering. Additionally, it achieved an overall accuracy of 64.77%, making it the preferred method over the K-means configurations tested.

Table 3-6. Classification accuracy and kappa coefficients for adopted classification techniques.

Classification	Overall Accuracy	Kappa
Unsupervised- 10 Clusters	61.41%	0.4581
Unsupervised- 14 Clusters	59.95%	0.4878
Unsupervised- 18 Clusters	58.51%	0.4703
Supervised - RF	64.77%	0.5637

3.4.4   Overlay and Zonal Statistics

The study employed ArcGIS software to overlay the flood extent map onto the LULC map and the administrative boundaries of Sindh province, merging three essential geographic layers to analyze the flood's impact across various land uses and administrative zones. This overlay process facilitated the analysis of flood coverage across various LULC classes within each administrative zone. By employing zonal statistics, detailed insights were garnered regarding the spatial distribution and intensity of flood impacts within distinct LULC categories and administrative units. The integration of these layers allowed for an assessment of the flood's impact on each LULC class in the Sindh province and is shown in Table 3-7.

Table 3-7. Flood extent over various Land cover/Land use classes (km^2).

Land Class	Total Area	Flooded Area	% Flooded
BUP	1384	250	18.1
ORC	1723	125	7.3
HCI	61471	19296	31.4
HCR	9494	269	2.8
HCF	0	0	0
HER	0	0	0
SOP	30572	1028	3.4
SCL	30	1	3.3
SWE	451	76	16.9
TOP	548	69	12.6
TCL	1413	68	4.8
SDU	14532	53	0.4
BSO	26365	3335	12.6
BTW	1288	141	10.9
WBO	3407	124	3.6
WET	3343	315	9.4

The flooded extent over the LULC classes, and each district within Sindh were identified and the results have been presented in Table 3-8 and Table 3-9 respectively. Among the districts, Kambar Shahdad Kot experienced the highest flood coverage in built-up areas, with an extent of

28.75 km^2 affected. This was closely followed by Jacobabad and Larkana, where the floodwaters impacted 23.79 and 25.01 km^2 of built-up regions, respectively. These districts represent the areas with the most inundations on urban and built-up land, highlighting a critical need for targeted flood mitigation and urban planning strategies.

In the district of Dadu, the herbaceous crops irrigated (HCI) class experienced the most significant flood impact, with 1384.90 km^2 affected (63.5% of the total flooded area in the district). This is followed by the Kambar Shahdad Kot district, where the HCI class again faced a substantial inundation, covering an area of 3167.34 km^2 (84.4% of the total flooded area in the district). The third highest district, Jacobabad, showed a similar trend with 1946.87 km^2 the HCI class impacted by flooding (88.7% of the total flooded area in the district). The HCI class stands out as the most affected, with 31.4% of its total area inundated, highlighting the vulnerability of irrigated agricultural lands to flood events in Sindh. The second most impacted class is the shrubs sparse natural vegetation (SOP), with significant inundation in Kambar Shahdad Kot (157.18 km^2 - 4.2% of the total flooded area in the district) and Dadu (113.30 km^2 - 5.2% of the total flooded area in the district). The third class, bearing significant flood impact, is the bare soil (BSO) class, especially within the Jamshoro district, where 1062.82 km^2 were affected (66.1% of the total flooded area in the district).

The analysis of the flood extent over built-up areas across different districts of Sindh province, as depicted in the provided data, revealed significant inundation within specific regions. Jacobabad, Larkana and Kambar Shahdad Kot are the most affected districts with 71.6%, 66.9% and 59.2% of the land affected. In terms of affected area, Kambar Shahdad Kot, Jacobabad, Dadu are the most affected districts with 3753 km^2, 2195 km^2 and 2179 km^2 of the land affected.

Table 3-8. Flood extent by District (km2) over various Land cover/Land use classes (%).

District	Total Area (km^2)	Flooded Area (km^2)	% Flooded	% of Land Class/Land use in the Flooded Area in the District												
				BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET
Badin	7270	1881	25.9	0.3	0.2	84.5	0	2.8	0	0.1	0.1	0.1	0.1	10.5	0.2	0.3	0.9
Central Karachi	72	0	0.6	28.9	0	37.8	0	8.9	0	0	0	0	0	22.2	0	2.2	2.2
Dadu	9024	2179	24.2	0.7	0.1	63.5	8.3	5.2	0	0.1	0.1	0.1	0.2	21.1	0.2	0.3	0.1
East Karachi	225	2	0.9	12.1	0	32	1.5	25.2	0	0	0.5	0	1.5	24.8	1	1	0.5
Ghotki	7227	620	8.6	1.2	0.9	84.2	0	4.9	0	0.3	0.4	0.1	1.7	5.1	0.2	0.7	0.4
Hyderabad	1132	200	17.6	1.5	3.8	81.5	0.5	2.9	0	0.3	0.3	0.1	0.1	6.9	0.7	0.8	0.5
Jacobabad	3065	2195	71.6	1.1	0.1	88.7	0	3.6	0	0.1	0.3	0	0.1	4.7	0.2	0.5	0.5
Jamshoro	12520	1608	12.8	0.7	0.1	25.2	2.8	4.5	0	0	0.1	0	0.1	66.1	0.2	0.2	0.1
Kambar Shahdad Kot	6342	3753	59.2	0.8	0.2	84.4	0.7	4.2	0	0.5	0.3	0.2	0.2	7.9	0.1	0.4	0.2
Kashmore	2980	800	26.8	1	0.2	88.6	0	3.5	0	0.2	0.3	0	0.1	4.9	0.4	0.5	0.4
Khairpur	18008	1163	6.5	1.7	0.9	82.1	0.1	5.7	0	0.7	0.5	0.2	0.3	6	0.4	0.9	0.4

District	Total Area (km²)	Flooded Area (km²)	% Flooded	% of Land Class/Land use in the Flooded Area in the District												
				BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET
Korangi Karachi	115	2	1.9	6.3	0.5	49.8	1.4	16.7	0	0.9	0.9	0.9	1.4	16.7	0.9	1.8	1.4
Larkana	2175	1455	66.9	1.7	0.2	88.1	0.1	4.3	0	0.2	0.4	0	0.1	3.9	0.4	0.5	0.2
Malir Karachi	3025	97	3.2	0.8	0.1	8.4	1.4	12.1	0	0.1	0.2	1	0.8	69.8	1.6	0.7	3
Matiari	1626	463	28.5	1.3	3.7	87.2	0	2.4	0	0.4	0.4	0.1	0	3.4	0.7	0.3	0.2
Mirpur Khas	3848	495	12.9	0.3	1.5	88.4	0	1.2	0	0.2	0.2	0.1	0	7.4	0.1	0.3	0.3
Naushahro Feroze	3418	1069	31.3	1.9	0.7	85	0.1	5.3	0	0.3	0.5	0.1	0.1	4.3	0.9	0.5	0.2
Sanghar	11345	832	7.3	0.9	1.1	86.6	0	2.1	0	0.8	0.2	0.6	0	5.5	0.2	1	1
Shaheed Benazir Abad	5135	1199	23.4	1.5	0.7	84.6	0.1	4.3	0	0.3	0.3	0.1	0.1	6.8	0.2	0.6	0.3
Shikarpur	2900	1624	56	1.3	0.1	86.2	0.1	6.5	0	0.2	0.4	0	0.1	4.5	0.1	0.2	0.1
South Karachi	562	10	1.8	3.7	0.3	26.8	0.5	12.7	0	0.1	0.1	0.8	4.6	36.2	4.3	5.8	4.1
Sujawal	9571	1306	13.6	0.5	0.1	46	0	2.2	0	0.4	0.1	2.7	0.5	24.9	5.4	0.7	16.4

111

District	Total Area (km²)	Flooded Area (km²)	% Flooded	% of Land Class/Land use in the Flooded Area in the District													
				BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET
Sukkur	5861	541	9.2	1.4	0.5	83.4	0.2	3.9	0	0.3	0.5	0.1	0.2	8.4	0.2	0.5	0.3
Tando Allahyar	1718	211	12.3	0.3	5.9	89.5	0	0.8	0	0.4	0.2	0.6	0	1.3	0	0.5	0.4
Tando Muhammad Khan	1858	457	24.6	0.6	1.3	89.8	0	1.5	0	0.2	0.2	0.1	0	5.5	0.2	0.3	0.4
Tharparkar	21775	76	0.3	0.8	0.2	35.4	1.2	7.3	0	0.1	0.1	0	5.1	44.3	2.7	1.7	1.3
Thatta	8438	586	6.9	1	0.9	53.4	0.8	4.7	0	1.1	0.3	1	0.2	30.1	1.5	1.6	3.4
Umer Kot	6102	324	5.3	0.5	0.3	87.8	0	2.1	0	0.1	0.1	0.1	0	7.9	0.1	0.3	0.5
West Karachi	268	5	1.7	6.9	0.2	55.3	0.9	15.1	0	0.4	0.2	0.2	0.4	14.2	0.9	3.4	1.7

Table 3-9. Flooded area over various Land Use/Land Cover classes by District (km^2).

District	BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET	Total (km^2)
Badin	6.58	3.66	1588.26	0.16	51.82	0.02	2.38	2.7	1.08	1.64	196.99	3.62	5.55	16.03	1880.54
Central Karachi	0.13	0	0.17	0	0.04	0	0	0	0	0	0.1	0	0.01	0.01	0.45
Dadu	14.74	2.5	1384.9	180.85	113.3	0.05	1.66	2.82	1.28	4.38	459.97	3.63	6.1	3.26	2179.44
East Karachi	0.25	0	0.66	0.03	0.52	0	0	0.01	0	0.03	0.51	0.02	0.02	0.01	2.06
Ghotki	7.45	5.41	521.89	0.27	30.64	0.04	1.7	2.23	0.61	10.56	31.52	1.32	4.06	2.39	620.11
Hyderabad	3.08	7.69	162.79	1.03	5.87	0	0.58	0.58	0.23	0.21	13.77	1.42	1.6	0.92	199.78
Jacobabad	23.79	1.73	1946.87	0.29	79.45	0.02	2.66	6.55	0.57	1.42	104.21	3.73	11.52	11.89	2194.7
Jamshoro	10.75	0.81	405.12	45.11	72.58	0.01	0.59	0.94	0.1	1.63	1062.82	3.92	2.6	1.36	1608.35

District	BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET	Total (km²)
Kambar Shahdad Kot	28.75	7.82	3167.34	25.79	157.18	0.25	18.36	11.51	7.42	8.57	296.48	2.85	14.57	6.02	3752.93
Kashmore	8.26	1.58	708.76	0.26	28.09	0.03	1.42	2.07	0.33	0.68	38.81	2.8	3.96	2.86	799.92
Khairpur	20.04	10.4	954.59	1.65	66.58	0.04	8.57	5.78	1.88	3.67	69.41	4.97	10.59	4.51	1162.67
Korangi Karachi	0.14	0.01	1.1	0.03	0.37	0	0.02	0.02	0.02	0.03	0.37	0.02	0.04	0.03	2.21
Larkana	25.01	2.79	1281.98	1.23	61.92	0.03	2.28	5.4	0.5	0.87	56.36	5.99	7.78	2.48	1454.62
Malir Karachi	0.74	0.07	8.06	1.39	11.64	0	0.12	0.2	0.92	0.8	67.41	1.57	0.71	2.87	96.51
Matiari	5.8	17	403.37	0.11	10.89	0.01	1.78	1.68	0.57	0.23	15.74	3.26	1.43	0.74	462.61

District	BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET	Total (km²)
Mirpur Khas	1.66	7.65	437.21	0.08	5.76	0.01	0.95	0.9	0.72	0.03	36.73	0.31	1.24	1.51	494.76
Naushahro Feroze	20.67	7.69	908.33	1.16	56.3	0.02	3.31	4.94	1.18	1.53	46.07	9.93	5.25	2.47	1068.86
Sanghar	7.56	9.24	720.29	0.04	17.16	0.02	6.42	1.66	4.94	0.15	45.53	1.95	8.65	7.97	831.59
Shaheed Benazir Abad	17.47	8.46	1014.43	1.78	51.96	0.01	3.96	3.4	1.43	1	81.21	2.66	7.63	3.91	1199.33
Shikarpur	20.85	1.99	1400.81	1.07	106.35	0.04	3.77	7.15	0.48	2.3	72.45	1.86	3.66	1.57	1624.35
South Karachi	0.38	0.03	2.77	0.05	1.31	0	0.01	0.01	0.08	0.48	3.74	0.44	0.6	0.42	10.33
Sujawal	6.29	1.68	600.22	0.09	28.5	0.02	5.41	1.57	35.33	6.94	324.78	70.78	9.61	214.82	1306.05

District	BUP	ORC	HCI	HCR	SOP	SCL	SWE	TOP	TCL	SDU	BSO	BTW	WBO	WET	Total (km²)
Sukkur	7.56	2.57	451.02	1.26	20.92	0.04	1.7	2.82	0.5	0.87	45.52	1.2	2.96	1.53	540.5
Tando Allahyar	0.69	12.41	188.69	0.01	1.75	0	0.9	0.47	1.24	0	2.71	0.09	0.99	0.85	210.8
Tando Muhammad Khan	2.8	5.77	410.57	0.02	6.87	0.01	0.87	1.07	0.24	0.06	25.05	0.72	1.26	1.67	456.97
Tharparkar	0.59	0.14	26.85	0.9	5.52	0	0.05	0.05	0.03	3.89	33.64	2.05	1.28	0.95	75.95
Thatta	5.86	5.06	312.94	4.61	27.37	0.02	6.16	1.75	6.05	1	176.51	9.01	9.38	20.02	585.73
Umer Kot	1.61	1.1	283.9	0.02	6.92	0	0.25	0.43	0.42	0.07	25.61	0.43	1.07	1.68	323.51
West Karachi	0.32	0.01	2.57	0.04	0.7	0	0.02	0.01	0.01	0.02	0.66	0.04	0.16	0.08	4.65

3.5 Discussion

Unsupervised K-means and supervised Random Forest classification were employed to evaluate the Land Use/Land Cover (LULC) dataset and extent of flooding on land cover classes within Sindh province. This analysis highlighted the flood's extensive reach across different land cover classes, emphasizing the distinct vulnerability of built-up areas, herbaceous crops irrigated, and shrubs sparse natural vegetation to flood events. Notably, Kambar Shahdad Kot, Jacobabad, and Larkana districts emerged as significantly affected areas, underlining the urgent need for focused flood mitigation strategies in urban planning endeavors.

The evaluation of K-means clustering, and Random Forest (RF) classification brought out essential information regarding the efficiency of these techniques in delineating LULC across the study region. The Random Forest classification showed higher accuracy and alignment with the reference dataset compared to several results of K-means classification. The accuracy of K-means clustering was impacted by the choice of the number of clusters and the initial seed values, which could lead to variations in the resulting classifications. In K-means, the reliance on spectral similarities without ground data can lead to inaccuracies, whereas Random Forest uses a labeled dataset for training and employs an ensemble approach with multiple decision trees to manage spectral dissimilarities within the same cluster and reduce overfitting, thereby enhancing prediction accuracy in complex datasets. Furthermore, the use of ground truth training data facilitates informed and accurate categorization of land cover types, an advantage not available with K-means clustering.

The accuracy of K-means clustering, an unsupervised learning method, can be significantly affected by several factors. Firstly, the initialization of centroids plays a crucial role; different

initializations can lead to varying outcomes due to the algorithm's sensitivity to starting points. The selection of the number of clusters, as informed by the Elbow Method in this study, also impacts quality of the classification. Moreover, the presence of mixed pixels, where a single spatial unit displays characteristics of multiple land cover types, poses a challenge for K-means, complicating the classification process. These factors collectively contribute to the moderate accuracy levels observed in the analysis, with overall accuracy ranging from 58.51% to 61.41% for different cluster numbers (k=10, 14, and 18) while kappa values were 0.4638 to 0.4878.

Random Forest (RF), a supervised classification method, demonstrated better accuracy due to its inherent advantages over unsupervised methods like K-means. Factors affecting RF accuracy include the quality and quantity of training points, feature selection, and the algorithm's capability to handle high-dimensional data and mixed pixels more effectively. RF's ensemble approach, generating multiple decision trees and aggregating their predictions, reduces the risk of overfitting and enhances classification accuracy. The kappa statistic (0.5602) and overall accuracy (64.29%) in the study underscore the effectiveness of RF in handling the complexity of LULC classification.

The quality and quantity of training data are pivotal in both K-means clustering and Random Forest classification. In K-means, although unsupervised, the interpretation of clusters as specific LULC classes can benefit indirectly from knowledge in selecting and validating clusters. For Random Forest, the direct use of training points significantly influences the model's ability to learn and generalize. Accurate, well-distributed training data covering the range of LULC classes ensures that RF models can correctly classify the land cover. Also, there is a lack of reliable and extensive ground data for Sindh province, because of which this study was heavily reliant on the FAO's land cover dataset. While the FAO dataset provides a valuable resource for such contexts,

the lack of detailed information on its development and processing raises questions about the assumptions and accuracy of the underlying data.

3.6 Conclusion

Remote sensing serves as a fundamental tool in flood studies, providing essential data for inventorying, monitoring, mapping, and responding to disasters. This study underscored the critical role of remote sensing technologies in assessing flooding extents, identifying affected regions, and comprehending flood dynamics under varying weather conditions. Moreover, remote sensing aids in analyzing changes in vegetation cover, influencing water dynamics and runoff patterns, thereby supporting urban planning, vulnerability assessment, and development of flood mitigation strategies. Disaster damage assessment plays a crucial role in emergency management, providing vital insights for resource allocation and informed decision-making. Land use and land cover (LULC) mapping, enabled by remote sensing, GIS, and machine learning algorithms, contributes to understanding human-nature interactions and supports sustainable development initiatives. Unsupervised and supervised classification techniques, such as K-means clustering and Random Forest classification, offer systematic categorization of remote sensing data, aiding in accurate land cover mapping. The study aims to assess flood impacts on various land cover classes in Sindh, Pakistan, utilizing remote sensing-based flood maps and LULC datasets within a GIS framework, thereby enhancing spatial understanding of flooding across different land classes during flood events.

The exploration of the impact of flooding on various land cover classes within Sindh province through remote sensing-based flood maps and Land Use/Land Cover (LULC) datasets culminates in revealing the significant role and application of both unsupervised (K-means clustering) and supervised (Random Forest classification) methods in assessing environmental

119

changes and disaster impacts. The study highlights the vulnerability of specific land cover classes such as built-up areas, irrigated herbaceous crops, and sparse natural vegetation to flooding, highlighting the necessity for targeted mitigation strategies and urban planning interventions. The comparative analysis of K-means clustering and Random Forest classification in this study illuminates the strengths and limitations of each method in the context of LULC mapping. The superior performance of RF, driven by its ensemble nature and effective handling of training data, emphasizes the value of supervised learning in complex classification tasks. However, the challenges posed by limited ground-truth data and the reliance on secondary datasets like FAO's LULC map emphasize the critical necessity for robust ground-truth datasets to refine the precision of remote sensing analyses. This study's findings contribute to a deeper understanding of the dynamics of flood impacts across different land cover classes in Sindh, offering valuable insights for future mitigation and planning efforts.

In conclusion, this study enhances our understanding of flood dynamics and their impact in Sindh province, while also broadening the knowledge of the applicability and limitations of remote sensing techniques in environmental monitoring and disaster management. The findings provide a deeper insight into how floods affect various land cover classes in Sindh and underscore the importance of combining advanced remote sensing technologies with reliable and accessible ground-truth datasets, like LULC data. This integration offers valuable insights for future mitigation and planning efforts.

3.7 References

Amatya, P., Kirschbaum, D., & Stanley, T. (2019). Use of very high-resolution optical data for landslide mapping and susceptibility analysis along the Karnali highway, Nepal. *Remote Sensing, 11(19)*. https://doi.org/10.3390/rs11192284

Berg, P., Pham, M. T., & Courty, N. (2022). Self-Supervised Learning for Scene Classification in Remote Sensing: Current State of the Art and Perspectives. *In Remote Sensing (Vol. 14, Issue 16)*. https://doi.org/10.3390/rs14163995

Bioresita, F., Puissant, A., Stumpf, A., & Malet, J. P. (2018). A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery. *Remote Sensing, 10(2)*. https://doi.org/10.3390/rs10020217

Bolanos, S., Stiff, D., Brisco, B., & Pietroniro, A. (2016). Operational surface water detection and monitoring using Radarsat 2. *Remote Sensing, 8(4)*. https://doi.org/10.3390/rs8040285

Bui, D. H., & Mucsi, L. (2021). From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal landsat data, a random forest classifier, and decision rules. *Remote Sensing, 13(9)*. https://doi.org/10.3390/rs13091700

Chen, Z., Wagner, M., Das, J., Doe, R. K., & Cerveny, R. S. (2021). Data-driven approaches for tornado damage estimation with unpiloted aerial systems. *Remote Sensing, 13(9)*. https://doi.org/10.3390/rs13091669

Cian, F., Marconcini, M., & Ceccato, P. (2018). Normalized Difference Flood Index for

rapid flood mapping: Taking advantage of EO big data. *Remote Sensing of*

Environment, 209. https://doi.org/10.1016/j.rse.2018.03.006

de Lima, R. P., & Marfurt, K. (2020). Convolutional neural network for remote-sensing

scene classification: Transfer learning analysis. *Remote Sensing, 12(1).*

https://doi.org/10.3390/rs12010086

Ghaffarian, S., Kerle, N., & Filatova, T. (2018). Remote sensing-based proxies for urban

disaster risk management and resilience: A review. *In Remote Sensing (Vol. 10, Issue*

11). https://doi.org/10.3390/rs10111760

Hamidi, E., Peter, B. G., Munoz, D. F., Moftakhari, H., & Moradkhani, H. (2023). Fast

Flood Extent Monitoring With SAR Change Detection Using Google Earth Engine.

IEEE Transactions on Geoscience and Remote Sensing, 61.

https://doi.org/10.1109/TGRS.2023.3240097

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep

learning benchmark for land use and land cover classification. *IEEE Journal of*

Selected Topics in Applied Earth Observations and Remote Sensing, 12(7).

https://doi.org/10.1109/JSTARS.2019.2918242

Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural

networks for the scene classification of high-resolution remote sensing imagery.

Remote Sensing, 7(11). https://doi.org/10.3390/rs71114680

Huang, W., Wang, D., & Ratilal, P. (2016). Diel and spatial dependence of humpback song

and non-song vocalizations in fish spawning ground. *Remote Sensing, 8(9)*.
https://doi.org/10.3390/rs8090712

Hütt, C., Koppe, W., Miao, Y., & Bareth, G. (2016). Best accuracy land use/land cover
(LULC) classification to derive crop types using multitemporal, multisensor, and multi-
polarization SAR satellite images. *Remote Sensing, 8(8)*.
https://doi.org/10.3390/rs8080684

Irwin, K., Braun, A., Fotopoulos, G., Roth, A., & Wessel, B. (2018). Assessing single-
polarization and dual-polarization TerraSAR-X data for surface water monitoring.
Remote Sensing, 10(6). https://doi.org/10.3390/rs10060949

Karim, M. F., & Zhang, X. (2021). Analysis of vegetative cover vulnerability in rohingya
refugee camps of bangladesh utilizing landsat and per capita greening area (Pcga)
datasets. *Remote Sensing, 13(23)*. https://doi.org/10.3390/rs13234922

Kitajima, N., Seto, R., Yamazaki, D., Zhou, X., Ma, W., & Kanae, S. (2021). Potential of a
sar small-satellite constellation for rapid monitoring of flood extent. *Remote Sensing,
13(10)*. https://doi.org/10.3390/rs13101959

Langford, Z. L., Kumar, J., Hoffman, F. M., Breen, A. L., & Iversen, C. M. (2019). Arctic
vegetation mapping using unsupervised training datasets and convolutional neural
networks. *Remote Sensing, 11(1)*. https://doi.org/10.3390/rs11010069

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification
algorithms and training sample sizes in urban land classification with landsat thematic
mapper imagery. *Remote Sensing, 6(2)*. https://doi.org/10.3390/rs6020964

Li, L., Xu, T., & Chen, Y. (2016). Improved urban flooding mapping from remote sensing images using generalized regression neural network-based super-resolution algorithm. *Remote Sensing, 8(8).* https://doi.org/10.3390/rs8080625

Martinis, S., Twele, A., Strobl, C., Kersten, J., & Stein, E. (2013). A multi-scale flood monitoring system based on fully automatic MODIS and terraSAR-X processing chains. *Remote Sensing, 5(11)*. https://doi.org/10.3390/rs5115598

Mellor, A., Haywood, A., Stone, C., & Jones, S. (2013). The performance of random forests in an operational setting for large area sclerophyll forest classification. *Remote Sensing, 5(6).* https://doi.org/10.3390/rs5062838

Morin, N., Masse, A., Sannier, C., Siklar, M., Kiesslich, N., Sayadyan, H., Faucqueur, L., & Seewald, M. (2021). Development and application of earth observation based machine learning methods for characterizing forest and land cover change in dilijan national park of armenia between 1991 and 2019. *Remote Sensing, 13(15).* https://doi.org/10.3390/rs13152942

Nex, F., Duarte, D., Tonolo, F. G., & Kerle, N. (2019). Structural building damage detection with deep learning: Assessment of a state-of-the-art CNN in operational conditions. *Remote Sensing, 11(23).* https://doi.org/10.3390/rs11232765

Saini, R., & Ghosh, S. K. (2018). CROP CLASSIFICATION ON SINGLE DATE SENTINEL-2 IMAGERY USING RANDOM FOREST AND SUPPOR VECTOR MACHINE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII–5*. https://doi.org/10.5194/isprs-archives-xlii-5-683-

2018

Santaga, F. S., Agnelli, A., Leccese, A., & Vizzari, M. (2021). Using sentinel-2 for
simplifying soil sampling and mapping: Two case studies in Umbria, Italy. *Remote
Sensing, 13(17)*. https://doi.org/10.3390/rs13173379

Scarth, P., Armston, J., Lucas, R., & Bunting, P. (2019). A structural classification of
Australian vegetation using ICESat/GLAS, ALOS PALSAR, and Landsat sensor data.
Remote Sensing, 11(2). https://doi.org/10.3390/rs11020147

Schumann, G. J. P. (2015). Preface: Remote sensing in flood monitoring and management.
In Remote Sensing (Vol. 7, Issue 12). https://doi.org/10.3390/rs71215871

Silva-Cardoza, A. I., Vega-Nieva, D. J., Briseño-Reyes, J., Briones-Herrera, C. I., López-
Serrano, P. M., Corral-Rivas, J. J., Parks, S. A., & Holsinger, L. M. (2022). Evaluating
a New Relative Phenological Correction and the Effect of Sentinel-Based Earth Engine
Compositing Approaches to Map Fire Severity and Burned Area. *Remote Sensing,
14(13)*. https://doi.org/10.3390/rs14133122

Sonobe, R. (2019). Parcel-based crop classification using multi-temporal TerraSAR-X dual
polarimetric data. *Remote Sensing, 11(10).* https://doi.org/10.3390/rs11101148

Sublime, J., & Kalinicheva, E. (2019). Automatic post-disaster damage mapping using deep-
learning techniques for change detection: Case study of the Tohoku tsunami. *Remote
Sensing, 11(9).* https://doi.org/10.3390/rs11091123

Sun, Y., Liu, C., Du, X., Yang, F., Yao, Y., Soomro, S. e. hyde., & Hu, C. (2022). Urban

storm flood simulation using improved SWMM based on K-means clustering of parameter samples. *Journal of Flood Risk Management, 15(4).* https://doi.org/10.1111/jfr3.12826

Tzepkenlis, A., Marthoglou, K., & Grammalidis, N. (2023). Efficient Deep Semantic Segmentation for Land Cover Classification Using Sentinel Imagery. *Remote Sensing, 15(8).* https://doi.org/10.3390/rs15082027

Viana-Soto, A., Aguado, I., Salas, J., & García, M. (2020). Identifying post-fire recovery trajectories and driving factors using landsat time series in fire-prone mediterranean pine forests. *Remote Sensing, 12(9).* https://doi.org/10.3390/RS12091499

Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A., & White, E. (2019). Individual tree-crown detection in rgb imagery using semi-supervised deep learning neural networks. *Remote Sensing, 11(11).* https://doi.org/10.3390/rs11111309

White, L., Brisco, B., Dabboor, M., Schmitt, A., & Pratt, A. (2015). A collection of SAR methodologies for monitoring wetlands. *In Remote Sensing (Vol. 7, Issue 6).* https://doi.org/10.3390/rs70607615

Wulder, M. A., Li, Z., Campbell, E. M., White, J. C., Hobart, G., Hermosilla, T., & Coops, N. C. (2018). A national assessment of wetland status and trends for Canada's forested ecosystems using 33 years of earth observation satellite data. *Remote Sensing, 10(10).* https://doi.org/10.3390/rs10101623

Xie, Z., Chen, Y., Lu, D., Li, G., & Chen, E. (2019). Classification of land cover, forest, and tree species classes with Ziyuan-3 multispectral and stereo data. *Remote Sensing, 11(2).*

https://doi.org/10.3390/rs11020164

Yang, Y., Yang, D., Wang, X., Zhang, Z., & Nawaz, Z. (2021). Testing accuracy of land

cover classification algorithms in the qilian mountains based on gee cloud platform.

Remote Sensing, 13(24). https://doi.org/10.3390/rs13245064

Chapter 4 - Contributions and Recommendations

4.1 Summary

Climate change has emerged as a critical global issue, intensifying the impact of natural disasters like floods. These calamities are influenced by various factors, including alterations in atmospheric conditions such as temperature, precipitation, humidity, and air circulation. Despite Pakistan's minimal contribution to greenhouse gas emissions, the nation remains highly vulnerable to climate-induced disasters. This vulnerability is exacerbated by the country's diverse topography and human interventions such as urbanization and inadequate water management practices. The devastating floods of 2022, triggered by heavy rainfall and glacier melting, resulted in significant economic losses, particularly affecting provinces like Sindh, Balochistan, and Punjab. Understanding the complexities and dynamics of floods is essential for effective emergency responses and mitigation efforts and for formulating flood mitigation strategies. To explore and understand the potential of open-access remote sensing datasets, two objectives were defined in this thesis focusing on Pakistan Floods 2022 in the Sindh province.

The first objective was to investigate and validate the utility of the Synthetic Aperture Radar (SAR) dataset to identify the extent of inundated areas. To do so, the indices relevant to open water pixel identification such as Sentinel-1 SAR backscatter-based Difference Image Index and Sentinel-2 reflectance-based Desert Flood Index and Modified Normalized Difference Water Index were used to develop the change detection method. This approach exploited the unique capabilities of SAR and optical satellite data to detect and map the spatial extent of flooding leveraging the computation capabilities of Google Earth Engine.

The Difference Image Index (DII), derived from the backscatter values of Sentinel-1 data, was used to distinguish the flooded regions. Sentinel-2-based Desert Flood Index (DFI) and Modified Normalized Difference Water Index (MNDWI) images were used to optimize the threshold value of the SAR-based DII index. A region with minimal cloud cover during the acquisition of Sentinel-2 images was carefully selected to ensure the reliability of the flood extent mappings. The process of threshold fine-tuning, which was based on exhaustive grid search methods, played a crucial role in refining the detection process. This approach ensured the highest level of agreement between flood maps derived from SAR imagery and those obtained from optical imagery.

The second objective of this thesis is to evaluate the effects of floods across diverse land cover categories by leveraging flood maps and Land Use/Land Cover (LULC) datasets obtained through remote sensing. The effects of flooding on different land cover types within the administrative regions of Sindh were investigated utilizing Sentinel-2 images. The inundated land classes were identified which quantified the extent of flood damage and provided insights on the vulnerability of different land cover classes. The findings help us understand the long-term implications for land use planning and disaster management in flood-prone regions.

A comprehensive LULC dataset was prepared to reflect the pre-flood land cover conditions of the study area using Sentinel-2 imagery. Two sets of LULC datasets were created using unsupervised K-means clustering and supervised Random Forest classification techniques. The unsupervised K-means algorithm was optimized using the Elbow plot method to identify the most appropriate number of clusters, ensuring that the classification results were both accurate and reflective of the diverse land cover types present in the region. The supervised ensemble Random Forest classification utilized training data to classify the Sentinel-2 imagery into

detailed land cover classes. The Land Cover dataset for Sindh published by FAO-Pakistan has been used as ground-truth data for LULC development. The spatial distribution and magnitude of flood impacts on various land cover classes were studied with overlay and zonal statistical analyses conducted within a GIS environment.

Optimal threshold values for mapping flood extents were established using imagery from Sentinel-1 SAR and Sentinel-2. Calibration of thresholds was performed in Dadu district, Sindh province, on pre- and post-flood composite images, revealing darker regions post-flood due to reduced backscatter signals. Flood extent mappings were generated for both imagery types, utilizing the Difference Image Index (DII), Desert Flood Index (DFI), and Modified Normalized Difference Water Index (MNDWI). Comparative analysis between DII and DFI showed significant agreement, with an optimal DII threshold of 0.414 and an agreement of 79.31% at a DFI value of 0.25. Similarly, a comparison between DII and MNDWI indicated substantial agreement, with an optimal DII threshold of 0.427 and an agreement of 83.86% at an MNDWI value of 0.15. Flood extent mapping of Sindh province revealed extensive inundation areas for August and September, totaling 25229 km² and 19181 km² respectively for DFI-DII analysis, and 30582 km² and 21851 km² for MNDWI-DII analysis.

The Elbow method was applied to identify the ideal number of clusters for unsupervised classification. By applying the elbow method, a graph of the number of clusters versus the Within-Cluster Sum of Squares (WCSS) from the K-means clustering analysis was plotted. This plot usually shows a distinct elbow point occurring between 10 and 20 clusters. K-means clustering was applied to Sentinel-2 satellite imagery to delineate patterns of land use and cover. Evaluation using confusion matrices showed varied accuracies across different classes, with overall accuracies ranging from 59.95% to 61.41%. In contrast, Random Forest classification

outperformed K-means, boasting a kappa statistic of 0.5602 and an overall accuracy of 64.29%, thus becoming the preferred method for mapping land use and cover. Overlaying flood extent maps onto LULC maps and administrative boundaries of Sindh province uncovered significant inundation in built-up areas, particularly in districts like Kambar Shahdad Kot, Jacobabad, and Larkana. Additionally, irrigated herbaceous crops emerged as notably impacted, highlighting agricultural vulnerability to floods. This integrated approach provided comprehensive insights into flood impacts, aiding targeted mitigation and urban planning strategies in the region.

4.2 Contributions

Optical sensors are popular in flood mapping and monitoring studies as they can provide high-resolution, multispectral data that captures detailed surface characteristics, allowing for the identification and analysis of various land cover types affected by flooding. Most of the studies employ the Modified Normalized Difference Water Index (MNDWI) in delineating water features, especially in the context of river basin flooding. One significant limitation of the MNDWI index is its reduced sensitivity to areas with sparse vegetation or complex urban areas and arid landscapes, which can lead to less accurate identification of water bodies, especially in cases of flooding. Such scenarios are commonplace in regions like Sindh which can benefit from use of the indices such as the Desert Flood Index (DFI). This study demonstrated that DFI outperforms the more established Modified Normalized Difference Water Index (MNDWI) by minimizing the influence of non-water pixels. Familiarity with the application of the index is crucial to regions like Sindh.

This study also introduces the application of an exhaustive grid search method for determining threshold values applicable to both SAR and optical imagery indices for flood extent mapping,

with a focus on the Sindh region. This approach is characterized by its easier reproducibility, flexibility, and capability for precise threshold determination, suited to the dynamic and complex nature of flood conditions. Through the exhaustive grid search, the method evaluates each potential threshold value, offering critical insights that contribute to the refinement of flood mapping methodologies. Additionally, this technique demonstrates resilience to the influence of clouds on flood detection, primarily by leveraging the agreement with SAR-based indices, which are unaffected by cloud cover or adverse weather conditions. This approach hence effectively mitigates the impact of clouds and shadows and is particularly beneficial for flood studies in regions prone to heavy cloud cover during monsoon seasons, where optical images may be compromised.

4.3 Limitations

Despite this study's diligent efforts to understand the impacts of flooding on various land cover classes in Sindh province using remote sensing-based flood maps and LULC datasets, it is not without its limitations. These limitations are inherent to the nature of remote sensing analysis, data availability, and methodological constraints, and outline the aspects that require improvement and consideration in future research.

The cloud cover in Sentinel-2 images limits the effectiveness of flood extent mapping, as it frequently occurs during rainfall-induced flooding conditions. This limitation was partially addressed by focusing on smaller, cloud-free regions for thresholding analysis. The presence of cloud cover often necessitates the removal of a substantial amount of optical data, which could otherwise obscure the flood analysis if included in larger areas.

The 5 to 6-day revisit cycles of the Sentinel satellites are insufficient to capture the rapid changes in flood events fully. This limitation means that the dynamics of flood progression or recession, which can alter significantly within a single day, may not be completely observed. Inaccuracies arise particularly at the edges of images, where flood extents might not uniformly represent the current state of flooding by the time new images are acquired.

Discrepancies arise when combining data from Sentinel-1 and Sentinel-2, due to their different acquisition times, sensor characteristics, and processing levels. These discrepancies can introduce errors in flood mapping, especially when data from these sensors are integrated to create a comprehensive flood extent map.

The thresholds used for flood detection were not precisely calibrated across the entire region, which could have improved using an exhaustive grid search technique despite its high computational demands. Applying a single threshold value uniformly assumes consistent conditions across diverse geographical areas, which is often not the case due to varying climate, land cover, and hydrological characteristics.

The reliance on the FAO's LULC dataset for 2021 introduces uncertainties due to unclear data development and processing details, and the lack of post-processing in the study's datasets. These factors create discrepancies between the LULC maps used and potentially affect the accuracy of the flood impact assessments. Additionally, mismatches between the timing of image acquisitions and actual flooding events further complicate the accuracy of these assessments.

Google Earth Engine (GEE) has been extensively utilized for image processing, primarily due to the large area covered by Sindh and GEE's computational capability in processing vast amounts of geospatial data. However, considering the extensive area covered by Sindh, approximately

140,910 km², and the wide array of data collected by Sentinel images, significant challenges were presented. To address this, many computations had to be downscaled within manageable parameters in GEE which had to be performed for a small region. This necessity for downscaling raises concerns about the representativeness of the findings of the study.

4.4 Recommendations for the Future Work

The study has provided valuable insights into the spatial dynamics of flooding and its influence on land cover, emphasizing the necessity for ongoing progress in remote sensing techniques, computational resources, and data availability. Subsequent recommendations are proposed to enhance future flood assessment studies and tackle the limitations identified in this research.

1. Google Earth Engine possessed significant challenges during the study in processing such a large area with state-of-the-art techniques. Future studies should explore alternatives such as increased processing capacity or other processing systems so that machine learning models can be improved.

2. The major setback regarding the use of optical data for thresholding the SAR-based flooding extent was the presence of clouds. This can be abated by increasing reliance on datasets that are less affected by clouds and bad weather.

3. Different environmental conditions can lead to non-uniform backscatter values within a single ROI, causing a single threshold value to be ineffective across the whole region. By dividing the ROI into smaller segments, each can be analyzed under a threshold value that is optimized for its specific conditions. This method enhances the overall precision of

the thresholding process as each smaller ROI is more likely to exhibit homogenous characteristics.

4. In this study, K-means classification and Random Forest were employed to develop the land use/Land cover dataset. Future studies should delve into the application of advanced machine learning and deep learning techniques to improve the accuracy of flood extent mapping and land cover classification. These techniques can better handle the complexities of mixed pixels and cloud shadows in satellite imagery.

5. The remote sensing data can be integrated with hydrological models to predict flood behaviors accurately.

6. To overcome the limitations associated with reliance on secondary LULC datasets, future research should prioritize the collection of ground-truth data, and allowing open access to such datasets will encourage researchers to pursue the topic. Continued monitoring is crucial to generate accurate LULC maps.

7. To assess the impact of floods on livelihood, future studies should incorporate socio-economic datasets such as housing conditions, and income levels. This will allow quantification of direct and indirect costs of floods such as lost income and housing displacement, thereby providing a more detailed assessment of the overall economic damage.

4.5 Recommendations for Water Management Authorities

This study highlights the importance of the approaches used for analyzing historical flooding. Employing such approaches is critical for precisely identifying flood-prone locations, allowing

for the ideation and implementation of targeted flood management frameworks and solutions. This will lead to collaborative and informed approaches to decrease the effect of floods, and thus increase the resilience of people and ecosystems to such natural disasters.

1. Based on the findings of this study, it is recommended to upgrade or construct flood protection infrastructure. Such infrastructures are vital for protecting vulnerable communities from the negative consequences of flooding.

2. The findings of this study should be used to redesign and improve drainage systems to improve the ability of existing systems to control floodwaters more effectively, hence lowering the danger of flooding.

3. Water management authorities should conduct ongoing studies to detect and analyze flood trends for a more nuanced understanding of changing trends, which is required for evolving flood management methods to future concerns.

4. The identification of flood-prone areas can be used to develop policies regarding land use in these locations. Such policies act as a preventative strategy, reducing communities' susceptibility to flooding risks.

5. Communities and stakeholders should be educated on the construction of flood-resistant infrastructures, thereby reducing vulnerabilities in the future.

6. The flood extent maps can be integrated with Digital Elevation Models (DEM) to determine the suitable areas for the construction of detention basins, which aid in faster drainage during flood disasters, reducing the risk of water collection and subsequent damage.

7. The outcome of the study can be used for the promotion of low-impact development and green infrastructures in appropriate regions so that they help, not only mitigate flood but also contribute to water conservation and the environment, especially for residential areas.

8. The outcome of the study should be used to understand the impact of urbanization on flood risks, and thus promote planned urban development. Carefully designed urban spaces can significantly reduce the impermeable surfaces that contribute to runoff, thereby minimizing the risk of flooding.

9. Coordination between different stakeholders in the public and private sectors should be encouraged for developing and implementing comprehensive flood management strategies. This leads to more sustainable and effective solutions that leverage diverse expertise and resources.

References

Adnan, S., Ullah, K., & Gao, S. (2015). Characterization of drought and its assessment over

    Sindh, Pakistan during 1951-2010. *Journal of Meteorological Research*, *29*(5).

    https://doi.org/10.1007/s13351-015-4113-z

Ali, A. (2013). Indus basin floods: Mechanisms, impacts, and management. In *Asian*

    Development Bank.

Amatya, P., Kirschbaum, D., & Stanley, T. (2019). Use of very high-resolution optical data for

    landslide mapping and susceptibility analysis along the Karnali highway, Nepal.

    Remote Sensing, *11*(19). https://doi.org/10.3390/rs11192284

Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., & Ruello, G. (2018). Unsupervised Rapid

    Flood Mapping Using Sentinel-1 GRD SAR Images. *IEEE Transactions on*

    Geoscience and Remote Sensing, *56*(6). https://doi.org/10.1109/TGRS.2018.2797536

Annis, A., Nardi, F., Petroselli, A., Apollonio, C., Arcangeletti, E., Tauro, F., Belli, C.,

    Bianconi, R., & Grimaldi, S. (2020). UAV-DEMs for small-scale flood hazard

    mapping. *Water (Switzerland)*, *12*(6). https://doi.org/10.3390/w12061717

Anusha, N., & Bharathi, B. (2020). Flood detection and flood mapping using multi-temporal

    synthetic aperture radar and optical data. *Egyptian Journal of Remote Sensing and*

    Space Science, *23*(2). https://doi.org/10.1016/j.ejrs.2019.01.001

Baig, M. H. A., Zhang, L., Wang, S., Jiang, G., Lu, S., & Tong, Q. (2013). COmparison of

    MNDWI and DFI for water mapping in flooding season. *International Geoscience and*

    Remote Sensing Symposium (IGARSS). https://doi.org/10.1109/IGARSS.2013.6723425

Belabid, N., Zhao, F., Brocca, L., Huang, Y., & Tan, Y. (2019). Near-real-time flood

    forecasting based on satellite precipitation products. *Remote Sensing*, *11*(3).

    https://doi.org/10.3390/rs11030252

Berg, P., Pham, M. T., & Courty, N. (2022). Self-Supervised Learning for Scene Classification

    in Remote Sensing: Current State of the Art and Perspectives. In *Remote Sensing* (Vol.

    14, Issue 16). https://doi.org/10.3390/rs14163995

Bioresita, F., Puissant, A., Stumpf, A., & Malet, J. P. (2018). A method for automatic and

    rapid mapping of water surfaces from Sentinel-1 imagery. *Remote Sensing*, *10*(2).

    https://doi.org/10.3390/rs10020217

Bolanos, S., Stiff, D., Brisco, B., & Pietroniro, A. (2016). Operational surface water detection

    and monitoring using Radarsat 2. *Remote Sensing*, *8*(4).

    https://doi.org/10.3390/rs8040285

Bui, D. H., & Mucsi, L. (2021). From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal landsat data, a random forest classifier, and decision rules. *Remote Sensing*, *13*(9). https://doi.org/10.3390/rs13091700

Chen, Z., Wagner, M., Das, J., Doe, R. K., & Cerveny, R. S. (2021). Data-driven approaches for tornado damage estimation with unpiloted aerial systems. *Remote Sensing*, *13*(9). https://doi.org/10.3390/rs13091669

Cian, F., Marconcini, M., & Ceccato, P. (2018). Normalized Difference Flood Index for rapid flood mapping: Taking advantage of EO big data. *Remote Sensing of Environment*, *209*. https://doi.org/10.1016/j.rse.2018.03.006

Clement, M. A., Kilsby, C. G., & Moore, P. (2018). Multi-temporal synthetic aperture radar flood mapping using change detection. *Journal of Flood Risk Management*, *11*(2), 152–168. https://doi.org/10.1111/jfr3.12303

Cohen, S., Raney, A., Munasinghe, D., Derek Loftis, J., Molthan, A., Bell, J., Rogers, L., Galantowicz, J., Robert Brakenridge, G., Kettner, A., Huang, Y. F., & Tsang, Y. P. (2019). The Floodwater Depth Estimation Tool (FwDET v2.0) for improved remote sensing analysis of coastal flooding. *Natural Hazards and Earth System Sciences*, *19*(9). https://doi.org/10.5194/nhess-19-2053-2019

Dasari, K., Anjaneyulu, L., Jayasri, P. V., & Prasad, A. V. V. (2016). Importance of speckle

   filtering in image classification of SAR data. *2015 International Conference on

   Microwave, Optical and Communication Engineering, ICMOCE 2015*.

   https://doi.org/10.1109/ICMOCE.2015.7489764

de Lima, R. P., & Marfurt, K. (2020). Convolutional neural network for remote-sensing scene

   classification: Transfer learning analysis. *Remote Sensing*, *12*(1).

   https://doi.org/10.3390/rs12010086

Food and Agriculture Organization of the United Nations. (2023). Land cover (Sindh

   Province, Pakistan - 10m). FAO - Pakistan.

Frappart, F., Seyler, F., Martinez, J. M., León, J. G., & Cazenave, A. (2005). Floodplain water

   storage in the Negro River basin estimated from microwave remote sensing of

   inundation area and water levels. *Remote Sensing of Environment*, *99*(4).

   https://doi.org/10.1016/j.rse.2005.08.016

Gebrehiwot, A., Hashemi-Beni, L., Thompson, G., Kordjamshidi, P., & Langan, T. E. (2019).

   Deep convolutional neural network for flood extent mapping using unmanned aerial

   vehicles data. *Sensors (Switzerland)*, *19*(7). https://doi.org/10.3390/s19071486

Ghaffarian, S., Kerle, N., & Filatova, T. (2018). Remote sensing-based proxies for urban

    disaster risk management and resilience: A review. In *Remote Sensing* (Vol. 10, Issue

    11). https://doi.org/10.3390/rs10111760

Giustarini, L., Hostache, R., Matgen, P., Schumann, G. J. P., Bates, P. D., & Mason, D. C.

    (2013). A change detection approach to flood mapping in Urban areas using

    TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(4).

    https://doi.org/10.1109/TGRS.2012.2210901

Gupta, P. K., Dubey, A. K., & Pradhan, R. (2022). *Pakistan Flood of 2022: Assessment using*

    suite of satellite sensors and hydrological modelling. https://doi.org/10.21203/rs.3.rs-

    2257969/v1

Hamidi, E., Peter, B. G., Munoz, D. F., Moftakhari, H., & Moradkhani, H. (2023). Fast Flood

    Extent Monitoring With SAR Change Detection Using Google Earth Engine. *IEEE*

    Transactions on Geoscience and Remote Sensing, *61*.

    https://doi.org/10.1109/TGRS.2023.3240097

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep

    learning benchmark for land use and land cover classification. *IEEE Journal of*

    Selected Topics in Applied Earth Observations and Remote Sensing, *12*(7).

    https://doi.org/10.1109/JSTARS.2019.2918242

Hidayah, E., Pranadiarso, T., Halik, G., Indarto, I., Lee, W. K., & Maruf, M. F. (2022). Flood

    mapping based on open-source remote sensing data using an efficient band

    combination system. *Acta Geographica Slovenica*, *62*(3).

    https://doi.org/10.3986/AGS10598

Hirabayashi, Y., Alifu, H., Yamazaki, D., Imada, Y., Shiogama, H., & Kimura, Y. (2021).

    Anthropogenic climate change has changed frequency of past flood during 2010-2013.

    Progress in Earth and Planetary Science, *8*(1). https://doi.org/10.1186/s40645-021-

    00431-w

Houze, R. A., Rasmussen, K. L., Medina, S., Brodzik, S. R., & Romatschke, U. (2011).

    Anomalous atmospheric events leading to the summer 2010 floods in Pakistan. *Bulletin*

    of the American Meteorological Society, *92*(3).

    https://doi.org/10.1175/2010BAMS3173.1

Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural

    networks for the scene classification of high-resolution remote sensing imagery.

    Remote Sensing, *7*(11). https://doi.org/10.3390/rs71114680

Huang, W., Wang, D., & Ratilal, P. (2016). Diel and spatial dependence of humpback song

    and non-song vocalizations in fish spawning ground. *Remote Sensing*, *8*(9).

    https://doi.org/10.3390/rs8090712

Hütt, C., Koppe, W., Miao, Y., & Bareth, G. (2016). Best accuracy land use/land cover

  (LULC) classification to derive crop types using multitemporal, multisensor, and

  multi-polarization SAR satellite images. *Remote Sensing*, *8*(8).

  https://doi.org/10.3390/rs8080684

Inoue, Y., Peñuelas, J., Miyata, A., & Mano, M. (2008). Normalized difference spectral

  indices for estimating photosynthetic efficiency and capacity at a canopy scale derived

  from hyperspectral and CO2 flux measurements in rice. *Remote Sensing of*

  Environment, *112*(1). https://doi.org/10.1016/j.rse.2007.04.011

Irwin, K., Braun, A., Fotopoulos, G., Roth, A., & Wessel, B. (2018). Assessing single-

  polarization and dual-polarization TerraSAR-X data for surface water monitoring.

  Remote Sensing, *10*(6). https://doi.org/10.3390/rs10060949

Ishaque, W., Tanvir, R., & Mukhtar, M. (2022). Climate Change and Water Crises in Pakistan:

  Implications on Water Quality and Health Risks. *Journal of Environmental and Public*

  Health, *2022*. https://doi.org/10.1155/2022/5484561

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive

  change detection method for updating the National Land Cover Database to circa 2011.

  Remote Sensing of Environment, *132*. https://doi.org/10.1016/j.rse.2013.01.012

Kapilaratne, R. G. C. J., & Kaneta, S. (2020). Towards an automated flood area extraction

from high resolution satellite images. *International Archives of the Photogrammetry,*

Remote Sensing and Spatial Information Sciences - ISPRS Archives, *43*(B3).

https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-97-2020

Karim, M. F., & Zhang, X. (2021). Analysis of vegetative cover vulnerability in rohingya

refugee camps of bangladesh utilizing landsat and per capita greening area (Pcga)

datasets. *Remote Sensing*, *13*(23). https://doi.org/10.3390/rs13234922

Kazi, A. (2014). A review of the assessment and mitigation of floods in Sindh, Pakistan.

Natural Hazards, *70*(1). https://doi.org/10.1007/s11069-013-0850-4

Khokhar, M. A. H., Javed, S., Awan, H. B. H., Yousaf, I., Iqbal, A., & Waseem, L. A. (2021).

Damage Assessment and Spatio-Temporal Mapping Using Object Based Image

Analysis Technique on Floods-2015 in Southern Punjab, Pakistan. *European Journal*

of Geosciences, *3*(1). https://doi.org/10.34154/2021-ejgs-0013/euraass

Kitajima, N., Seto, R., Yamazaki, D., Zhou, X., Ma, W., & Kanae, S. (2021). Potential of a sar

small-satellite constellation for rapid monitoring of flood extent. *Remote Sensing*,

13(10). https://doi.org/10.3390/rs13101959

Kumar, H., Karwariya, S. K., & Kumar, R. (2022). Google Earth Engine-Based Identification

of Flood Extent and Flood-Affected Paddy Rice Fields Using Sentinel-2 MSI and

Sentinel-1 SAR Data in Bihar State, India. *Journal of the Indian Society of Remote Sensing*, *50*(5). https://doi.org/10.1007/s12524-021-01487-3

Langford, Z. L., Kumar, J., Hoffman, F. M., Breen, A. L., & Iversen, C. M. (2019). Arctic vegetation mapping using unsupervised training datasets and convolutional neural networks. *Remote Sensing*, *11*(1). https://doi.org/10.3390/rs11010069

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. *Remote Sensing*, *6*(2). https://doi.org/10.3390/rs6020964

Li, L., Xu, T., & Chen, Y. (2016). Improved urban flooding mapping from remote sensing images using generalized regression neural network-based super-resolution algorithm. *Remote Sensing*, *8*(8). https://doi.org/10.3390/rs8080625

Liang, J., & Liu, D. (2020). A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *159*. https://doi.org/10.1016/j.isprsjprs.2019.10.017

Long, S., Fatoyinbo, T. E., & Policelli, F. (2014). Flood extent mapping for Namibia using change detection and thresholding with SAR. *Environmental Research Letters*, *9*(3), 035002. https://doi.org/10.1088/1748-9326/9/3/035002

Manzoor, A., & Adesola, R. O. (2022). Disaster in public health due to flood in Pakistan in 2022. In *Health Science Reports* (Vol. 5, Issue 6). https://doi.org/10.1002/hsr2.903

Manzoor, Z., Ehsan, M., Khan, M. B., Manzoor, A., Akhter, M. M., Sohail, M. T., Hussain, A., Shafi, A., Abu-Alam, T., & Abioui, M. (2022). Floods and flood management and its socio-economic impact on Pakistan: A review of the empirical literature. In *Frontiers in Environmental Science* (Vol. 10). https://doi.org/10.3389/fenvs.2022.1021862

Markert, K. N., Chishtie, F., Anderson, E. R., Saah, D., & Griffin, R. E. (2018). On the merging of optical and SAR satellite imagery for surface water mapping applications. *Results in Physics*, *9*. https://doi.org/10.1016/j.rinp.2018.02.054

Martinis, S., Twele, A., Strobl, C., Kersten, J., & Stein, E. (2013). A multi-scale flood monitoring system based on fully automatic MODIS and terraSAR-X processing chains. *Remote Sensing*, *5*(11). https://doi.org/10.3390/rs5115598

Matheswaran, K., Alahacoon, N., Pandey, R., & Amarnath, G. (2019). Flood risk assessment in South Asia to prioritize flood index insurance applications in Bihar, India. *Geomatics, Natural Hazards and Risk*, *10*(1). https://doi.org/10.1080/19475705.2018.1500495

Mellor, A., Haywood, A., Stone, C., & Jones, S. (2013). The performance of random forests in

    an operational setting for large area sclerophyll forest classification. *Remote Sensing*,

    5(6). https://doi.org/10.3390/rs5062838

Meresa, H., Tischbein, B., & Mekonnen, T. (2022). Climate change impact on extreme

    precipitation and peak flood magnitude and frequency: observations from CMIP6 and

    hydrological models. *Natural Hazards*, *111*(3). https://doi.org/10.1007/s11069-021-

    05152-3

Moortgat, J., Li, Z., Durand, M., Howat, I., Yadav, B., & Dai, C. (2022). Deep learning

    models for river classification at sub-meter resolutions from multispectral and

    panchromatic commercial satellite imagery. *Remote Sensing of Environment*, *282*.

    https://doi.org/10.1016/j.rse.2022.113279

Morin, N., Masse, A., Sannier, C., Siklar, M., Kiesslich, N., Sayadyan, H., Faucqueur, L., &

    Seewald, M. (2021). Development and application of earth observation based machine

    learning methods for characterizing forest and land cover change in dilijan national

    park of armenia between 1991 and 2019. *Remote Sensing*, *13*(15).

    https://doi.org/10.3390/rs13152942

Muhadi, N. A., Abdullah, A. F., Bejo, S. K., Mahadi, M. R., & Mijic, A. (2020). The use of

    LiDAR-derived DEM in flood applications: A review. In *Remote Sensing* (Vol. 12,

    Issue 14). https://doi.org/10.3390/rs12142308

Munawar, H. S., Hammad, A. W. A., & Waller, S. T. (2022). Remote Sensing Methods for

  Flood Prediction: A Review. In *Sensors* (Vol. 22, Issue 3).

  https://doi.org/10.3390/s22030960

Nanditha, J. S., Kushwaha, A. P., Singh, R., Malik, I., Solanki, H., Chuphal, D. S., Dangar, S.,

  Mahto, S. S., Vegad, U., & Mishra, V. (2023). The Pakistan Flood of August 2022:

  Causes and Implications. *Earth's Future*, *11*(3). https://doi.org/10.1029/2022EF003230

Nasirzadehdizaji, R., Akyuz, D. E., & Cakir, Z. (2019). Flood mapping and permanent water

  bodies change detection using sentinel sar data. *International Archives of the*

  Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives,

  42(4/W18), 797–801. https://doi.org/10.5194/isprs-archives-XLII-4-W18-797-2019

Nasr, A., Abdelkareem, M., & Moubark, K. (2022). Integration of remote sensing and GIS for

  mapping flash flood hazards, Wadi Queih, Egypt. *SVU-International Journal of*

  Agricultural Sciences, *4*(4). https://doi.org/10.21608/svuijas.2023.186199.1264

Nex, F., Duarte, D., Tonolo, F. G., & Kerle, N. (2019). Structural building damage detection

  with deep learning: Assessment of a state-of-the-art CNN in operational conditions.

  Remote Sensing, *11*(23). https://doi.org/10.3390/rs11232765

Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., & Song, C. (2017). A regional-scale

    assessment of Himalayan glacial lake changes using satellite observations from 1990 to

    2015. *Remote Sensing of Environment*, *189*. https://doi.org/10.1016/j.rse.2016.11.008

Notti, D., Giordan, D., Caló, F., Pepe, A., Zucca, F., & Galve, J. P. (2018). Potential and

    limitations of open satellite data for flood mapping. *Remote Sensing*, *10*(11).

    https://doi.org/10.3390/rs10111673

Otto, F. E. L., Zachariah, M., Saeed, F., Siddiqi, A., Kamil, S., Mushtaq, H., Arulalan, T.,

    AchutaRao, K., Chaithra, S. T., Barnes, C., Philip, S., Kew, S., Vautard, R., Koren, G.,

    Pinto, I., Wolski, P., Vahlberg, M., Singh, R., Arrighi, J., … Clarke, B. (2023). Climate

    change increased extreme monsoon rainfall, flooding highly vulnerable communities in

    Pakistan. *Environmental Research: Climate*, *2*(2). https://doi.org/10.1088/2752-

    5295/acbfd5

Porter, J., & Demeritt, D. (2012). Flood-risk management, mapping, and planning:

    Theinstitutional politics of decision support in England. *Environment and Planning A*,

    44(10). https://doi.org/10.1068/a44660

Qamer, F. M., Ahmad, B., Hussain, A., Salman, A., Muhammad, S., Nawaz, M., Shrestha, S.,

    Iqbal, B., & Thapa, S. (2022). *The 2022 Pakistan floods: Assessment of crop losses in

    Sindh Province using satellite data*. https://doi.org/10.53055/ICIMOD.1015

Rahman, M. S., & Di, L. (2020). A systematic review on case studies of remote-sensing-based

    flood crop loss assessment. In *Agriculture (Switzerland)* (Vol. 10, Issue 4).

    https://doi.org/10.3390/agriculture10040131

Rahman, M. S., Di, L., Yu, E., Lin, L., & Yu, Z. (2021). Remote Sensing Based Rapid

    Assessment of Flood Crop Damage Using Novel Disaster Vegetation Damage Index

    (DVDI). *International Journal of Disaster Risk Science*, *12*(1).

    https://doi.org/10.1007/s13753-020-00305-7

Rajpar, H., Zhang, A., Razzaq, A., Mehmood, K., Pirzado, M. B., & Hu, W. (2019).

    Agricultural land abandonment and farmers' perceptions of land use change in the

    indus plains of Pakistan: A case study of Sindh province. *Sustainability (Switzerland)*,

    11(17). https://doi.org/10.3390/su11174663

Saini, R., & Ghosh, S. K. (2018). CROP CLASSIFICATION ON SINGLE DATE

    SENTINEL-2 IMAGERY USING RANDOM FOREST AND SUPPOR VECTOR

    MACHINE. *The International Archives of the Photogrammetry, Remote Sensing and*

    Spatial Information Sciences, *XLII–5*. https://doi.org/10.5194/isprs-archives-xlii-5-683-

    2018

Sajjad, A., Lu, J., Aslam, R. W., & Ahmad, M. (2023). *Flood Disaster Mapping Using*

    Geospatial Techniques: A Case Study of the 2022 Pakistan Floods.

    https://doi.org/10.3390/ecws-7-14312

Sajjad, A., Lu, J., Chen, X., Chisenga, C., Saleem, N., & Hassan, H. (2020). Operational

    monitoring and damage assessment of riverine flood-2014 in the lower chenab plain,

    Punjab, Pakistan, using remote sensing and gis techniques. *Remote Sensing*, *12*(4).

    https://doi.org/10.3390/rs12040714

Santaga, F. S., Agnelli, A., Leccese, A., & Vizzari, M. (2021). Using sentinel-2 for simplifying

    soil sampling and mapping: Two case studies in Umbria, Italy. *Remote Sensing*,

    13(17). https://doi.org/10.3390/rs13173379

Scarth, P., Armston, J., Lucas, R., & Bunting, P. (2019). A structural classification of

    Australian vegetation using ICESat/GLAS, ALOS PALSAR, and Landsat sensor data.

    Remote Sensing, *11*(2). https://doi.org/10.3390/rs11020147

Schumann, G. J. P. (2015). Preface: Remote sensing in flood monitoring and management. In

    Remote Sensing (Vol. 7, Issue 12). https://doi.org/10.3390/rs71215871

Shastry, A., Carter, E., Coltin, B., Sleeter, R., McMichael, S., & Eggleston, J. (2023). Mapping

    floods from remote sensing data and quantifying the effects of surface obstruction by

    clouds and vegetation. *Remote Sensing of Environment*, *291*.

    https://doi.org/10.1016/j.rse.2023.113556

Shehzad, K. (2023). Extreme flood in Pakistan: Is Pakistan paying the cost of climate change?

A short communication. *Science of the Total Environment*, *880*.

https://doi.org/10.1016/j.scitotenv.2023.162973

Silva-Cardoza, A. I., Vega-Nieva, D. J., Briseño-Reyes, J., Briones-Herrera, C. I., López-

Serrano, P. M., Corral-Rivas, J. J., Parks, S. A., & Holsinger, L. M. (2022). Evaluating

a New Relative Phenological Correction and the Effect of Sentinel-Based Earth Engine

Compositing Approaches to Map Fire Severity and Burned Area. *Remote Sensing*,

14(13). https://doi.org/10.3390/rs14133122

Sonobe, R. (2019). Parcel-based crop classification using multi-temporal TerraSAR-X dual

polarimetric data. *Remote Sensing*, *11*(10). https://doi.org/10.3390/rs11101148

Suarez, E. L., & Meeroff, D. E. (2021). *Interdisciplinary approach to flood risk and the*

Consequence of Flooding. https://doi.org/10.3311/floodrisk2020.11.19

Sublime, J., & Kalinicheva, E. (2019). Automatic post-disaster damage mapping using deep-

learning techniques for change detection: Case study of the Tohoku tsunami. *Remote*

Sensing, *11*(9). https://doi.org/10.3390/rs11091123

Sun, Y., Liu, C., Du, X., Yang, F., Yao, Y., Soomro, S. e. hyde., & Hu, C. (2022). Urban

storm flood simulation using improved SWMM based on K-means clustering of

parameter samples. *Journal of Flood Risk Management*, *15*(4).

https://doi.org/10.1111/jfr3.12826

Sunar, A. F., Yagmur, N., & Dervisoglu, A. (2019). Flood analysis with remote sensing data

– A case study: Maritsa River, Edirne. *International Archives of the*

Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives,

42(3/W8). https://doi.org/10.5194/isprs-archives-XLII-3-W8-497-2019

Tazmul Islam, M., & Meng, Q. (2022). An exploratory study of Sentinel-1 SAR for rapid

urban flood mapping on Google Earth Engine. *International Journal of Applied Earth*

Observation and Geoinformation, *113*, 103002.

https://doi.org/10.1016/J.JAG.2022.103002

Ticehurst, C. J., Dyce, P., & Guerschman, J. P. (2009). Using passive microwave and optical

remote sensing to monitor flood inundation in support of hydrologic modelling. *18th*

World IMACS Congress and MODSIM 2009 - International Congress on Modelling

and Simulation: Interfacing Modelling and Simulation with Mathematical and

Computational Sciences, Proceedings.

Tzepkenlis, A., Marthoglou, K., & Grammalidis, N. (2023). Efficient Deep Semantic

Segmentation for Land Cover Classification Using Sentinel Imagery. *Remote Sensing*,

15(8). https://doi.org/10.3390/rs15082027

Venkatappa, M., Sasaki, N., Han, P., & Abe, I. (2021). Impacts of droughts and floods on

croplands and crop production in Southeast Asia – An application of Google Earth

Engine. *Science of the Total Environment*, *795*.

https://doi.org/10.1016/j.scitotenv.2021.148829

Viana-Soto, A., Aguado, I., Salas, J., & García, M. (2020). Identifying post-fire recovery

trajectories and driving factors using landsat time series in fire-prone mediterranean

pine forests. *Remote Sensing*, *12*(9). https://doi.org/10.3390/RS12091499

Wang, Y., Li, G., Ding, J., Guo, Z., Tang, S., Liu, R., & Chen, J. (2016). A combined GLAS

and MODIS estimation of the global distribution of mean forest canopy height. *Remote

Sensing of Environment*, *174*. https://doi.org/10.1016/j.rse.2015.12.005

Wang, J., Li, K., Hao, L., Xu, C., Liu, J., Qu, Z., Yan, X., Sajjad, M. M., & Sun, Y. (2023).

Disaster mapping and assessment of Pakistan's 2022 mega-flood based on multi-source

data-driven approach. Natural Hazards. https://doi.org/10.1007/s11069-023-06337-8

Waqas, M. A. (2022). Pakistan's floods flow from climate injustice. In *Science* (Vol. 378,

Issue 6619). https://doi.org/10.1126/science.ade8490

Ward, D. P., Petty, A., Setterfield, S. A., Douglas, M. M., Ferdinands, K., Hamilton, S. K., &

Phinn, S. (2014). Floodplain inundation and vegetation dynamics in the Alligator

Rivers region (Kakadu) of northern Australia assessed using optical and radar remote

sensing. *Remote Sensing of Environment*, *147*. https://doi.org/10.1016/j.rse.2014.02.009

Waseem, H. Bin, & Rana, I. A. (2023). Floods in Pakistan: A state-of-the-art review. *Natural Hazards Research*. https://doi.org/10.1016/j.nhres.2023.06.005

Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A., & White, E. (2019). Individual tree-crown detection in rgb imagery using semi-supervised deep learning neural networks. *Remote Sensing*, *11*(11). https://doi.org/10.3390/rs11111309

White, L., Brisco, B., Dabboor, M., Schmitt, A., & Pratt, A. (2015). A collection of SAR methodologies for monitoring wetlands. In *Remote Sensing* (Vol. 7, Issue 6). https://doi.org/10.3390/rs70607615

Wulder, M. A., Li, Z., Campbell, E. M., White, J. C., Hobart, G., Hermosilla, T., & Coops, N. C. (2018). A national assessment of wetland status and trends for Canada's forested ecosystems using 33 years of earth observation satellite data. *Remote Sensing*, *10*(10). https://doi.org/10.3390/rs10101623

Xie, Z., Chen, Y., Lu, D., Li, G., & Chen, E. (2019). Classification of land cover, forest, and tree species classes with Ziyuan-3 multispectral and stereo data. *Remote Sensing*, *11*(2). https://doi.org/10.3390/rs11020164

Yang, Y., Yang, D., Wang, X., Zhang, Z., & Nawaz, Z. (2021). Testing accuracy of land cover classification algorithms in the qilian mountains based on gee cloud platform. *Remote Sensing*, *13*(24). https://doi.org/10.3390/rs13245064

Yuan, X., Yang, K., Sun, J., Wang, Y., He, Y. Y., Zou, M. J., & Jiang, Y. Z. (2023). Why was Pakistan extreme precipitation stronger in 2022 than in 2010? *Advances in Climate Change Research*, *14*(6), 913–920. https://doi.org/10.1016/j.accre.2023.11.016

Zhang, F., Zhu, X., & Liu, D. (2014). Blending MODIS and Landsat images for urban flood mapping. *International Journal of Remote Sensing*, *35*(9). https://doi.org/10.1080/01431161.2014.903351

Zhang, J., Liu, K., & Wang, M. (2023). Flood detection using Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage and extreme precipitation data. *Earth System Science Data*, *15*(2). https://doi.org/10.5194/essd-15-521-2023

Curriculum Vitae

Alina Shrestha

Date: May 2024

Graduate College

University of Nevada, Las Vegas

Department of Civil and Environmental Engineering and Construction

Email address: alina.stha222@gmail.com

Education

University of Nevada, Las Vegas

M.S.E. Civil and Environmental Engineering, May 2024

Tribhuvan University, Kathmandu, Nepal

Bachelor's Degree in Civil Engineering, 2015

Thesis Title: Evaluation of the Extent of Floods in 2022 in Different Land Cover and Land Use Classes in Sindh Province, Pakistan Using Sentinel-1 and Sentinel-2 Imageries

Thesis Examination Committee:

 Chairperson, Dr. Sajjad Ahmad

 Co-chairperson, Dr. Haroon Sahotra

 Committee Member, Dr. Marie-Odile Fortier

 Graduate College Representative, Dr. Ashok Singh