ASSESSING HAILSTORM DAMAGES IN CROPS USING MULTI-TEMPORAL REMOTE SENSING DATA AND MACHINE LEARNING SOLUTIONS
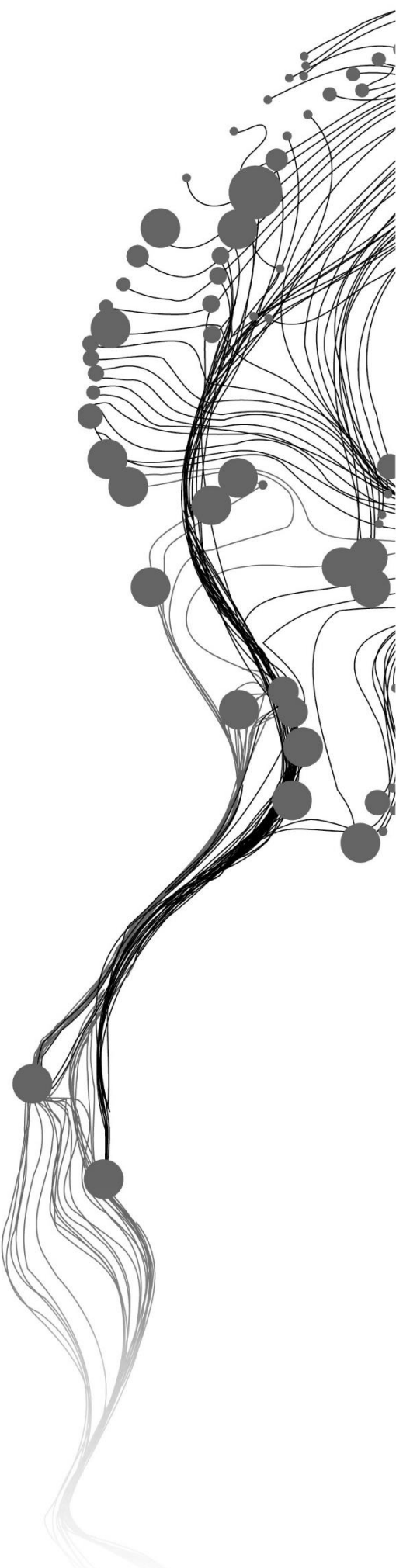
RISHI PAREEK
[June 2023]

SUPERVISORS:
Dr. ing. Hossein Aghababaei
Dr. ir. Wietske Bijker

ASSESSING HAILSTORM DAMAGES IN CROPS USING MULTI-TEMPORAL REMOTE SENSING DATA AND MACHINE LEARNING SOLUTIONS

RISHI PAREEK
Enschede, The Netherlands, June 2023

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Spatial Engineering.

SUPERVISORS:
Dr. ing. Hossein Aghababaei
Dr. ir. Wietske Bijker

THESIS ASSESSMENT BOARD:
Prof. dr. ir. Claudio Persello (Chair)

ABSTRACT

Extreme weather events such as hailstorms pose a substantial risk due to their unpredictable nature and potential to damage infrastructure and agricultural yield severely. In India, these hailstorms have emerged as a significant concern due to the extensive damage they inflict on crops. Evaluating this damage in terms of location and severity is pivotal in formulating appropriate strategies to mitigate crop loss and adequately compensate the affected farmers. However, current assessment methods, predominantly field surveys and visual inspections are expensive, labour-intensive, and deficient in transparency and efficiency.

This research aimed to create a Machine Learning (ML) model using multi-temporal Remote Sensing (RS) data to quantify the crop damage caused by hailstorms. The model was trained and tested using in-situ data collected from the Kathumar block in the Alwar district of Rajasthan province, India. We explored the performance of both Pixel Based Classification (PBC) and Object Based Classification (OBC) techniques using Random Forest (RF) model. The analysis utilized open-source Sentinel datasets, particularly the Sentinel-1 Synthetic Aperture Radar (S1) and the Sentinel-2 Multispectral Instrument (S2). Following this, we combined S1 and S2 data, prioritizing variables based on their importance for the model based on Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG). We extracted a total of 15 features, inclusive of polarimetric features from the pre and post-hail event S1 data, and the differences in these features were used as input variables for the RF model.

Furthermore, we extracted the original S2 channels, excluding B1 and B10, along with 19 vegetation indices, and their difference band was used as input variables. The model underwent training and testing procedures employing both PBC and OBC methods. For the PBC, 40,157 in-situ data samples derived from 3,394 polygons were utilized. The same polygons were also used in the training and testing processes for the OBC approach. To ensure a balanced representation across three damage categories, i.e. low (0-33%), moderate (33-50%), and high (50-75%) damage, we allocated 70% of the data for training and 30% for testing using stratified random sampling. The research also encompassed stakeholder analysis and the impacts of the intervention on various stakeholders. This was achieved by interviewing a farmer from the same region and reviewing relevant literature and news articles.

The research results suggest that the PBC approach produced significant results when using combined S1 and S2 data. It accurately identified damage classes with F1 scores of 0.97, 0.93, and 0.94 for low, moderate, and high damage classes, respectively, after majority classification. In contrast, when using a single dataset for damage assessment, the OBC approach with the S1 data yielded the highest accuracy, with F1 scores of 0.96, 0.90, and 0.85 for the low, moderate, and high damage classes, respectively. The study further revealed that the accuracy in PBC was enhanced when both S2 and S1 data were combined, compared to their individual usage. However, for the OBC approach, where statistical parameters were employed, a marginal decrease in accuracy was noted when compared to the exclusive use of the SAR data.

This research has social significance as it attempts to mitigate the wicked problem associated with current field-based crop damage evaluations and their effects on the farming community. This impact can lead to financial instability and, in extreme cases, suicide. The current process tends to perpetuate a sense of inequality and unfairness. Consequently, the broader aim of this research is to present a proof of concept aimed at enhancing the current scenario using RS based approach for crop damage assessment. Stakeholder analysis indicates that the developed model has the potential to act as an instructive geoprocessing tool, which could significantly reduce the necessity for field surveys and accelerate the evaluation process. Moreover, it offers a robust solution to the lack of transparency in damage assessment categorization.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI - Artificial Intelligence
BoR - Board of Revenue
DMR - Disaster Management, Relief & Civil Defence Department
DpRVI - Dual-pol Radar Vegetation Index
DVI - Difference Vegetation Index
GEMI - The Global Environmental Monitoring Index
GNDVI - Green Normalized Difference Vegetation Index
HDZ - Homogeneous Damage Zones
IPCC - The Intergovernmental Panel on Climate Change
IPVI - Infrared Percentage Vegetation Index
IRECI - Inverted Red-Edge Chlorophyll Index
MCARI - Modified Chlorophyll Absorption in Reflectance Index
ML - Machine Learning
MDA - Mean Decrease Accuracy
MDG - Mean Decrease Gini
MSAVI - Modified Soil-Adjusted Vegetation Index
MSI - Multi-Spectral Instrument
MTCI - Meris Terrestrial Chlorophyll Index
NDI45 - Normalized Difference Index using Bands 4 and 5
NDVI - Normalized Difference Vegetation Index
NGOs - Non-Governmental Organizations
OBC-Object Based Classification
OOB - Out-of-Bag
PBC- Pixel Based Classification
PSSRA - Pigment Specific Simple Ratio A
PVI - Perpendicular Vegetation Index
REIP - Red Edge Inflection Point
RF - Random Forest
RoR - Revenue Record or Registry of Records
RS - Remote Sensing
RVI - Ratio Vegetation Index
S1 - Sentinel-1
S2 - Sentinel 2
S2REP - Sentinel-2 Red-Edge Position
SAVI - Soil-Adjusted Vegetation Index
SAR - Synthetic Aperture Radar
SLC - Single Look Complex
TNDVI - Transformed Normalized Difference Vegetation Index
TSAVI - Transformed Soil-Adjusted Vegetation Index
WDVI - Weighted Difference Vegetation Index

1.   INTRODUCTION

1.1.   Background

The global surface temperature has experienced a significant upward trend from 2011 to 2020, showing an increase of 1.1°C compared to the period between 1850 and 1900 (Portner et al., 2022). This temperature rise is largely attributed to the greenhouse gas emissions caused by human activities (NASA, 2023). The planet's temperature is influenced by the presence of carbon dioxide in the atmosphere, leading to climate change. Within a span of fewer than 200 years, human actions have significantly increased the concentration of carbon dioxide in the atmosphere by 50% (NASA, 2023). The graph presented in Figure 1 depicts the variations in global surface temperature from 1880 to 2020 compared to the long-term average observed from 1951 to 1980 and atmospheric CO2 levels measured from 1960 to 2020 (NASA, 2023).
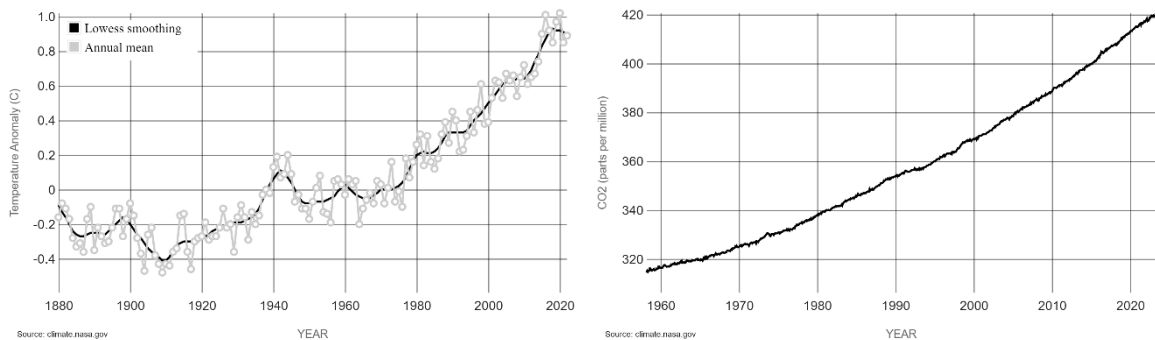
Figure 1. Change in global surface temperature (Left) from 1880 to 2020 and atmospheric CO2 levels (Right) measured from 1960 to 2020 (Source: NASA, 2023)

This man-made climate change is triggering changes in weather and climate extremes across the globe, resulting in significant harm to both the natural world and human societies (Portner et al., 2022). Farmers, policymakers, and government authorities are constantly concerned about the state of the climate and its effects on crop production. Changes in the world's hydrological cycle due to climate change, including increased variability in rainfall and occurrences of floods and droughts, are already affecting agriculture. These impacts have resulted in reduced yields of key crops. For instance, from 1981 to 2010, changes in precipitation led to yield decreases of 4.1% for maize, 4.5% for soybeans, 1.8% for rice, and 1.8% for wheat (Portner et al., 2022).

The risk associated with crop yield varies from region to region and is influenced by factors such as soil type, climate, farming techniques and water sources. One of these factors that has the most influence on agricultural production is the weather. In rain-fed agriculture, weather conditions can account for up to 80% of the variation in agricultural productivity (Wang et al., 2012). Climate change, including increases in frequency and intensity of weather extremes like floods, droughts, hailstorms, rainfall, wind, heat waves and continued sea level rise, have reduced food and water security, obstructing efforts to meet Sustainable Development Goals (Portner et al., 2022). Although agricultural productivity has increased, climate change has slowed this growth over the past 50 years (Portner et al., 2022).

Increasing weather and climate extremes have exposed millions of people to acute food insecurity and reduced water security (Portner et al., 2022). The most significant impacts have been observed in many locations and communities in Africa, Asia, Central and South America, Small Islands and the Arctic (Portner et al., 2022). The Intergovernmental Panel on Climate Change (IPCC) suggests that severe weather events in the Indian subcontinent may rise (Portner et al., 2022). The advancements in climate research have facilitated the prediction of weather patterns, a critical development offering significant opportunities for agricultural stakeholders. They can proactively respond to potential risks or exploit beneficial conditions by utilising climate forecast data. However, the successful prediction of inherently unpredictable meteorological events, including hailstorms, droughts, floods, and frigid temperatures, remains a considerable challenge (Wang et al., 2012). Unpredictable extreme weather conditions like hailstorms can gravely damage property and agricultural yields (Wang et al., 2012).

Hail consists of solid ice chunks originating in thunderstorm updrafts, where raindrops freeze into hailstones. These hailstones can expand as liquid water droplets freeze onto their surface. The slow freezing process allows air bubbles to escape, producing clear ice. Hailstones begin to fall when they become sufficiently large, or the thunderstorm's updraft weakens (Rao et al., 2014). Smaller hailstones with less than 2.5 cm in diameter can fall at speeds between 14 and 40 km/h, while larger ones, between 2.5 and 4.5 cm in diameter, can fall at 40 to 64 km/h (NSSL, 2022). Huge hailstones of 4.5 to 10 cm in diameter, produced by the strongest supercells, can fall at speeds ranging from 71 to 116 km/h (NSSL, 2022). There's some uncertainty in these estimates because factors like hailstone shape, melting, fall orientation, and environmental conditions can influence their fall speed. Nonetheless, hailstones exceeding 10 cm can fall at rates greater than 160 km/h (NSSL, 2022). Changing climate patterns, potentially due to rising global temperatures, could intensify hailstorms, resulting in larger ice chunks and heavier rainfall (BBC Future, 2022). Countries that experience the most destructive hailstorms include the United States, China, Russia, Italy, and India (NSSL, 2022).

Hailstorms, possessing significant impulse and kinetic energy, frequently result in substantial physical harm to agricultural yield (BBC Future, 2022). The severity and extent of crop damage hinge on various elements, such as the hailstone's size, the duration of the hailstorm, the wind speed, and the crop category (Angearu et al., 2022). The susceptibility of crops and their recovery capability after hail event depends on the period of agriculture season, development level and phenological stage of the crop (Ha et al., 2022).

Generally, hail damage to field crops is classified as (i) leaf defoliation, (ii) plant stand loss, (iii) stem amputation, and (iv) grain loss (Ha et al., 2022). Hailstorms may cause severe plant damage, including significant leaf area loss and laceration. Foliage, flowers, and fragile stem tissues are susceptible to hail damage, manifesting as bruising, shredding, defoliation, or physical mutilation (Ha et al., 2022). Some studies have shown that even small hailstones can damage plants, but the damage to fruits, flowers, leaf buds, and seedlings is not always apparent in the early stages (Leite et al., 2002). Hail may cause several problems for plants, such as the lodging and breaking of stems, the breaking of branches, the threshing of grain, the loss of leaf area, and the destruction of leaves (Fernandes et al., 2011). However, torn holes may only be visible on larger leaves (Schubert, 1991). In addition to the direct damage triggered by hail, there are indirect consequences, such as the loss of photosynthetically active areas and the facilitation of disease entry produced by the impact of hailstones on branches and leaves. These indirect consequences raise the probability of disease, rot, and insect infestation (Leite et al., 2002; Schubert, 1991; Singh et al., 2017).

Hailstorms pose a substantial challenge in India, causing considerable crop damage (Ray et al., 2016). The spatial pattern of hail events in India over 39 years (1972–2011) indicated that incidents were typically confined and occurred predominately between January to March (Rao et al., 2014).

Some districts are affected frequently more than others in all parts of India, especially in north India (Rao et al., 2014). In 2015, hailstorms, thunderstorms, heavy rain, and strong winds hit several areas of northern India, causing an estimated 106730 km² in agricultural damage across 14 provinces in India (Prabhakar et al., 2019). the primary states impacted included Rajasthan, which was affected over an area of 45,530 km², Uttar Pradesh, with damage spanning across 26,790 km², Haryana, which experienced impact over 18,750 km², Madhya Pradesh, where the affected area extended across 5,700 km², and Maharashtra, which saw damage over a region of 3,950 km², most severely affected districts were found in Haryana, Rajasthan and Madhya Pradesh (Ray et al., 2016).

Figure 2. Most affected districts by hailstorm during February -March 2015 (Source: Ray et al., 2016)

Figure 2 shows the districts most impacted by hailstorms, as identified in the research conducted by (Ray et al., 2016). After a hailstorm, it is essential to evaluate the damage in terms of its location and severity so that appropriate actions can be taken to lessen the loss of crops and compensate the impacted farmers. In India, the damage is documented through damage assessment surveys using field observation and sometimes aerial photography or videography; however, there have been limited research efforts focused on utilizing Remote Sensing (RS) data to evaluate the extent of destruction caused by hailstorms (Prabhakar et al., 2019).

1.2. Remote Sensing for crop damage assessment

RS is an efficient and distinctive method for conducting quick and accurate studies of natural disasters (Sosa et al., 2021). Due to extensive coverage, precise geocoding, frequent revisit, rapid data distribution, comparatively low data cost, and robust crop and land discrimination, remotely sensed data have been proven useful in crop loss assessments (Shrestha & Rahman, 2021). With the introduction of advanced platforms and sensors offering improved spatial, temporal, and spectral capacities, there has been a significant increase in RS studies offered to agricultural applications in recent years. (Weiss et al., 2020).

RS involves acquiring information from a distance, facilitated by remote sensors mounted on satellites or aircraft that detect and record reflected or emitted energy (NASA (ESDS), 2022). Optical RS gathers reflected and emitted radiation from the observed surfaces within wavelengths that span from visible to shortwave infrared (VSWIR, 0.4-2.5 μm) and thermal infrared (TIR, 8-14 μm) (Berger et al., 2022). At different wavelengths, different materials reflect and absorb light differently. Thus, the targets' spectral reflectance characteristics in multispectral satellite images can be used to distinguish them (Sosa et al., 2021). Concerning discerning plant responses to various crop stresses, reflectance in the visible (VIS, 0.4–0.7 μm), near-infrared (NIR, 0.7–1.3 μm), and shortwave infrared (SWIR, 1.3–2.50 μm) spectra, alongside TIR and solar-induced fluorescence (SIF, commonly at 0.687 μm and 0.76 μm, or across the entire emission wavelength from 0.65 to 0.8 μm), have been the most utilized passive sensing signals (Gerhards et al., 2019).

Multispectral sensors are more responsive to physiological changes appearing in canopy colour (Sosa et al., 2021). Hailstorms induce stress in plants, leading to reduced photosynthetic efficiency, cellular damage, increased vulnerability to pathogens, and yield reduction (Szabó et al., 2021).

The visible spectrum (VIS, 0.4–0.7 μm) is significant as plants under stress, possessing reduced chlorophyll, tend to reflect more red and blue light. Sensors can capture this variation and identify stressed crop areas (NASA EO, 2000). The near-infrared range also plays a crucial role, as healthy plant leaves demonstrate strong light reflection due to their structure and water content (NASA EO, 2000). Any stress-induced changes in plant structure or hydration levels would decrease this NIR reflectance, enabling the detection of stressed plants. SWIR is critical for gauging plant water content. As plants undergoing stress lose water, their reflectance increases in the SWIR range, facilitating the identification of stressed plants. The utility of vegetation indices such as the Advanced Vegetation Index (AVI), Enhanced Vegetation Index (EVI), Green Normalized Difference Vegetation Index (GNDVI), Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Modified Soil Adjusted Vegetation Index (MSAVI), Plant Senescence Radiation Index (PSRI), Moisture Stress Index (MSI), Green Normalized Difference Water Index (NDWIg) and Normalized Pigment Chlorophyll Ratio Index (NPCRI) for evaluating crop damage due to hail has been highlighted in prior research (Zhou et al., 2016; Prabhakar et al., 2019; Sosa et al., 2021; Ha et al., 2022).

Furthermore, Synthetic Aperture Radar (SAR) is a valuable tool for monitoring crops due to its sensitivity to geometrical structures and dielectric properties of targeted objects. The backscattered microwave signal reflected to the sensor from a vegetated surface is directly influenced by the plant coverage, the soil (mediated through the plant canopy), and the interactions between the vegetation and the soil beneath (Karam et al., 1992). This interaction suggests that microwave signals are sensitive to fluctuations in the architecture and geometric structure of the canopy, irrespective of whether these changes are brought about by ordinary plant maturation or a hailstorm incident (Sosa et al., 2021). Microwave signals are also sensitive to variations in surface roughness and the dielectric coefficient associated with the observed materials (Arciniegas et al., 2007). In agricultural applications, changes in the backscattering between two consecutive SAR images acquired under identical conditions, such as the same wavelength, identical look angle, equal slant range distance, and identical polarization, can be attributed to either standard plant growth patterns (Karam et al., 1992) or significant weather events such as a hailstorm.

In SAR imaging, the backscattering coefficient measures how much a radar signal is scattered back towards the source when it hits a target or surface. The backscattering coefficient's value is influenced by a combination of system and target parameters, as outlined by (Ackermann, 2015). System parameters encompass the frequency, incidence angle, and polarization, while target parameters include the dielectric constant, temperature, orientation, and surface roughness (Periasamy, 2018). In biomass assessment, including crop growth tracking, crop stress monitoring, and yield estimation, the system parameter known as 'Polarization' assumes a more significant role than other parameters (McNairn & Shang, 2016).

Following intense hail events, multispectral imagery procured from passive RS satellites offers valuable assistance during crop damage assessment (Jedlovec et al., 2006). While optical data offers many benefits, its application faces certain constraints, including its dependency on weather and daylight conditions. Furthermore, it necessitates rectifying sun glint effects induced by atmospheric disturbances and identifying clouds and their shadows (Sosa et al., 2021). Conversely, SAR provides neither weather-dependent nor time-restricted data making it capable of effectively penetrating cloud cover, giving it an edge over optical data as it's typically not susceptible to atmospheric interference and, therefore, does not generally need atmospheric corrections (Mulla, 2013). However, SAR data's complexities in processing and its limited ability to provide visual and colour information remain challenges (Mulla, 2013). The combination of both optical and SAR data can offset these limitations, offering researchers a more in-depth and comprehensive overview of a given study area (Mulla, 2013; Sosa et al., 2021).

Recent studies have focused on evaluating the potential of various multispectral and microwave indices and parameters for assessing hailstorm damage. Research conducted by Zhou et al. (2016) used multispectral imagery captured by unmanned aerial vehicles (UAVs) to evaluate the extent of hail damage in potato crops at different growth stages. The damage levels ranged from 0% to 33%, 66%, and 99%, and the assessment was conducted during the tuber initiation, early bulk, and late bulk stages of growth for two potato varieties. Vegetation indices, specifically GNDVI, NDVI, and SAVI, were computed. Results indicated that potato plants damaged during the tuber initiation stage exhibited substantial recovery, as evidenced by high GNDVI values 32 days after damage, which correlated well with final crop yield data. Conversely, severe damage during the early bulk stage (99% defoliation) resulted in significantly lower GNDVI values and reduced yield. The study also found a strong correlation (r > 0.77) between vegetation indices and crop yield at the early bulk stage than at other stages.

In the study conducted by Prabhakar et al. (2019), a differential analysis of the NDVI (before and after hail) events was utilized, with Landsat 8 data for identifying hail-induced crop damage. The study included examining six hail-affected sites, ranging in width from 3 to 8 km and length from 6 to 33 km. The crops were classified into four categories: young grape orchards (grape-1, less than 3 years old), old grape orchards (grape-2, more than 3 years old), papaya, and sugarcane. By monitoring changes in the NDVI profiles across these crop types, a simple linear model was developed to estimate the changes resulting from hail damage. High-resolution LISS-IV satellite data from IRS-Resourcesat-2, along with in situ data, was used to classify crops using a maximum likelihood classifier. The study findings highlighted grape crops as the most damaged, followed by sugarcane and papaya, with a crop damage assessment accuracy of 69.6%.

In their study, Sosa et al. (2021) focused on identifying Homogeneous Damage Zones (HDZ)- sections of a crop field that have incurred comparable levels of damage following a hailstorm event. They performed unsupervised K-means clustering to delineate HDZ using remotely sensed data automatically. They used Sentinel-1 SAR (S1) and Sentinel-2 MSI (S2) data. They analysed ten indices, five microwave indices, including Dual-Pol Diagonal Distance (DPDD), Inverse Dual-Pol Diagonal Distance (IDPDD), Vertical Dual Depolarization Index (VDDPI), Microwave Polarization Difference Index (MPDI), and Dual Polarization SAR Vegetation Index (DPSVI), and five spectral indices, namely NDVI, EVI, SAVI, AVI, and NPCRI before and after a hailstorm across areas exhibiting various degrees of damage. The study established that DPSVI and NPCRI were most responsive to hail-induced changes. Temporal data and change rates of these indices served as input variables for the K-means clustering algorithm, which aimed to classify pixels into consistent damage zones. The algorithm's efficiency in identifying homogeneous hail damage areas was demonstrated through the validation process, which involved analysing in situ data from 91 soybean, wheat, and corn fields. The results showed a significant classification output of HDZ, indicating its successful performance in 87.01% of cases. The authors suggested future algorithm enhancements, including crop-specific index selection and consideration of crop phenological stages and varying soil types within the cultivation region, to improve the accuracy of hailstorm event classification.

Recent research conducted by Ha et al. (2022) utilized the capabilities of S2 data within a cloud computing platform to assess field crops damaged by hailstorms. Multiple indices were employed, including the NDVI, NDWI, EVI, GNDVI, MSAVI, PSRI, MSI and NDWIg. It was found that S2 time-series data effectively monitored the hail-induced damage and subsequent recovery of canola, wheat, and lentil crops over time. NDVI temporal profiles revealed that crops affected by hail recuperated approximately a month post-damage. The area under the curve (AUC) dataset for one month could precisely quantify hail-impacted crops, accounting for both direct damage and plant recovery, without cloud cover hindrance.

The 32-day AUC of NDVI, NDWI, and PSRI showed a robust correlation with ground-estimated hail damage for canola (r = -0.90, RMSE = 8.24), wheat (r = -0.86, RMSE = 12.27), and lentil (r = 0.80, RMSE = 17.41). The research was confined to a single hailstorm event with a limited number of fields under observation, suggesting that future studies should validate the methodology across larger areas, multiple crop species, and different growing stages.

1.3. Machine Learning (ML)

Artificial Intelligence (AI) involves developing computer systems to perform tasks requiring human-like intelligence. It encompasses algorithms and models enabling machines to perceive and interpret data, make decisions, and adapt to different situations (Camps Valls, 2008). ML, a subfield of AI, focuses on algorithms and models that enable computers to learn from data training. ML algorithms automatically identify patterns, make predictions, and develop methods that iteratively enhance their performance by learning from data (Camps Valls, 2008; Maxwell et al., 2018). ML provides an opportunity for accurate and efficient classification of remotely sensed imagery.

The advantages of ML lie in its ability to handle high-dimensional data and effectively classify classes with intricate characteristics (Maxwell et al., 2018). ML offers improved accuracy, efficiency, adaptability, and scalability in RS applications, making it preferred over conventional approaches like visual interpretation and parametric classifiers such as maximum likelihood classifiers (Maxwell et al., 2018). It leverages the power of data-driven models and automated learning to extract meaningful information from RS data and aid in better decision-making for various environmental and geospatial applications (Camps Valls, 2008; Maxwell et al., 2018).

Research done by (Maxwell et al., 2018) employing various ML algorithms, including Random Forest (RF), Support Vector Machines (SVM), Decision Trees (DT), Boosted Decision Trees (BTD), Artificial Neural Network (ANN), and k-Nearest Neighbors (k-NN), for a classification problem. The research findings highlighted that the choice of ML algorithm depends on several factors. Firstly, the sample size and quality of training data significantly impact classification accuracy. Secondly, training data imbalance can affect accuracy, especially for rare classes, thus necessitating the selection of an appropriate method that effectively handles data imbalance. Thirdly, feature selection is crucial in simplifying the model while maintaining sufficient accuracy.

RF, in particular, exhibits notable advantages in managing large datasets with high-dimensionality and multiple input variables, making it well-suited for complex classification problems (Belgiu & Drăgu, 2016; Breiman, 2001; Liaw & Wiener, 2002). RF also excels in handling the inherent non-linearity often encountered in classification tasks. Additionally, RF serves as a powerful tool for dimensionality reduction, as it assesses the significance of each feature relative to the problem at hand (Breiman, 2001; Izquierdo-Verdiguier et al., 2017). Unlike other techniques, RF effectively handles the diverse, dynamic ranges of parameters without requiring scaling or normalization (Hariharan et al., 2022). This characteristic is advantageous in crop classification scenarios where polarimetric parameters exhibit varied ranges (Hariharan et al., 2022). An additional strength of RF is its inherent resistance to overfitting, a common issue in many ML models (Belgiu & Drăgu, 2016; Breiman, 2001). Furthermore, RF demonstrates robustness against noise and outliers, contributing to the stability of the model (Belgiu & Drăgu, 2016).

As delineated in section 1.2, combining optical and SAR data can counteract their respective constraints, presenting researchers with a more exhaustive and comprehensive overview of a given study area (Mulla, 2013; Sosa et al., 2021). In the context of our present research, the utilization of the SAR and optical data with RF model is justified due to its advantageous characteristics compared to other ML models. Given the complexity of our dataset, which incorporates both optical and SAR data, RF model offers a robust solution that can handle diverse data types and effectively capture complex relationships.

1.4.　　Research Identification

The literature review revealed a gap in research focusing on utilising multi-temporal data from different satellite sensors, such as optical and SAR, using ML techniques like RF to assess crop damage caused by hailstorms. To address this gap, our research aims to leverage both optical and SAR time series data, in combination with various vegetation indices and polarimetric parameters, to investigate the applicability of the RF method. By examining Pixel Based Classification (PBC) and Object Based Classification (OBC) individually on optical and SAR data, as well as on combined datasets, we aim to assess the performance and accuracy of the RF model in different situations.

In addition, this research acknowledges the importance of promoting open science and enhancing reproducibility. To achieve these objectives, a focus will be placed on utilizing open-source satellite data, open-source software, and openly shared codes throughout the research process.

1.5.　　Research Objectives & Research Questions

This research aims to develop a RS-based crop damage assessment model in order to support decision-makers and farmers to access the hail damage in crops quickly and accurately for large geographical areas. A supervised ML method is developed to map RS image information to the field and in-situ crop damage data. The model developed in this study was trained and tested using in-situ data from hail-affected regions. The dataset encompasses about 2036 hectares and roughly 3400 parcels from five villages, as listed in Table 2, in the Kathumar tehsil, Alwar district, from the year 2020. The three sub-objectives and six research questions (RQs) outlined below have been formulated in alignment with the primary objective of this research.

1.5.1.　Main objective.

The primary objective is to propose a remote sensing-based crop damage assessment method using multi-temporal data from S1 and S2 and a RF classifier.

1.5.2.　Sub-objectives.

1) **To evaluate crop damage caused by hailstorms using a RF classifier on S1 and S2 data separately.**

   RQ 1: How can using RF classifier on S1 data improve the present field survey-based crop damage assessment approach?

   RQ 2: How can using RF classifier on S2 data improve the present field survey-based crop damage assessment approach?

   RQ 3: Which classification approach, PBC or OBC, results in a more accurate assessment of hailstorm-related crop damage when utilizing S1 and S2 data separately?

2) **To evaluate crop damage caused by hailstorms using a RF classifier on a combination of S1 and S2 data.**

   RQ 4: How can using RF classifier on combined S1 and S2 data improve the present field survey-based crop damage assessment approach?

   RQ 5: Which classification approach, PBC or OBC, results in a more accurate assessment of hailstorm-related crop damage when utilizing a combination of S1 and S2 data?

3) **To assess the potential impact of the new procedure on different stakeholders involved.**

   RQ 6: How would the proposed remote sensing-based crop damage assessment method impact different stakeholders?

1.6. Study Area

In the past five years, various districts in Rajasthan, India, had severely affected by hailstorms. Hail event dates and affected districts can be seen in
Table 1. The study focuses on five villages listed in Table 2 of tehsil Kathumar, district Alwar, which is situated in the northeast part of Rajasthan, as depicted in Figure 3.
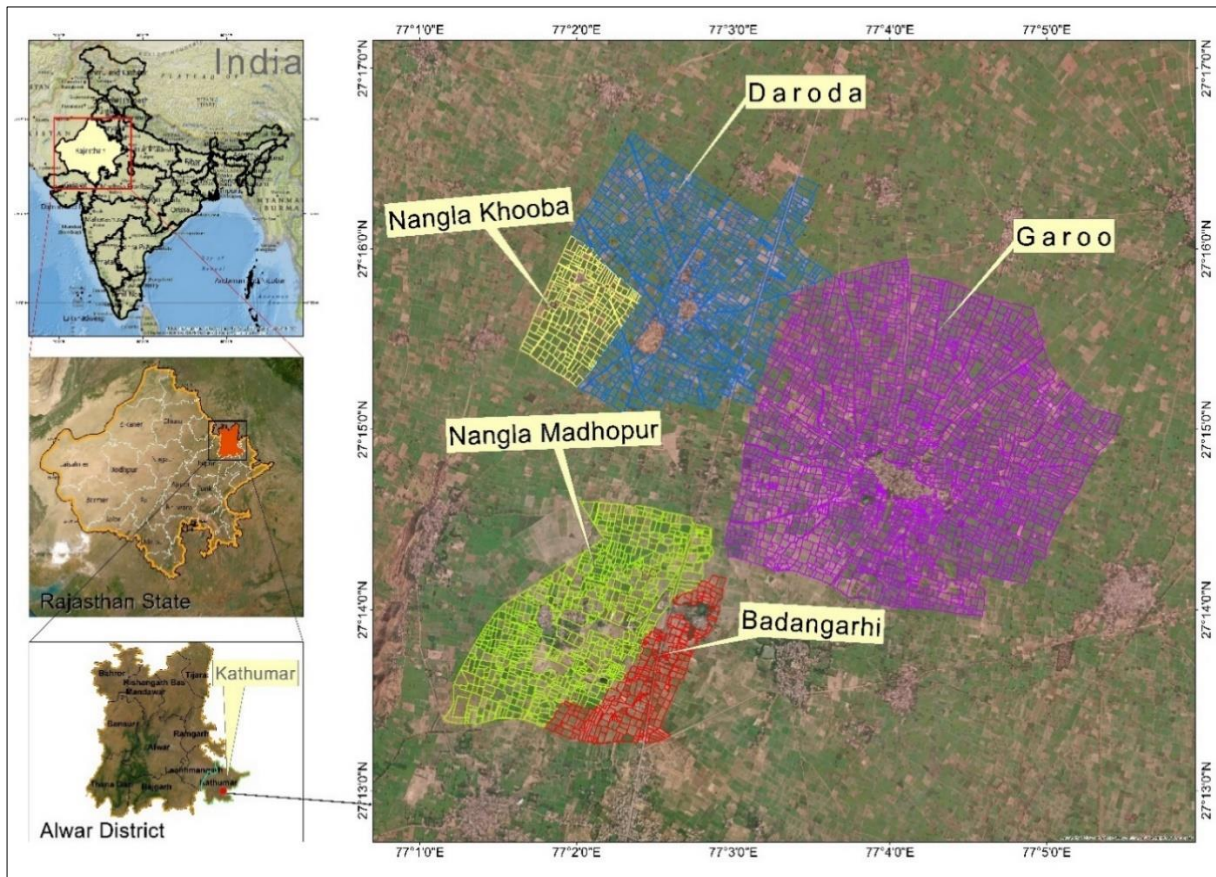
Figure 3. Study Area (Parcel map: Revenue Department, Rajasthan; Satellite Data: QGIS base map)

Alwar district plays a significant role in Rajasthan's agricultural production. The district encompasses a total area of 7,82,897 ha, or approximately 2.5% of the state (KVK, 2023). The district's total cultivated

area is 7,81,615 ha. In the Kharif season (July-October), maize, sorghum (jowar), millet, pigeon pea (arhar), sesamum, cotton, etc., are sowed on approximately 3,29,088 ha (42%), whereas in the Rabi season (November-March), wheat, mustard, gram, barley, rabi pulses, etc., are planted on approximately 4,525,250 hectares (58%) (KVK, 2023). The ten-year average precipitation is 724 mm. The agriculture in the district is largely dependent on the distribution of precipitation (KVK, 2023).

Table 1. Hailstorm events in the past five years in Rajasthan, India

Sr. No.	hailstorm date	Districts affected	Source
1	14 February 2019	Bundi, Baran, Jhalawar, Kota, Nagaur, Churu, Sriganganagar, Bikaner, Jaisalmer, Barmer and Jodhpur	(*DNA*, 2019)
2	05 March 2020	Alwar, Bharatpur, Dausa, Jaipur	(India Today, 2020)
3	5 January 2021 and 12 March 2021	Bundi, Baran, Jhalawar, Kota	(Times of India, 2021)
4	8 January 2022 and 8 March 2022	Alwar, Bharatpur, Jaipur, Sikar	(Times of India, 2022)
5	29 January 2023 and 8 March 2023	Ajmer, Alwar, Jaipur, Jaisalmer, Udaipur	(Patrika News, 2023)

Rabi (November-March) and Kharif (July-October) are the two main crop seasons in the region. In India, crops cultivated during the rainy season are known as Kharif crops, while those cultivated during the winter are known as rabi crops. Kharif crops' seeds are planted at the beginning of the monsoon season (the start of July). These crops are harvested after the monsoon season (end of October). Rabi season crop seeds are planted in mid-November at the beginning of winter and harvested by the end of the winter season (end of February or March). Mustard, wheat, barley and gram are the principal rabi crops in the region. In the present study, we are focusing on rabi crops because rabi crops are the most affected crops by hailstorm events in the region (Bal et al., 2014; Interview, 2023).

Table 2. Study area villages

Sr.	Village Name	Area (ha)	Tehsil, District	Province, Country
1	Badangarhi	115.17		
2	Daroda	448.61		
3	Nagla Khooba	105.60	Kathumar, Alwar	Rajasthan, India
4	Nagla Madhopur	339.89		
5	Garoo	1029.32		

1.7.   Stakeholders

In order to determine the main stakeholders, a review of relevant literature is carried out, involving thorough reviews of relevant ministry websites and newspaper articles. Stakeholder identification was further enriched through an interview conducted with a local farmer from the study area. The purpose of this section is to convey the understanding of the stakeholder's roles concerning the problem at hand. The stakeholders are identified, and their roles are subsequently outlined. They are classified into primary and secondary groups, depending on the importance pertaining to the problem. In the context of the present research, the term 'importance' signifies the level of attention that needs to be given to meeting the requirements and concerns of various stakeholders involved in the project (Olander & Landin, 2005). Five primary and eight secondary stakeholders are identified, which are listed below.

Primary Stakeholders:

I. **Farmers:** They are the main stakeholders in hailstorm damage assessment research as they are the ones who bear significant impacts from such incidents (News NCR, 2022; Patrika News, 2023; Times of India, 2020; Times of India, 2021).

II. **Agricultural Labourers:** They are landless farmers who lease land from other farmers to carry out their farming activities. They utilize their own resources for cultivation and face an added burden as they are typically not eligible for compensation in the event of crop damage. Instead, such compensation is usually directed towards the land's registered owners. They have the most significant impact from such incidents (News NCR, 2022; Patrika News, 2023; Times of India, 2020; Times of India, 2021).

III. **Rajasthan State Government:** The state government has control over fund allocation and departmental directives, which allows them to ensure that all processes work smoothly and the farmers receive their due compensation. They hold ultimate control over the budget and policy direction, making them highly influential and important. Despite not being directly involved in damage assessment or compensation distribution, their decisions significantly affect the overall process (DMRD, 2023; State Portal Rajasthan, 2023).

IV. **Disaster Management, Relief & Civil Defence Department (DMR)**: The DMR's role extends to the efficient processing of damage reports received from the Board of Revenue and compensation disbursement. They are to operate within the constraints of a fixed budget while ensuring the compensation process is efficient, and the farmers are adequately compensated for their loss. Their direct involvement in the compensation process explains their high influence and importance scores (*DMRD*, 2023).

V. **Board of Revenue (BoR)**: The organization has the direct responsibility of evaluating the degree of damage and recording it via its regional revenue departments in accordance with government guidelines (BoR, 2023).

Secondary Stakeholders:

VI. **Local Communities:** The inhabitants of the affected areas often depend on the local agricultural sector for their livelihood, sustenance, and economic progression. This includes agriculture workers on daily wages who assist with all stages of crop cultivation, from sowing to post-harvest, some of whom are temporarily or seasonally employed labourers. Local workers of food processing, biofuel production, or seed and fertilizer companies, whose operations are reliant on farm outputs. Additionally, local small vendors, such as vegetable and fruit or grain sellers, source their goods from local farms.

VII. **Agriculture Department**: The department serves multiple functions focused on improving agricultural productivity and enhancing the income of farmers (Agriculture Dept., 2022).

VIII. **Central Government:** This authority is responsible for allotting central funds to the DMR (Annual Progress Report DMR, 2023).

IX. **Print and Electronic Media**: Print and electronic media play a significant role in the event of crop damage due to hailstorms, as they are primarily responsible for disseminating information

and raising awareness about the situation. They provide critical updates on the extent of damage, the response from authorities, and the impact on the affected communities. They also publish important weather forecasts and predictions, helping communities prepare for potential hailstorms (News NCR, 2022; Patrika News, 2023; Times of India, 2020; Times of India, 2021).

X.   **Crop Insurance Providers: These stakeholders** play an indirect yet significant role in the situation. As insurance companies, their interest primarily lies in maintaining a balanced portfolio of risks to ensure profitability. This balance is achieved by ensuring a sufficient number of farmers are insured or that there is a large enough area of farmland covered under their policies. (PMFBY Crop Insurance, 2023).

XI.   **Research Institutions:** These include universities and other organizations that conduct research on hail events, crop damage, and agricultural resilience to shape future policies or mitigation plans.

XII. **Non-Governmental Organizations (NGOs):** NGOs often step in to provide support to communities affected by hailstorm-induced crop damage. This can include distributing emergency aid, providing support in navigating compensation procedures, and offering resources for recovery and rebuilding. NGOs may also advocate for affected communities, pushing for fair practices and sufficient aid.

XIII. **Local Supply Chain Entities:** These businesses, including seed suppliers, fertilizer companies, and agricultural equipment providers, might experience changes in demand following a hailstorm.

1.8.   Problem Analysis

1.8.1.   Problem tree.

In order to come up with an intervention that reduces a part of wickedness in the present situation, a problem tree has been developed. The problem tree illustrated in Figure 4 showcases what the author sees as the main causes contributing to the problem. One key issue is climate change, which is causing hailstorms to become more severe, producing larger ice pieces and heavier rain due to the increasing global temperatures (BBC Future, 2023). Hailstorm events cause an abrupt loss of harvestable crops and, in some cases, the complete destruction of mature crops (DNA, 2019). Hail-affected areas are occasionally inaccessible, and damage assessment accuracy mostly depends on the visual evaluation of HDZs in the field (Sosa et al., 2021).

Farmers attempt to limit the effects of natural disasters by implementing management measures inside their agricultural operations, but certain risks, such as hailstorms, are beyond their financial resources (Sosa et al., 2021). Enumerating crop loss and compensating impacted farmers has proven to be challenging (Prabhakar et al., 2019). Currently, the process of assessing crop damage largely depends on field surveys and visual checks, which are costly, time-consuming, and require a lot of human resources. Moreover, these methods can be less transparent, as visual checks might sometimes be biased for personal gain (Prabhakar et al., 2019; Rao & Raju, 2010).

The existing process for assessing damage is identified as the central research problem in this study and can be seen as represented by the light orange box in Figure 4. Delays in evaluating crop damage lead to delays in compensation, which can result in farmers losing trust in the government. These delays cause financial instability for farmers and can also cause delays in planting the next crop.

This impacts not only farmers but also the local economy, as farmers are major buyers of farming inputs and the main providers of agricultural products to local markets. Delays in planting the next crop can also negatively impact food security in the area, affecting not only the farmer's family but also the wider community. Sudden financial losses due to hail-damaged crops can cause health problems for farmers and make them feel financially insecure. These factors can lead to severe outcomes, including suicide among farmers (Times of India, 2023).
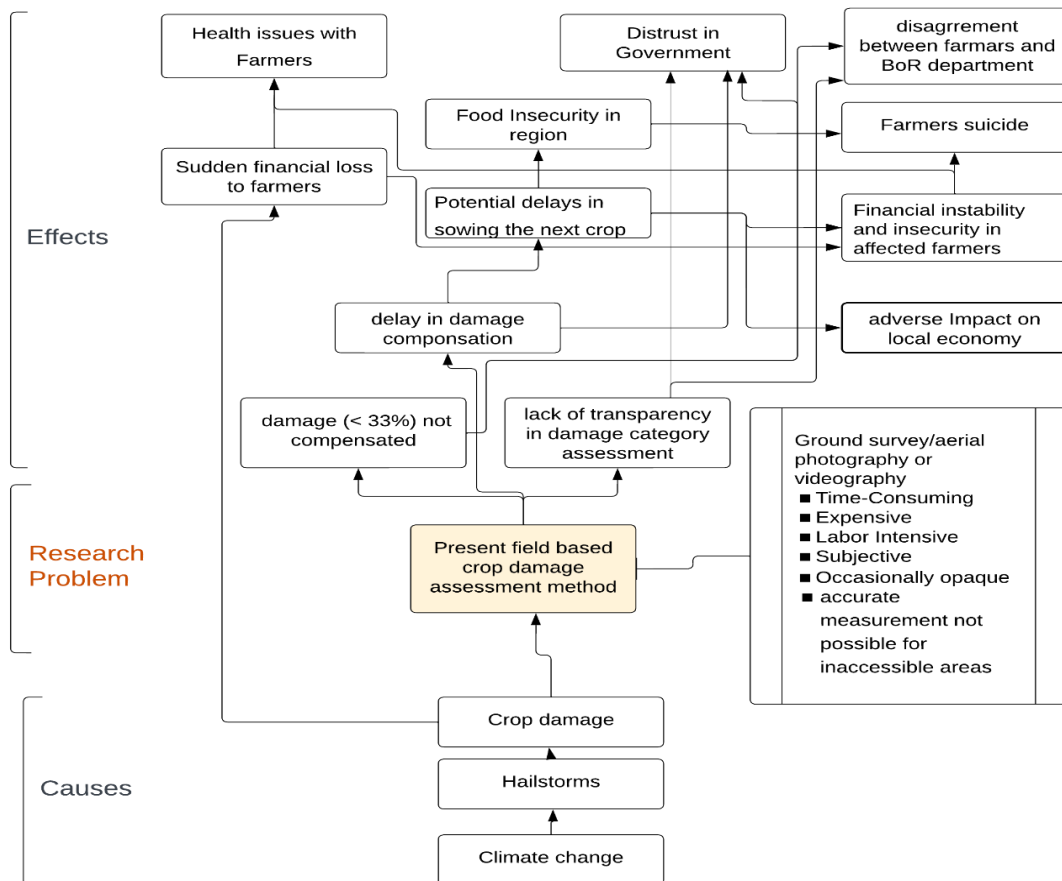
Figure 4. Problem tree in crop damage assessment

The assessment of crop damage is categorized into four segments: 0-33%, 33-50%, 50-75%, and 75-100%. However, compensation is only provided to farmers when the damage exceeds 33% of the total crop area (DMRD, 2023). also, there is a lack of transparency in the assessment of the damage category, leading to disagreements between officials from the BoR and farmers. Farmers have expressed their discontent with the assigned damage categories and the compensation provided when the damage surpasses the 33% threshold, as well as the predetermined compensation per hectare established by the government (Interview, 2023). To address this, farmers are advocating for a revision of the categorization, proposing narrower intervals such as 0-15%, 15-30%, 30-45%, and so on or 10% bin increments for more precise assessment (Interview, 2023).

Farmers express discontent with the drawn-out procedures and time required for damage assessment and distribution. The wicked nature of this issue is causing considerable disturbances in the workings of the government (Times of India, 2020).

1.8.2.   Wicked problem analysis.

The analysis of the wicked problem was formulated utilizing the complex problem framework, as represented in Table 3, proposed by Georgiadou & Reckien (2018). The table displays the four categories of policy problems alongside the tools designed to address them. When it comes to wicked problems, stakeholders often engage in ongoing debates regarding the exact nature of the problem and the best strategies for resolution. The optimal outcome is to lessen the wickedness of the problem, which involves reducing uncertainty in knowledge and facts and fostering agreement among stakeholders (Georgiadou & Reckien, 2018).

The existing system in Rajasthan employs a method known as 'girdavari', executed by the BoR. This process involves village-level officials recording crop losses and inputting essential details such as land ownership, irrigation source, cultivator's name, and other relevant data. The DMR uses this report form each regional office of BoR to determine the compensation amount for each region (DMRD, 2023; The Hindu, 2023). Compensation is awarded to farmers when damage surpasses 33% of the total crop area. The compensation disbursed aligns with the pre-approved amount per hectare, with damage estimates segmented into four categories: less than 33%, 33-50%, 50-75%, and 75-100% (DMRD, 2023). Farmers have expressed discontent with the prolonged damage evaluation process and the government's predetermined compensation per hectare (Interview, 2023; Times of India, 2020).

The field-based assessment techniques do not clearly delineate whether the damage results from the hailstorm or from other factors such as suboptimal farming practices, inferior soil quality, inadequate water supply, crop diseases, tainted seeds, inappropriate use of fertilizer, or unsuitable cultivation materials (CMDA, 2022). Additionally, the funds allocated for disaster management are restricted. For the fiscal year 2021-22, the state's share was 134,284,473 euros, and the central government's contribution was 44,761,491 euros, culminating in a total reserve fund of 179,045,964 euros (DMRD, 2023) for disaster management operations state-wide. The ambiguity inherent in the visual assessment process, coupled with the lengthy timeframe for damage assessment, financial constraints, and lack of personnel and spatial information, significantly intensifies the problem.

Disagreements have emerged between farmers and the government for various reasons, like lack of transparency in assessing the damage category, as farmers contest the surveyors' identification method without any quantitative measurement, leaning more towards a subjective analysis (Interview, 2023). The current damage category divisions or bins (0-33, 33-50, 50-75, 75-100) are not aligning with farmers' perspectives. They advocate for smaller categories or bins to distribute proportional compensation among the affected (Interview, 2023).

The time lag in damage assessment and compensation disbursement by the department also raises disagreement, as farmers require funds promptly to sow the next crop (Interview, 2023). Farmers who have connections with government authorities frequently obtain increased compensation by being classified into higher damage categories, regardless of experiencing comparable losses (Interview, 2023). This leads to an unfair allocation of resources, introducing inequality in the process. Occasionally, they might even procure compensation for crops that are distressed and damaged due to the other reasons discussed above (Interview, 2023). Moreover, mustard crops are harvested about 20 to 25 days earlier than wheat. Thus, if a hailstorm happens within this period, farmers may submit compensation claims, mentioning hail-induced crop losses.

The success of these claims frequently depends on the farmers' existing networks and relationships (Interview, 2023). Thus, the presence of corruption within the system is a substantial cause for disagreement. The government's guideline that compensation is only provided when damage surpasses the 33% mark and pre-set compensation per hectare, as defined by the government, is another significant point of contention (Interview, 2023). However, this issue hinges on the government's financial resources and is a matter of debate.

Concerning the understanding of crop damages, a notable knowledge gap exists in assessing the amount of damage and proper damage categorisation within the limited time frame. This uncertainty complicates the formulation of effective solutions. The conflict of differing goals and values and the inadequate understanding of spatial knowledge make this issue wicked. As per the wicked problem framework by Georgiadou & Reckien (2018), this issue can be characterized as a wicked or unstructured problem, landing in quadrant four, as highlighted in green colour. Here, an RS-based model can be implemented as a key Geo-information tool to bridge this gap. Encouraging active dialogue among stakeholders concerning their goals and values is of utmost importance.

Table 3. Four types of policy problems and related tools (Source: Georgiadou & Reckien, 2018)

Spatial Knowledge	Policy Goals and Values	
	Consensus among Stakeholders	**Dissensus among Stakeholders**
Certain (Facts and cause-effects)	(1) Tame or structured problems ➢ Debate on technicalities ➢ Geo-information tools as problem solver	(3) Moderately structured problems ➢ Participation to debate goals and values ➢ Geo-information tools as mediator
Uncertain (Facts and cause-effects)	(2) Moderately structured problems ➢ Participation to debate cause-effects and optimize the collection of facts. ➢ Geo-information tools as analyst and/or advocate	(4) **Wicked or unstructured problems** ➢ Endless debate ➢ Geo-information tools as problem recognizer

The government is already showing interest in leveraging modern technologies to enhance the transparency and efficiency of the damage assessment process. This existing momentum provides a favourable environment for the proposed research. In a significant step forward, the Government of Rajasthan declared in March 2023 that they are developing a mobile application. This app will empower farmers in Rajasthan to independently evaluate and report the damage caused to their crops by natural disasters (The Hindu, 2023). And colouration of this app with the proposed RS model could make it more valuable for farmers as well as related departments. While farmers are not currently aware of such applications, they express a positive attitude towards the possibility of using such tools directly (Interview, 2023).

To better support farmers, policy strategists, government officials, and agricultural insurance firms in their decision-making, there is a critical need for an enhanced process to assess crop damage, specifically in monitoring and evaluating the impact of hailstorms. Given the diversity of stakeholder opinions and current debates surrounding the limitations of the existing system, as highlighted in the problem tree, this thesis principally seeks to establish a RS-based model. This innovative model can be utilised to assist authorities in performing rapid and precise assessments of crop damage.

2. DATA DESCRIPTION

The analysis was primarily focused on the identification of the impact of a significant hailstorm that took place after midday on the 5th of March 2020. This hailstorm event is selected as the subject of a detailed examination and analysis. The following section elaborates on the data utilized in this study. Figure 5 shows the picture of 5 March 2020 after the hailstorm.

Figure 5. Picture of 5 March 2020 after hailstorm event (Source: DNA, 2019)

2.1.   Sentinel-1 SAR

The Sentinel-1 (S1) mission is comprised of a two-satellite constellation, namely Sentinel-1A (S1A) and Sentinel-1B (S1B), in polar orbit. S1 operate continuously, day and night, utilizing the capabilities of C-band SAR imaging. The active sensing technology allows to capture images irrespective of prevailing weather conditions. Each S1 satellite provides a revisit time of 12 days at the equator. However, when working synergistically, they can reduce this to a revisit period of 6 days. The C-band SAR aboard the satellites can provide data in varying polarization modes, either single (HH or VV) or dual HH+HV or VV+VH (*ESA Sentinel Mission*, 2023).

The satellites have the capacity to acquire data in four distinct acquisition modes: Interferometric Wide Swath (IW), Extra Wide Swath (EW), Strip Map (SM), and Wave (WV). They yield two different kinds of data products: Single Look Complex (SLC) and Ground Range Detection (GRD). While the SLC product maintains both phase and amplitude information, the GRD product imparts amplitude information only (*ESA Sentinel Mission*, 2023). S1B mission was suspended on 03 august 2022 after an anomaly was detected on 23 December 2021. Now the Sentinel-IC mission, scheduled for launch in 2023, will carry on radar imaging for a wide variety of applications.

S1 is suitable for the present study because high-frequency radar penetrates cloud cover and doesn't dependent on solar light. Therefore, it can deliver the reliable and timely observations required for agricultural applications (Bush & Ulaby, 1978). Low-frequency radar is ideal for monitoring soil moisture, water content, and structural changes in crops (Khabbazan et al., 2019). S1 uses more than one polarisation to emit and record the signals. The polarisation characteristic of S1 allow extraction of ample information from the earth surface. For instance, surface scattering (e.g., sea surface) have different polarisation than volume scattering (e.g. forest). Polarimetric methods allow the separation of different scattering contribution and can be used to extract information about scattering process (*ESA - Sentinel-1*, 2022), which contribute to improved classification. The present research has been conducted utilizing data from the dual polarization mode VV+VH of the Sentinel-1A SLC product. The specifics of this data are comprehensively delineated in Table 4.

Table 4. Details of SAR data from Sentinel-1 used in the research.

Sensor	Sentinel-1A SAR
Product Type	SLC -IW
Bandwidth, Central Frequency, Resolution (rg x az)	C band 0-100 MHz, 5.405 GHz, (2.7x22 m to 3.5x22 m)
Polarisation	Dual (VV, VH)
Pre event date, time, and orbit	5 March 2020, 00:52:59 UTC, 031533
Post event date, time, and orbit	17 March 2020, 00:52:54 UTC, 031708

2.2. Sentinel-2 MultiSpectral Instrument (MSI)

The Copernicus Sentinel-2 (S2) initiative involves a pair of satellites - S2A and S2B - that orbit in a polar configuration, located in a shared sun-synchronous orbit, with a phasing of 180° between them. The mission's broad swath size (290 km) and frequent repeat intervals (once every 10 days with a single satellite and once every 5 days with both satellites in operation) are among its key features (*ESA Sentinel Mission*, 2023).

The S2 encompasses 13 unique spectral bands: four bands (B2: Blue, B3: Green, B4: Red, B8: NIR) with a spatial resolution of 10 metres, and six bands (B5, B6, B7, B8a, B11, B12) offering a resolution of 20 metres, and three bands (B1: Coastal aerosol, B9, B10: Cirrus) with a 60-metre resolution. The sensors possess a radiometric resolution of 12 bits, which facilitates the identification of minor differences in surface reflectance (*ESA Sentinel Mission*, 2023).

The spatial, spectral, radiometric, and temporal resolution and open access data policy of S2 is suitable for monitoring changes in agriculture and forest areas (Sanchez et al., 2020). The Sentinel 2 Level-2A product was utilized in the present research. This product provides surface reflectance that have been atmospherically corrected and are presented in cartographic geometry. All the bands were incorporated in the analysis except B1 band (Coastal aerosol) and the B10 band (Cirrus) because Band 1 primarily aids in coastal and aerosol studies, and Band 10 is sensitive to high-altitude cirrus clouds rather than terrestrial features. The specifics of this data are comprehensively delineated in Table 5.

Table 5. Details of Multi Spectral Instrument data from S2 used in the research.

Sensor	Sentinel-2A-MSI
Product Type	L2A
Spatial Resolution	10-60 m
Temporal Resolution	5 days
Spectral resolution	12 bands (Visible to SWIR)
Radiometric Resolution	12 bits
Pre event acquisition date, time, orbit no, tile no	05-03-2020, 05:27:11 UTC, R105, T43RGL
Post event acquisition date, time, orbit no, tile no	25-03-2020, 05:26:41 UTC, R105, T43RFL

2.3. Field Data

For the research at hand, field data was meticulously compiled from five villages in the Alwar district, as delineated in Table 2. This data encompassed three crucial elements: cadastral maps, Revenue Records, and Hailstorm Reports. Cadastral maps, which present parcel boundaries with corresponding parcel or survey numbers, were obtained from the Land Revenue Department, District Alwar. This valuable set of information is also readily available for online access via https://bhunaksha.raj.nic.in/.

Another source of vital data came in the form of the Revenue Record or Registry of Records (RoR), providing specific details about each parcel or survey number, including land area, ownership status (whether private or government-held), and land type (such as agricultural, residential, waterbody, grassland, forest land, etc.). This information can be conveniently accessed on the website http://apnakhata.raj.nic.in/.

Furthermore, communication was established with the DMR of the Government of Rajasthan in order to procure crop damage assessment reports. These reports are comprehensive, categorizing crop damage levels into four groups as per government guidelines: low (0-33%), moderate (33-50 %), high (50-75%), and severe (75-100%). This gradation of data is particularly useful for understanding the impact of the hailstorm event and subsequently shaping the analysis. The specifics of this data are comprehensively delineated in Table 6.

Table 6. Field data utilized in the research.

Sr.	Type	Description	Agency
1	Cadastral Map	Village cadastral map showing parcel boundary with parcel number/survey number	Land Revenue Department, District Alwar (works under BoR, Rajasthan Government) URL: https://bhunaksha.raj.nic.in/
2	Revenue Record/ Registry of Records (RoR)	Village RoR containing parcel number/survey number-specific information on land area, farmer details (name, father name), land ownership (private/government), and land type (agriculture, residential, waterbody, grassland, forest land, etc.).	Revenue Officer, Land Revenue Department Tehsil Kathumar, District Alwar (works under BoR, Rajasthan Government) URL: http://apnakhata.raj.nic.in/
3	Hailstorm damage report	The report comprises four categories of crop damage level information due to hailstorms according to government guidelines. I. Low (0-33%) I. Moderate (33-50 %) II. High (50-75%) III. Severe (75-100%)	Disaster Management, Relief & Civil Defence Department (DMR), Govt. of Rajasthan URL: https://dmrelief.rajasthan.gov.in/

2.4. Software & Packages used

To accomplish the objectives of this research, a variety of specialized software tools were utilized, each serving unique and essential functions within the data processing and analysis workflow. Table 7 provides a detailed overview of the functions of each software used in this study.

Table 7. Software utilized for data processing and analysis in the research.

Sr	Software/Packages	Developer (license)	Function
1	SNAP	European Space Agency (Open source)	Sentinel Application Platform (SNAP) was employed to conduct pre-processing operations on data derived from the S1 and S2 satellites.
2	QGIS	QGIS Development Team (Open source)	QGIS was a fundamental tool used in this research for georeferencing cadastral images. This ensured the spatial accuracy of the images in relation to real-world coordinates. Additionally, QGIS was used to prepare shapefiles that served as training and testing datasets, which was essential in creating and validating our spatial models.
3	R Studio	Posit, PBC (Open source)	R Studio was used specifically to develop a Random Forest Classification Model. This model was integral to our research, as it facilitated the detection of hail damage on a parcel-by-parcel basis, thereby enhancing the specificity and accuracy of our damage assessment.
4	PolSARpro	European Space Agency (Open source)	This specialized software was crucial in processing polarimetric SAR data. The usage of PolSARpro ensured comprehensive handling and interpretation of the SAR data, which was critical in achieving a detailed understanding of the radar returns within the context of my study.
5	Python	Python Software Foundation (Open source)	A Python script was utilized to assign coordinates to the outputs from PolSARpro.
6	CRIB	ITC Geospatial Computing Platform	GPU-assisted computing capabilities for ML involving geospatial data

3. METHODOLOGY

3.1. Methodological Workflow

As explained in Section 1.8, there is a essential requirement for an improved methodology to evaluate crop damage, particularly in tracking and assessing the ramifications of hailstorms. This need is essential to augment the decision-making capabilities of policy planners, governmental authorities, agricultural insurance companies and farmers. Accordingly, this study is designed with the intention of formulating a RS-centred crop damage assessment model. The goal is to expedite and enhance the accuracy of hail damage assessment for crops, particularly over broad geographical areas.
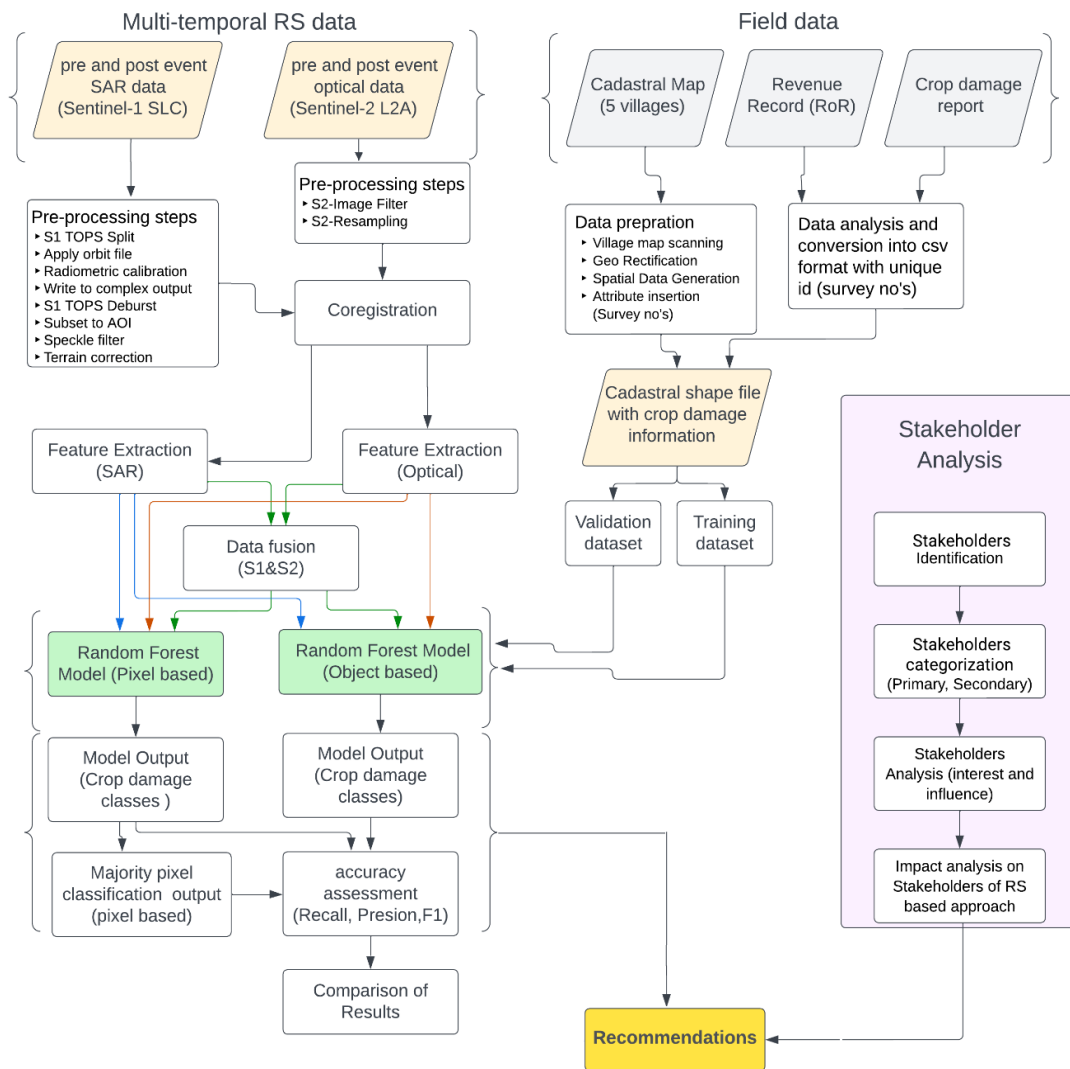
Figure 6. Methodological workflow for the present research.

The methodological workflow for the present study is visually represented in Figure 6. The research incorporates multi-temporal data from S1 (as catalogued in Table 4) and S2 (as outlined in Table 5). These data have been employed to extract features that serve as input to the RF model.

In addition, The field data summarised in Table 6 has been used to create training and testing datasets for supervised learning within the RF model. The model incorporates both PBC and OBC approaches, generates predictive outputs, and performs accuracy analyses. An analysis of stakeholders has also been undertaken to examine their interests and influence within the context of the current research, along with assessing the impact of the proposed intervention. The research concludes with recommendations derived from its findings. The software outlined in Table 7 is employed to execute the analysis stages depicted in the methodological workflow. The distinctive steps represented in this workflow are described comprehensively in Sections 3.2 to 3.8.

3.2.   Data Pre-Processing

3.2.1.   Sentinel-1 SAR.

In the scope of RS data, pre-processing holds significant importance. This study investigates the polarimetric features and microwave vegetation Dual-pol Radar Vegetation Index (DpRVI) derived from S1- Single Look Complex (SLC) data. A systematic series of steps is necessary to extract features from S1 data, which were meticulously followed in this research. The pre-processing and feature extraction from the S1 data were facilitated by the SNAP and PolSARpro software, as outlined in Table 7.

The methodological workflow (Figure 6) outlines the pre-processing steps undertaken for both the pre and post-hail event datasets. S1 SLC data is divided into three sub-swaths, each comprising of nine bursts. The split tool allowed me to isolate sub-swath IW1 pertinent to my study area. Thereafter, an orbit file was applied, a process which refines the preliminary state vectors in the SAR product's metadata with more accurate ones. The subsequent step involved radiometric calibration. Considering the use of both real and imaginary parts of the data, it was crucial to save the data in a complex output format, which retains both amplitude and phase information. After preserving the complex output, TOPSAR deburst procedure was used, which seamlessly fuses separate image bursts into one comprehensive image. Since data was derived from a single swath, there was no need for the TOPSAR merge tool. A subset operation was then performed, aligning with the study area. Figure 7 shows the workflow of processing steps.
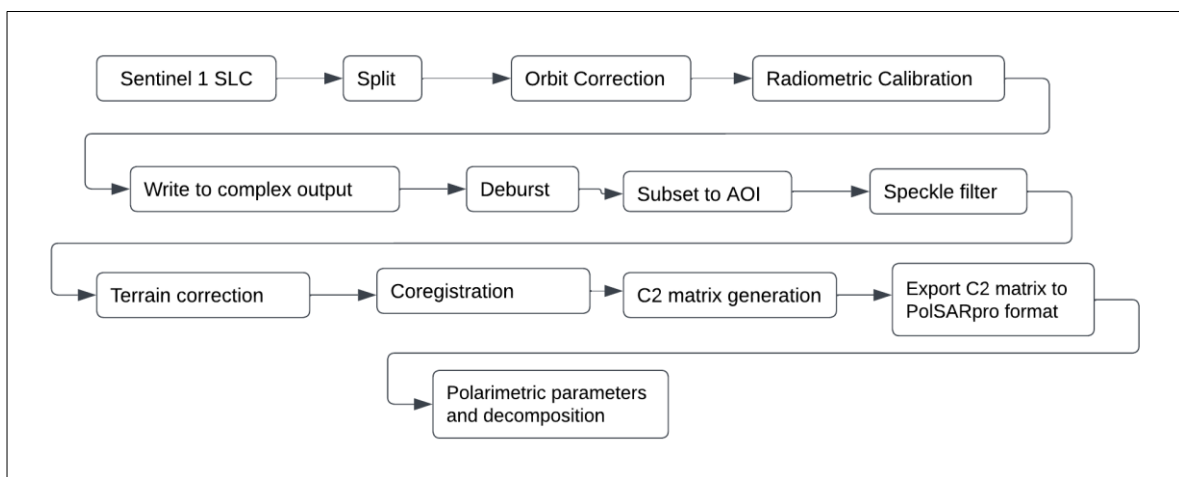
Figure 7 Flowchart of the steps followed during S1 pre-processing and generation of polarimetric and decomposition parameters.

Following operations were performed subsequently.

3.2.1.1. Polarimetric Speckle filter

Speckle filtering was applied to a subset image. Speckle is a granular, salt-and-pepper type noise that is inherent to SAR images. It is caused by the coherent nature of the radar signal, meaning that the returned waves interfere constructively and destructively, leading to bright and dark pixels in the radar image. (Emery & Camps, 2017). To address this, a Lee speckle filter (Rubel et al., 2021)was implemented in the research. Lee speckle filter is renowned for its adaptability and capability to preserve image details. The kernel size of 7x7 window was opted. This size was selected based on its effective balance between reducing noise and maintaining image detail. After a careful review of the results, the refined Lee filter with a 7x7 kernel proved to be an optimal choice for this application.

3.2.1.2. Terrain correction.

Terrain correction is a crucial step in SAR data processing. The primary purpose of terrain correction is to compensate for topographic distortion in the radar image. It transforms SAR data from the slant range or ground range geometry to map coordinates (P. Wang et al., 2013). This correction ensures that the ground features in the image are in their correct geographic locations and that the pixel values accurately represent the radar backscatter of the ground surface.

In the SNAP software, the range doppler terrain correction was performed under the radar section in the following sequence: Radar →Geometric → Terrain Correction → Range Doppler Terrain Correction. For precise topographic correction, the Copernicus 30 m global DEM was employed in the processing parameters, and the map projection utilized was the UTM WGS 84 system. Owing to its capability of generating smoother images than the nearest-neighbor method while ensuring satisfactory computational efficiency, the bilinear interpolation resampling method with a pixel spacing of 10 meters was chosen.

3.2.1.3. Co-registration.

Following terrain correction, co-registration was performed on the pre and post-hail event datasets. This step is vital when dealing with a sequence of images or when merging images from different dates, orbits, or sensors for analysis activities like change detection, interferometry, or multi-temporal studies. The implementation of the co-registration procedure is based on the cross-correlation technique (Wu et al., 2017). Co-registration guarantees that identical pixels in each image align with the same ground location, allowing for precise comparison or mathematical operations between the images. Even minor misalignment can lead to false outcomes. Given my research focus on detecting crop damage after hail events, this step is indispensable for achieving the most accurate results. SNAP software was used to perform co-registration in the following sequence: Radar →Coregistration.

3.2.1.4. Covariance Matrix.

Polarimetric data offers the capability to extract details such as orientation, shape, and dielectric properties from backscattered information. The dataset in use features two polarimetric channels: VV and VH. We utilized SNAP to produce a dual polarimetric covariance matrix. As described by (Mandal et al., 2020), we can generate the diagonal elements of the C_2 matrix for the VV and VH (dual pol) S1 data according to equation (1). In this equation, S_{VV} stands for the backscatter measured when both the transmitted and received electromagnetic waves are vertically polarized. Similarly, S_{VH} refers to the backscatter measured when the transmitted wave is vertically polarized but the received wave is horizontally polarized.

$$\mathbf{C}_2 = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} \langle |S_{VV}|^2 \rangle & \langle S_{VV} S_{VH}^* \rangle \\ \langle S_{VH} S_{VV}^* \rangle & \langle |S_{VH}|^2 \rangle \end{bmatrix} \tag{1}$$

S^*_{VV} is the complex conjugate of S_{VV}, and S^*_{VH} is the complex conjugate of S_{VH}, and the notation $\langle.\rangle$ signifies a spatial average over a moving window. The elements C_{11} and C_{22} in the covariance matrix are real numbers, representing backscatter measured of respective polarization channels. Conversely, C_{12} represents the complex covariance between the two polarizations (Mandal et al., 2020). It is a complex number composed of both a real and an imaginary part containing intensity and phase values. These elements C_{11}, C_{12}, and C_{22} are utilized later in the report as three separate bands serving as input features for a RF model.

The produced C_2 matrix was converted into PolSARpro format to facilitate the derivation of numerous polarimetric parameters, such as matrix elements, Stokes parameters, and H/Alpha decomposition. A more comprehensive discussion of the feature extraction process can be found in section 3.3. Figure 8 depicts the C_{11} channels before and after the hailstorm event extracted from S1.
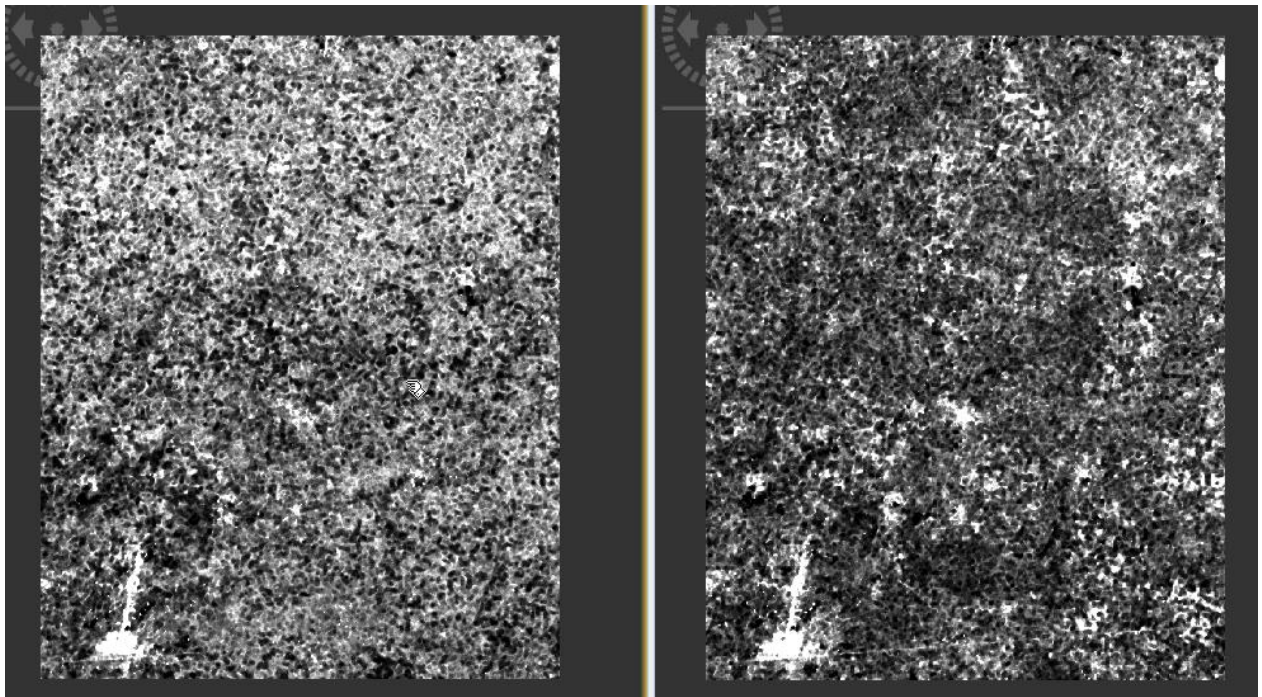
Figure 8. C11 image of pre-event (Left) and post-event (Right).

3.2.2. Sentinel-2 MSI.

S2 Level 2A images were acquired from the ESA Open Access Hub (*Copernicus Open Access Hub*, 2022). Level 2A images have undergone 'atmospheric correction', meaning the effects of atmospheric gases and aerosols have been mathematically removed from the data. This results in bottom-of-atmosphere (BOA) reflectance values that more accurately represent the Earth's surface.

The implementation of a Gaussian filter with a 7x7 kernel size was chosen for the S2 Level 2A datasets obtained before and after the hail event. It was found that this kernel size produced the most satisfactory results by suppress the high frequency without losing image information. The key motivation behind using the filter was to suppress the high-frequency noise present in the imagery, which was instrumental in enabling a smoother and more precise mapping of the damaged areas. The S2 Level 2A dataset includes bands with varying spatial resolutions: four bands (B2, B3, B4, B8) at 10 meters, six bands (B5, B6, B7, B8A, B11, B12) at 20 meters, and three bands (B1, B9, B10) at 60 meters resolution.

A resampling at a 10-meter resolution was carried out following the filtering process to maintain uniformity across all bands and enable accurate comparisons. This procedure is crucial given the distinct spatial resolutions of various bands in S2 Level 2A data. After the resampling operation, as outlined in the above section, co-registration step was performed using SNAP software under the raster section in the following sequence: Raster →Geometric →Collocation. This alignment operation is crucial, especially when working with time series data or combining data from different periods for a thorough and accurate analysis, like our current investigation of crop damage caused by the hail event. Figure 9 depicts the S2 healthy vegetation index RGB (B8, B11, B2) for before and after the hailstorm event.
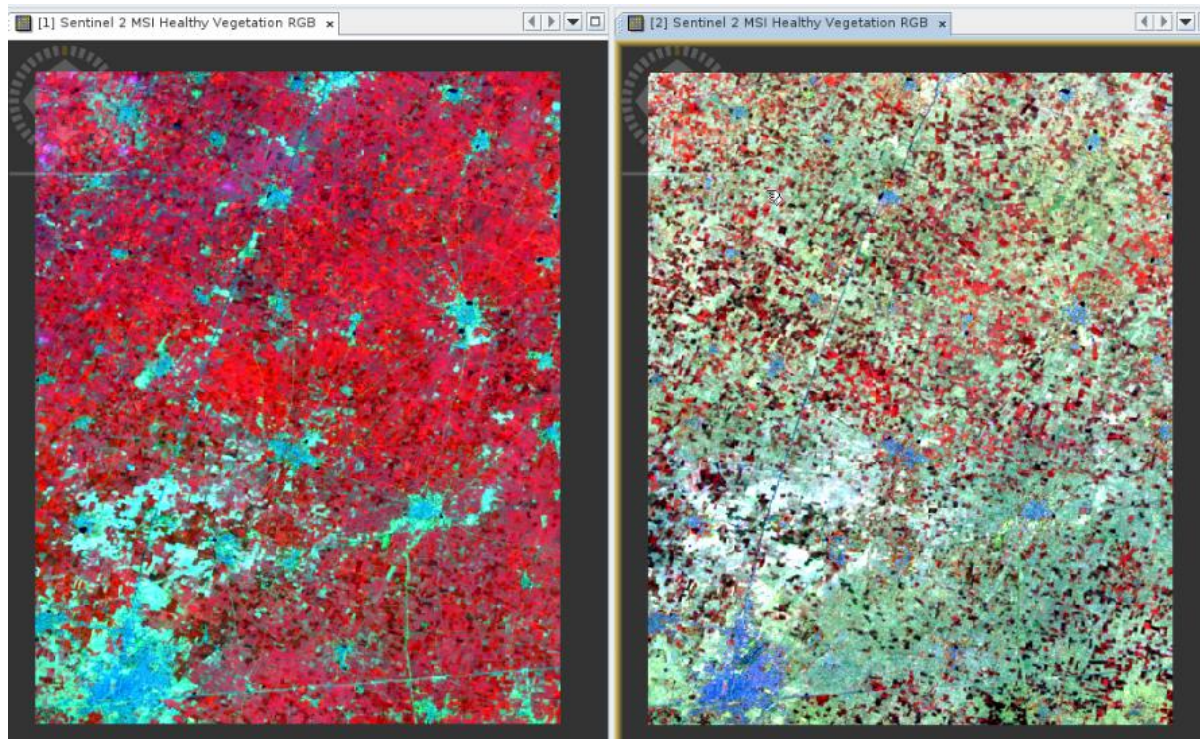
Figure 9. S2 image of pre-event (Left) and post-event (Right) with healthy vegetation index RGB (B8, B11, B2).

3.2.3. Field data.

The dataset employed for preparing the training and testing datasets required for the supervised learning process of the RF model is outlined in Table 6. The data preparation process for the cadastral data proceeded as follows. Initially, the cadastral maps for selected five villages, obtained in hardcopy format, were scanned and imported in tiff format. Following this, georectification of the digital map was performed utilizing QGIS software. Ground control points (GCPs), which were distinguishable and uniformly dispersed throughout the image, were identified within both the satellite data and the cadastral map. Notable features such as parcel boundary corners, and intersections of roads served as these GCPs. These identified GCPs were then linked, facilitating the alignment of the cadastral map points with the satellite data within the real-world geographic coordinate system. The Universal Transverse Mercator (UTM) World Geodetic System (WGS) 84 Projection system was employed during this stage of the process. Subsequently, the parcel boundaries contained within the georectified cadastral maps were digitized. Each digitized parcel was assigned a relevant parcel ID, which was then populated in an attribute field. Finally, a vector shapefile comprising these digitized parcels was created.

The RoR for five chosen villages were subsequently gathered. These records summarized information related to parcel or survey number, farmer's detail (name, father's name), land area, ownership (whether private or government-held), and land type, which include categories such as agricultural, residential, waterbody, grassland, and forest land. This data was organized according to parcel ID, which facilitated the extraction of parcel-specific information. Following this, the data was compiled into CSV files. CSV files, standing for Comma-Separated Values, are a simplistic file format used to house tabular data, like to databases or spreadsheets. Each record in the file is represented by a line, with individual fields within the record divided by commas. These prepared CSV files were partitioned by village. Using the parcel ID as a unique identifier within the CSV files made it easier to join the data with its associated parcel vector shapefile.

Subsequently, the hailstorm damage reports for the designated five villages were obtained. These reports included specific data regarding the extent of crop damage due to the hailstorm, which were classified into four distinct severity levels according to governmental standards: Low (0-33%), Moderate (33-50%), High (50-75%), and Severe (75-100%). It is noteworthy that the dataset used in this study did not contain instances of the Severe category. The link between the hailstorm damage reports and the revenue record was established via the farmer details, which served as a common attribute. A complexity was encountered when single farmers owned multiple parcels. In these circumstances, the total land area affected by the hailstorm was summarized into a single record in the hailstorm report. The process of identifying the distinct parcels that sustained damage was challenging and required the exploration of various combinations and permutations of land area to accurately discern the corresponding land parcels.

Ultimately, upon the integration of the revenue records and the hail damage report, a consolidated CSV file was produced. This file organized data on a per-parcel ID basis, which was then utilized for amalgamation with the vector shapefile. Figure 10 illustrates three components. In the top left, there is a scanned cadastral map of the village Nagla Khooba. The top right portion of the figure displays a sample of the Revenue Record. Finally, the bottom section presents a processed shapefile that includes an attribute field populated with pertinent information.

Figure 10. Scanned cadastral map of the village Nagla Khooba (Top Left), sample of the Revenue record (Top Right) and processed shapefile that includes an attribute field populated with pertinent information (Bottom).
Images have been blurred for ethical considerations

3.2.3.1.   Training and testing dataset preparation.

Polygons corresponding to non-agricultural classes such as roads, buildings, forests, water bodies, and pasture lands were excluded from the dataset. Similarly, within the agricultural parcels themselves, any parcels containing objects like buildings, huts, and iron sheds were removed. These measures were implemented to preserve the integrity and enhance the accuracy of the training and testing samples, ensuring that the focus remains strictly on agricultural land.

Moreover, for the creation of training and testing dataset parcels, the boundaries of each parcel were systematically reduced by a proportion of 10% area. This was to reduce the impact of extraneous vegetation, such as weeds and shrubs, which are often utilized to delineate parcel boundaries. These vegetative elements can introduce potential interference when undertaking damage assessments utilizing S1 and S2 data, specifically affecting backscattering and reflectance measures. By accounting for these factors and making necessary adjustments to the parcel boundaries, we can ensure a more accurate interpretation of the RS data within the training and testing dataset. Figure 11 showcases a three-part comparison. It includes the initial cadastral map, the refined parcel boundaries following the removal of irrelevant classes, and the final stage of generating the training dataset (yellow parcels), a process achieved through a systematic 10% area reduction of each parcel. A comprehensive collection of 3,394 field samples (polygon layer) was prepared for training and testing purposes.

Legend

- Reduced area (10%) used for training and validation
- Parcel boundary (after removing unwanted classes)
- Original cadastral map
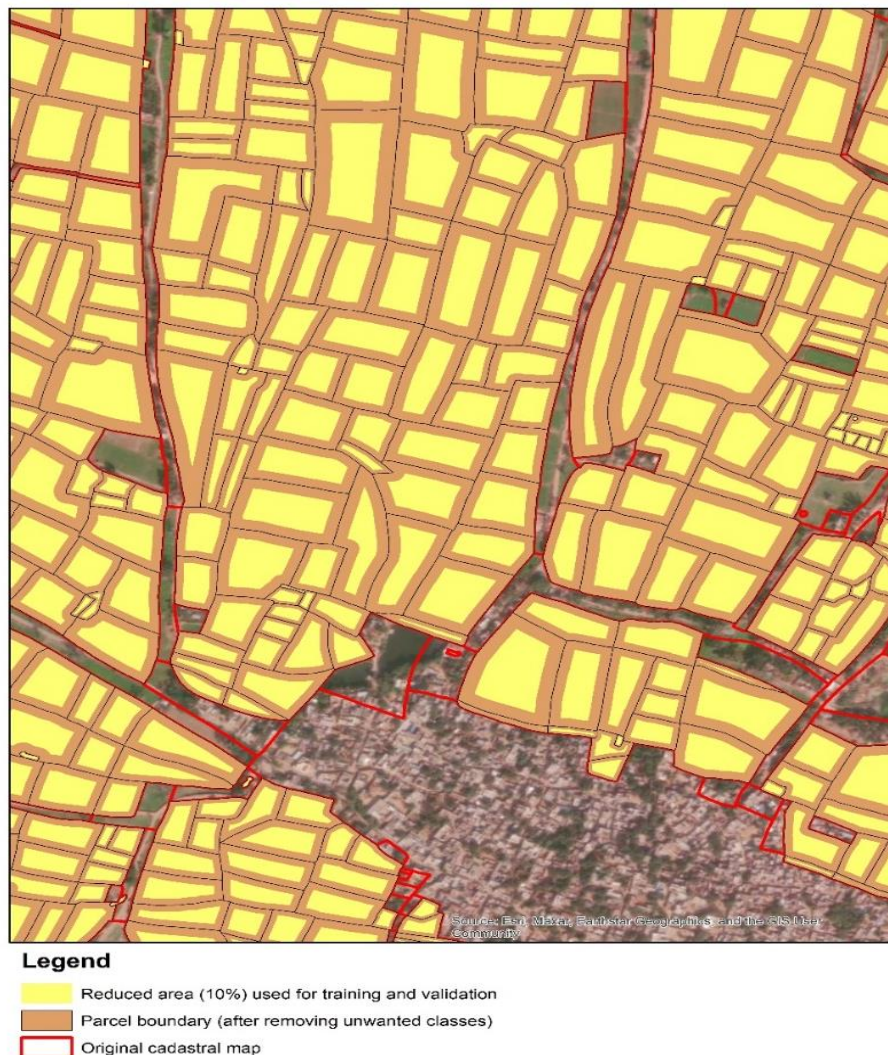
Figure 11. Parcel boundaries refinement for training and testing dataset generation.

3.3. Feature Extraction

3.3.1. Sentinel-1 SAR.

After polarimetric C2 matrix preparation as per steps mentioned in section 3.2.1, The produced C2 matrix was converted into PolSARpro format to facilitate the derivation of numerous polarimetric parameters, such as matrix elements, Stokes parameters, and H/Alpha decomposition. The following parameters were extracted into individual bands to serve as input for the model.

1. **Stokes Parameters**
    1.1. **Stokes components S0**
    1.2. **Stokes components S1**
    1.3. **Stokes components S2**
    1.4. **Stokes components S3**
    1.5. **Eigenvalue L_1**
    1.6. **Eigenvalue L_2**
    1.7. **Degree of Liner Polarization**
    1.8. **Linier Polarisation Ratio**

2. **Matrix elements**
    2.1. **C_{11}**
    2.2. **C_{22}**
    2.3. **C_{21} or $C12_{real}$**
    2.4. **Span**

3. **H/Alpha Decomposition**
    3.1. **Alpha**
    3.2. **Entropy**

4. **Dual-pol Radar Vegetation Index**
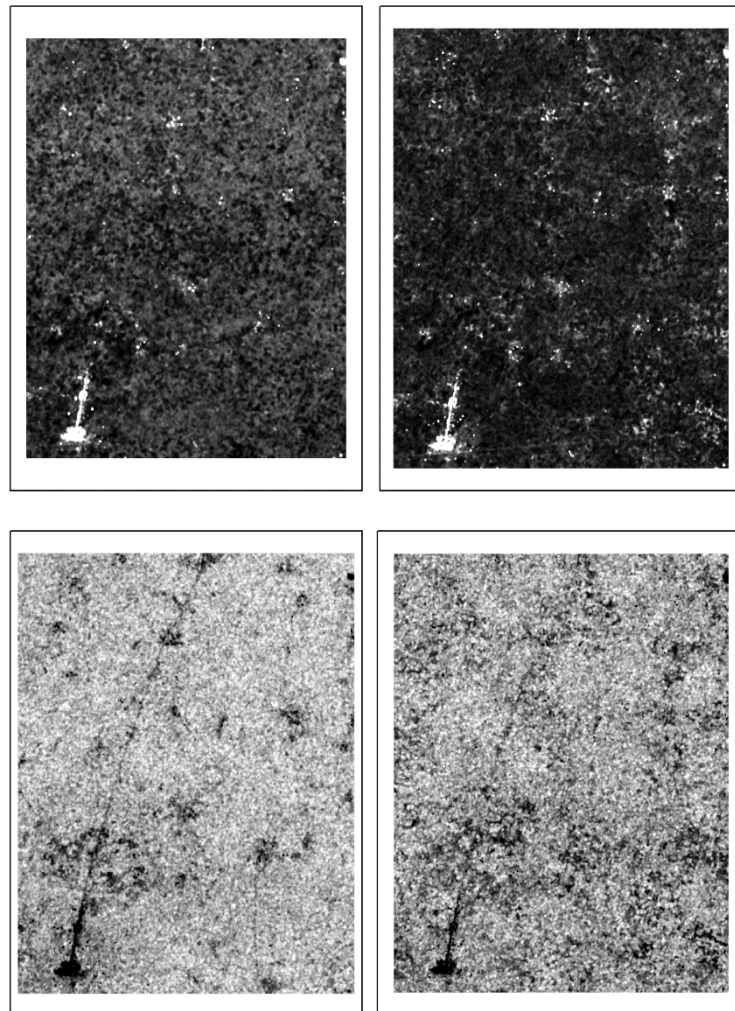
Figure 12. Showing the Eigenvalue L1 of pre-hailstorm (Upper Left) and post-hailstorm (Upper Right) conditions, and DpRVI of pre-hailstorm (Lower Left) and post-hailstorm (Lower Right) features of S1.

Figure 12 showcases the variation in Eigenvalue L1 and DpRVI for the pre and post features of S1. The top half of the figure, on the left and right, represents the Eigenvalue L1 for pre and post conditions, respectively. The bottom half, similarly divided, displays the DpRVI for the corresponding pre and post conditions.

3.3.1.1. Stokes Parameters.

Stokes Components: The Stokes parameters are used to describe the polarization state of electromagnetic waves (Lee & Pottier, 2017). For dual-polarized data, the Stokes components are calculated as follows:

$$S0 = |S_{VV}|^2 + |S_{VH}|^2 \tag{2}$$
$$S1 = |S_{VV}|^2 - |S_{VH}|^2 \tag{3}$$
$$S2 = 2 * (S_{VH} * S^*_{VV}) \tag{4}$$
$$S3 = -2 * (S_{VV} * S^*_{VH}) \tag{5}$$

3.3.1.2. Eigenvalues (L1 and L2).

The eigenvalues of the covariance matrix are found by solving the characteristic equation of the matrix (Lee & Pottier, 2017). The characteristic equation of a matrix is a polynomial equation that is derived from the determinant of the difference between the matrix and an identity matrix multiplied by a scalar, and the solutions to this equation are the eigenvalues of the matrix (Lee & Pottier, 2017). The eigenvalues are ordered in a way that L1 >= L2, and they represent the total scattering power (L1) and the difference in scattering power between different polarization states (L2).

3.3.1.3. Degree of Linear Polarization.

This measures the fraction of the electromagnetic wave energy, which is linearly polarized (Lee & Pottier, 2017). It's calculated as:

$$\text{Degree of linear polarization (m)} = \text{sqrt} (S1^2 + S2^2) / S0 \tag{6}$$

3.3.1.4. Linear Polarization Ratio.

The ratio of the minimum to the maximum scattering mechanism. It's calculated as:

$$\text{Linear polarization ratio} = L2 / L1 \tag{7}$$

3.3.1.5. Matrix elements.

C_2 matrix elements C_{11}, C_{22} and $|C_{12}|$ have been previously explained in section 3.2.1. Below are the definition and calculations of the additional parameter span. The span is the total power or simply the sum of the power in all the polarization channels (Lee & Pottier, 2017). For dual-polarized data, Span (L) is calculated as:

$$\text{Span} = |S_{VV}|^2 + |S_{VH}|^2 \tag{8}$$

3.3.1.6. H/Alpha Decomposition.

Polarimetric decomposition isolates distinct scattering elements, enabling a detailed analysis of the scattering process (Cloude & Pottier, 1997). Key parameters such as Anisotropy, Alpha (α), and Entropy, derived from eigenvalues and eigenvectors as previously detailed, offer insights into scattering mechanisms(Cloude & Pottier, 1997). Alpha reveals the type of dominant scattering mechanism with values from 0 to 90 degrees. Entropy illustrates scattering heterogeneity or randomness of the scattering process. Alpha (α) and Entropy are extracted into individual bands to serve as input for the model (Lee & Pottier, 2017).

3.3.1.7.   Dual-pol Radar Vegetation Index.

The DpRVI is derived from the dual-polarized 2x2 covariance matrix C2, as indicated in equation (10). The scattering information in terms of the degree of polarization and the eigenvalue spectrum is jointly utilized to derive the vegetation index from dual-pol SAR data. The state of polarization of an EM wave is characterized in terms of the degree of polarization (m) indicated in equation (9) as proposed by. (Barakat, 1977; Mandal et al., 2020).

$$m = \sqrt{1 - \frac{4|\mathbf{C}_2|}{(\mathrm{Tr}(\mathbf{C}_2))^2}} \tag{9}$$

.

In this context, Tr signifies the trace of the matrix, which is the sum of the diagonal elements, while |.| denotes the determinant of the matrix. The non-negative eigenvalues (L1>=L2>=0) are ascertained from the eigen-decomposition of the C2 matrix and subsequently normalized with the total power Span (Tr(C2) = L1 + L2). These eigenvalues illustrate the predominance of scattering mechanisms, leading to the introduction of the parameter beta, defined as beta = L1/Span.

This principal scattering information is modulated with the degree of polarization (m), which specifically characterizes anisotropy for dual-pol SAR data. Scattering randomness is then computed by subtracting the product of m and beta from unity, as given in the following equation by (Mandal et al., 2020).

$$DpRVI = 1 - m\beta, \quad 0 \leq DpRVI \leq 1 \tag{10}$$

It's primarily used to estimate vegetation biomass and monitor vegetation health. Its value ranges from -1 to +1, with positive values indicating higher vegetation density and negative values indicating lower vegetation density or non-vegetated areas. All the parameters described above are extracted into individual bands to serve as input for RF Model.

3.3.2.   Sentinel- 2 MSI.

In the study, S2 Level 2A datasets were employed, which underwent a sequence of pre-processing steps, including filtering, resampling, and co-registration as detailed in Section 3.2.2. A selection of 11 bands and channels was extracted from the processed data. Notably, Band 1 Coastal aerosol (443 nm) and Band 10 Cirrus (1375 nm) were excluded from this selection. Additionally, a total of 19 vegetation indices were computed for both pre and post-hail event data. In total, 30 channels/variables were used to serve as input for RF model. The selected original channels from the S2 data included the following parameters (showing band number, band name, central wavelength (nm) and spatial resolution n meters (m)):

1. Band 2 - Blue - 490 nm - 10 m
2. Band 3 - Green - 560 nm - 10 m
3. Band 4 - Red - 665 nm - 10 m
4. Band 5 - Vegetation red edge - 705 nm - 20 m
5. Band 6 - Vegetation red edge - 740 nm - 20 m
6. Band 7 - Vegetation red edge - 783 nm - 20 m
7. Band 8 - NIR - 842 nm - 10 m
8. Band 8a - Narrow NIR - 865 nm - 20 m
9. Band 9 - Water vapor - 945 nm - 60 m
10. Band 11 - SWIR - 1610 nm - 20 m
11. Band 12 - SWIR - 2190 nm - 20 m

The diverse spectral channels of S2 provide a wealth of data conducive to analysing vegetation health and detecting crop damage (Immitzer et al., 2016). Band 2, the blue band (490 nm), can differentiate between soil and vegetation and detect atmospheric particles and (Al-Gaadi et al., 2016). Band 3, or the green band

(560 nm), is advantageous during peak vegetation periods for biomass estimation and the recognition of plant stress (Delegido et al., 2011). The Red band, or Band 4 (665 nm), is especially absorbed by chlorophyll and is instrumental in computing vegetation indices such as NDVI, which can signal the vitality and potential damage or stress in plants (Hussain et al., 2013). Bands 5, 6, and 7 (Vegetation Red Edge - 705 nm, 740 nm, and 783 nm, respectively) are significant in identifying minor variances in plant health. These bands monitor the red edge, a segment in the vegetation reflectance spectrum where chlorophyll absorption shifts to leaf scattering. This can assist in identifying plant stress or damage, including hail damage (Fassnacht et al., 2016).

Band 8 (NIR - 842 nm) and Band 8a (Narrow NIR - 865 nm) are essential in gauging the structure of vegetation and evaluating its health (Löw et al., 2013). The NIR bands find use in multiple vegetation indices as they can discriminate between healthy and damaged vegetation. Band 11 (SWIR - 1610 nm) and Band 12 (SWIR - 2190 nm) are useful in distinguishing between clouds, snow, and ice and are able to identify the water content in soil and vegetation, a vital factor in comprehending plant stress or damage (Zhang et al., 2020).

Overall, the S2 bands contribute to the thorough monitoring of plant health, allowing for early crop damage detection and timely intervention and damage control. Band 1 (Coastal aerosol, 443 nm) and Band 10 (Cirrus, 1375 nm) of S2 data are not typically employed in such studies due to their lack of direct contribution to understanding vegetation health. Band 1 primarily aids in coastal and aerosol studies, capturing the violet-blue light, which is heavily scattered by the atmosphere and can be absorbed by water bodies. Conversely, Band 10 is sensitive to high-altitude cirrus clouds rather than terrestrial features and is predominantly used for atmospheric corrections. These original channels were integral to the study, as they were used to calculate the following vegetation indices. Table 8 shows the list of 19 Vegetation Indices used in classification which are extracted using SNAP software.

Table 8. Vegetation Indices used for classification.

Sr. no.	Vegetation Index	Formula (for S2)	Source
1	DVI (Difference Vegetation Index)	NIR - Red = B8 - B4	Tucker, 1979
2	GEMI (The Global Environmental Monitoring Index)	eta * (1 - 0.25 * eta) - ((B4 - 0.125) / (1 - B4)) where eta = (2 * (B8^2 - B4^2) + 1.5 * B8 + 0.5 * B4) / (B8 + B4 + 0.5)	Pinty & Verstraete, 1992
3	GNDVI (Green Normalized Difference Vegetation Index)	(NIR - Green) / (NIR + Green) = (B8 - B3) / (B8 + B3)	Gitelson & Merzlyak, 2010
4	IPVI (Infrared Percentage Vegetation Index)	NIR / (NIR + Red) = B8 / (B8 + B4)	Naji, 2018
5	IRECI (Inverted Red-Edge Chlorophyll Index)	(NIR / Red_edge - 1) = B8 / B5 − 1	(Padalia et al., 2020)
6	MCARI (Modified Chlorophyll Absorption in Reflectance Index)	[(B5 - B4) - 0.2 * (B5 - B3)] * (B5 / B4)	Daughtry et al., 2000a
7	MSAVI (Modified Soil-Adjusted Vegetation Index)	(1 + L) * (B8 - B4) / (B8 + B4 + L) where L = 1 - 2 * s * NDVI * WDVI (s is the soil line slope, NDVI is Normalized Difference Vegetation Index and WDVI is	Mulla, 2013

Sr. no.	Vegetation Index	Formula (for S2)	Source
		Weighted Difference Vegetation Index)	
8	MTCI (Meris Terrestrial Chlorophyll Index)	(B8 - B6) / (B6 - B4)	Mulla, 2013
9	NDI45 (Normalized Difference Index using Bands 4 and 5)	(B5 - B4) / (B5 + B4)	Thenkabail et al., 2000
10	NDVI (Normalized Difference Vegetation Index)	(B8 - B4) / (B8 + B4)	Lichtenthaler et al., 1996
11	PSSRA (Pigment Specific Simple Ratio A)	B7 / B4, where (Central wavelength/Bandwidth): B7 = 783 nm (15 nm) and B4 = 665 nm (30 nm)	Blackburn, 1998
12	PVI (Perpendicular Vegetation Index)	sin(a) * B8 - cos(a) * B4 where: a is the angle between the soil line and the NIR axis, in degrees.	Jordan, 1969
13	REIP (Red Edge Inflection Point)	705 + 35 * ((B4 + B7)/2 - B5) / (B6 - B5), where (Central wavelength/Bandwidth): B7 = 783 nm (15 nm), B6 = 740 nm (15 nm), B5 = 705 nm (15 nm), B4 = 665 nm (30 nm)	(Mutanga & Skidmore, 2007)
14	RVI (Ratio Vegetation Index)	NIR / Red = B8 / B4	Mulla, 2013
15	S2REP (S2 Red-Edge Position)	705 + 35 * ((B4 + B7)/2 - B5) / (B6 - B5), where (Central wavelength/Bandwidth): B7 = 783 nm (15 nm), B6 = 740 nm (15 nm), B5 = 705 nm (15 nm), B4 = 665 nm (30 nm)	(*SNAP S2*, 2023)
16	SAVI (Soil-Adjusted Vegetation Index)	((NIR - Red) / (NIR + Red + 0.5)) * (1 + 0.5) = ((B8 - B4) / (B8 + B4 + 0.5)) * (1 + 0.5)	Barati et al., 2011
17	TNDVI (Transformed Normalized Difference Vegetation Index)	sqrt ((B8 - B4) / (B8 + B4) + 0.5)	Gitelson & Merzlyak, 2010
18	TSAVI (Transformed Soil-Adjusted Vegetation Index)	TSAVI = s * (B8 - s * B4 - a) / (s * B8 + B4 - a * s + X * (1 + s * s)) where: - a is the soil line intercept - s is the soil line slope - X is the adjustment factor to minimize soil noise	Barati et al., 2011
19	WDVI (Weighted Difference Vegetation Index)	NIR - a*Red = B8 - aB4 (where 'a' is the slope of the soil line, usually determined empirically)	Naji, 2018

Spectral indices, derived from optical RS data, serve as effective tools for monitoring and quantifying characteristics such as chlorophyll content, structural composition, and moisture content of vegetation (Thenkabail et al., 2000). Indices such as the GNDVI and MCARI (Daughtry et al., 2000) are attuned to variations in chlorophyll content - a crucial parameter in detecting vegetation stress indicators such as hail-induced damage. The DVI and IPVI (Naji, 2018), on the other hand, are effective at identifying structural alterations in vegetation, a critical factor in pinpointing hail damage. Additionally, MSAVI and SAVI (Barati et al., 2011) adjust for soil reflectance, allowing for more accurate detection of stressed or damaged vegetation in areas with diverse soil backgrounds.

The NDVI, TNDVI (Gitelson & Merzlyak, 2010) and TSAVI (Barati et al., 2011), utilize the contrast between specific band reflectance to suggest potential compromises in plant health, such as those inflicted by hailstorms.

The WDVI (Naji, 2018), which accounts for soil background reflectance, is particularly useful in providing insight into vegetation health in regions with sparse vegetation. The GEMI (Pinty & Verstraete, 1992) and NDI45 (Thenkabail et al., 2000), on the other hand, reduce the effects of atmospheric disturbance and employ unique band combinations, respectively, could offer additional perspectives on potential hail-induced crop damage. Collectively, the above listed 19 indices could provide an array of tools for comprehensive hail damage detection and assessment using S2 data. The rationale for utilizing a range of bands and vegetation indices originates from the principles of the RF model. It suggests that the performance of the model can be enhanced by providing a more extensive set of variables. Figure 13 shows MCARI, GNDVI and REIP vegetation indices of pre and post hailstorm events.
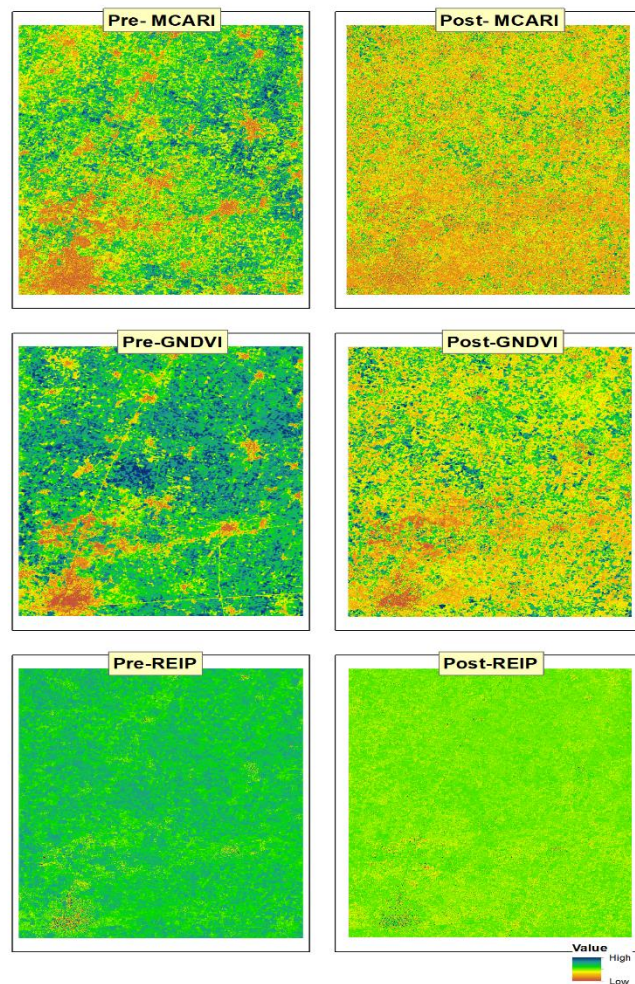
Figure 13. MCARI, GNDVI, REIP vegetation indices for pre-event (Left) and post-event (Right) of hailstorm.

3.4.   Random Forest Machine Learning Model

The RF algorithm (Breiman, 2001) is widely regarded as an effective ensemble learning method, frequently applied for both classification (Pal, 2007) and regression (Mutanga et al., 2012) tasks. The RF classifier operates by combining numerous tree classifiers. Each of these classifiers is created using a random vector, which is independently sampled from the input vector. For an input vector classification, every tree contributes a unit vote towards the class with the highest popularity (Breiman, 2001).

The architecture of the RF model is depicted in Figure 14, as detailed by (Park et al., 2020). here N are total number of training samples, and M are total features, wherein each node grew a tree using features randomly chosen individually or in combination. The method of bagging, a process that creates a training dataset by drawing n(N) (refers as *'sampsize'* parameter in R package) instances/samples at random with replacement, where N corresponds to the original training set's size (Breiman, 2001), was utilized for each selected feature combination m(M) (refers as *'mtry'* parameter in R package). Classification of any instances (or pixels) was achieved by identifying the class with the highest number of votes from all tree predictors within the forest (Breiman, 2001).

Two user-determined parameters are essential for creating a RF classifier: the number of features or variables used at each node to grow a tree and the number of trees (k) (refers as *'ntree'* parameter in the R package) to be developed. For classifying a new dataset, every instance of the dataset is introduced to each of the k trees. The forest then selects the class that obtains the majority of the k votes for that specific instance (Pal, 2007).

The RF classifier employs the Gini index as a tool to measure an attribute's impurity in relation to different classes. Mean Decrease Gini (MDG) measures the total decrease in node impurities, measured using the Gini index, from splitting on the variable and averages this across all trees. Mean Decrease Accuracy (MDA) is calculated by permuting the values of a particular feature, making predictions using the model, and measuring the decrease in accuracy caused by the permutation. A high value of either MDA or MDG indicates a variable of high importance (Díaz-Uriarte & Alvarez de Andrés, 2006). The Out-Of-Bag (OOB) error in RF is a predictive error estimate made from the instances left out during the bootstrapping process(Breiman, 1996). These unused instances, known as OOB instances, help to validate the model internally without needing a separate test set.
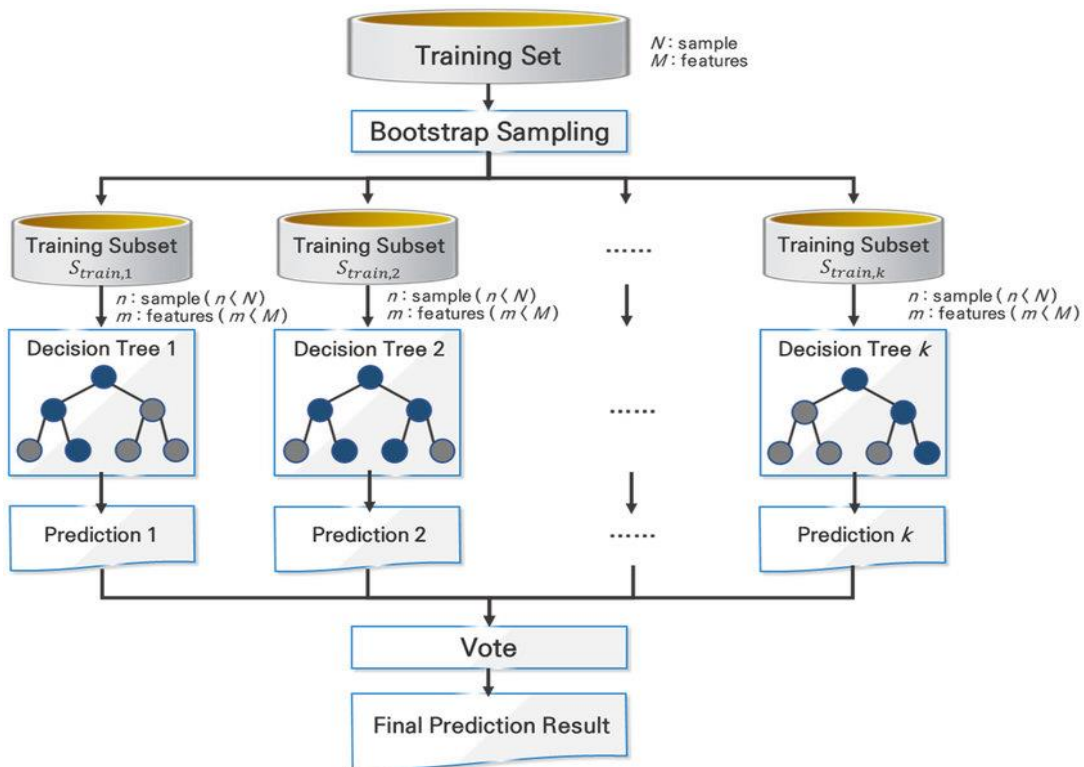
Figure 14. Architecture of the Random Forest model (Source: Park et al., 2020)

This method efficiently addresses the challenges associated with high-dimensional input data and the inherent non-linearity often found in classification problems. Additionally, RF has the added advantage of serving as a strong tool for dimensionality reduction, as it can ascertain the significance of each feature in relation to the problem under consideration (Izquierdo-Verdiguier et al., 2017). One notable advantage of this ensemble method is its inherent resistance to overfitting, a common pitfall in many ML models. This is mainly due to the diversified array of trees created during the model's construction (Breiman, 2001).

Further, RF demonstrates robustness against noise and outliers, contributing to the model's stability. They are capable of efficiently managing large datasets with high dimensionality and multiple input variables, making them versatile for complex classification problems (Liaw & Wiener, 2002). Another distinctive advantage of RF is its ability to measure and rank the importance of features, thereby providing critical insights into the underlying data and enhancing model interpretation (Díaz-Uriarte & Alvarez de Andrés, 2006).

3.5. Random Forest Pixel based Classification

3.5.1. Sentinel-1 SAR.

The input variables for the RF model were obtained by extracting features discussed in section 3.3.1. These variables were computed for both pre-event and post-event data, as well as ratio and difference bands derived from the post-event and pre-event data using the band math function in SNAP software. The formulas presented in
Table 9 were utilized to compute the difference and ratio bands. In these formulas, t2 represents the band data after the hail event, while t1 represents the band data before the hail event. The calculations involved logarithmic transformations with a base of 10, and absolute values (denoted by | |) were employed. The polarimetric matrix element bands, namely C11, C22, and C12 was explained in section 3.2.1

Table 9. S1 difference and ratio band calculation formula.

Sr	Difference band	Ratio bands	Description
1	$\|\log(C11\,t2) - \log C11\,t1\|$	$\|C11\,t2\| \div \|C11\,t1\|$	For band C11
2	$\|\log(C22\,t2) - \log C22\,t1\|$	$\|C22\,t2\| \div \|C22\,t1\|$	For band C22
3	$\|\log(\|C12\,t2\|) - \log(\|C12\,t1\|)\|$	$\|C12\,t2\| \div \|C12\,t1\|$	For band C12
4	$\|t2 - t1\|$	$\|t2 \div t1\|$	For Other bands

The dataset was processed using R-studio, involving the stacking of a total of 60 layers. Including 15 original features from pre data, 15 original features from post data, 15 features from difference data and 15 features from ratio data. Subsequently, the pixel values of these bands were normalized using min-max normalization, employing the formula $N_V= (O_V - M_V) / (MX_V - M_V)$, where N_V represents the normalized pixel value, O_V denotes the original pixel value, MX_V corresponds to the maximum pixel value, and M_V indicates the minimum pixel value in the band.

As outlined in section 3.2.3, a comprehensive collection of 3,394 field samples (polygon layer) was meticulously prepared. To facilitate PBC, approximately 12 points were randomly generated for each polygon, culminating in a total of 40,157 points. Subsequently, stratified random sampling was executed to ensure the representation of distinct damage classes, i.e. Low, medium, and high in a proportionate manner. This approach guaranteed that each damage class was adequately represented within the sample.

Moreover, the collected samples were subsequently partitioned into training and testing sets, allocating 70% for supervised training of the RF model, while the remaining 30% was designated for testing purposes.

Various combinations of bands were investigated, and different parameter values of the randomForest function in R software, including k= *'ntree'*, m(M)= *'mtry'*, and n(N)= *'sampsize'*, were tested to train the RF model and attain optimal accuracy. The importance parameter was utilized to identify significant variables based on their MDG and MDA values. Specifically, *'mtry'* determines the number of randomly selected input features considered at each split of a decision tree, often employing the square root of the total input features. *'sampsize'* indicates the number of randomly selected samples used for constructing each decision tree. *'ntree'* denotes the number of decision trees included in the RF, where increasing this value improves performance but also extends computation time.

Change detection techniques, specifically image differencing and image ratioing, are fundamental tools in RS studies (Lu et al., 2010). Given that our research focuses on identifying crop damage, a task essentially involving the detection of changes, we employed these methodologies in our investigation. As initially stated, we utilized 60 layers, including difference and ratio bands. However, we found that only the difference bands yielded the highest MDG and MDA values in every test. As a result, we ultimately decided to use only the difference band's 15 features to conduct our experiment.

3.5.2. Sentinel-2 MSI.

The RF model's input variables were obtained by extracting features, as discussed in section 3.3.2. These variables were computed for both pre-event and post-event data, including ratio and difference bands derived using the R software. A total of 30 variables consisting of 11 original channels and 19 vegetation indices were extracted for the pre-event data, post-event data, as well as their respective ratio and difference bands. The difference bands were calculated using the formula t1 - t2, where t1 represents the band data before the hail event, and t2 represents the band data after the hail event. The ratio bands were calculated using the formula t1 ÷ t2. (Lu et al., 2010)

The dataset was processed using R-studio, involving the stacking of 120 layers, with 30 layers each for the pre-data, post-data, difference data, and ratio data. Subsequently, the pixel values of these bands were normalized using min-max normalization. As described in section 3.5.1, supervised training was conducted to train the RF model. Various combinations of bands were investigated, and different parameter values of the randomForest function in R software, including *'ntree'*, *'mtry'*, and *'sampsize'*, MDG and MDA were tested to optimize the model's accuracy.

As initially stated, we utilized 120 layers. However, we found that only the difference bands produced the highest MDG and MDA values in every test. As a result, we ultimately decided to use only the difference band's 30 features to conduct our experiment.

3.5.3. Combination of Sentinel-1 and Sentinel-2.

RF has the advantageous capability of evaluating and ranking the importance of features, offering valuable insights into the underlying data and enhancing model interpretation(Díaz-Uriarte & Alvarez de Andrés, 2006). In this study, the importance parameter was employed to determine the significance of variables based on their MDG and MDA values. The model achieved the highest accuracy for both the S1 dataset and the S2 dataset.

From each dataset, the top four variables or bands with the utmost importance were selected. Subsequently, these eight bands were stacked together, and training and testing processes were conducted to attain optimal accuracies after generating the final prediction output.

After generating the final prediction output, the majority pixel classification approach was employed. This procedure incorporated a parcel boundary layer to determine the dominant crop damage class within the pixels encompassing the parcel boundaries. Subsequently, the identified majority class was assigned to the corresponding parcel boundary. This technique enables a more precise evaluation of the predominant crop damage class within each distinct parcel, thereby enhancing the accuracy of the classification results.

3.6. Random Forest Object based Classification

3.6.1. Sentinel-1 SAR.

In the OBC approach, each polygon in the field data was treated as an individual object, and 11 statistical parameters (Mean, Standard Deviation, Median, Mode, Minimum, Maximum, Quantiles, Interquartile Range, Skewness, Kurtosis) were computed. These parameters were selected based on their importance in analyzing datasets and understanding their characteristics. For instance, the mean provides an average value indicating the dataset's typical value, while the standard deviation measures variability around the mean. The median, less influenced by extreme values, offers insight into the dataset's central tendency. Mode identifies dominant characteristics, and minimum/maximum values establish the data range. Skewness measures asymmetry, and kurtosis assesses the distribution's shape. These parameters help in comprehending central tendency, spread, shape, and potential outliers.

In Section 3.5.1, a total of 15 difference band layers were finally used, while in Section 3.2.3, a comprehensive collection of 3,394 polygons of field data with damage level information was utilized for OBC. For each polygon, 11 statistical parameters were computed for each band, resulting in a total of 165 variables (11 parameters × 15 bands) per polygon. A data frame was created, comprising 3,394 rows and 165 columns. Stratified random sampling was then employed to ensure a proportional representation of distinct damage classes within the sample. This approach guaranteed adequate representation of each damage class. The collected samples were divided into training and testing sets, allocating 70% for supervised training of the RF model and reserving 30% for testing purposes.

Exploration of different column variables from the data frame was conducted, and various parameter values of the randomForest function in R software, specifically *'ntree'*, *'mtry'*, and *'sampsize'*, were tested to train the RF model and achieve optimal accuracy. The importance parameter was utilized to identify significant variables based on their MDG and MDA values. Finally, model accuracies and testing accuracies were assessed, and the class prediction was performed on the data frame to obtain the predicted output for the classified damage levels.

3.6.2. Sentinel-2 MSI.

As described in section 3.6.1, In the OBC approach, each polygon in the field data was treated as an individual object, and 11 statistical parameters (Mean, Standard Deviation, Median, Mode, Minimum, Maximum, Quantiles, Interquartile Range, Skewness, Kurtosis) were computed. As described in Section 3.5.2, 30 difference band layers were finally used for experiments. As mentioned in Section 3.2.3, a total of 3,394 polygons of field data with damage level information were utilized for OBC. For each polygon, 11 statistical parameters were computed for each band, resulting in a total of 330 variables (11 parameters × 30 bands) per polygon. A data frame was created, comprising 3,394 rows and 330 columns.

As discussed in section 3.6.1, stratified random sampling was performed. Subsequently, RF model was trained with optimal accuracies considering different parameters. Finally, the class prediction was performed on the data frame to obtain the predicted output for the classified damage levels.

3.6.3.   Combination of Sentinel-1 and Sentinel-2.

In this section, we selected the top four variables or bands from each dataset based on their MDA and MDG values. These eight bands were then stacked together, and for each polygon, we calculated 11 statistical parameters for each band. This resulted in a total of 88 variables per polygon, derived from the product of 11 parameters and 8 bands.

Following this, a RF model was trained, optimizing for accuracy with different parameters. Finally, we applied class prediction on the data frame to obtain the output, which was the predicted damage levels.

3.7.   Accuracy Assessment

The confusion matrix is a commonly used tool in assessing the performance of a model for classification tasks, whether it's binary or multiclass classification (Kulkarni et al., 2020). This matrix displays (Figure 15) the count of predicted versus real labels. True Negative, or "TN", represents the count of accurately classified negative instances, while True Positive, or "TP", represents the correctly classified positive instances. On the other hand, False Positive, "FP", shows the count of negative instances
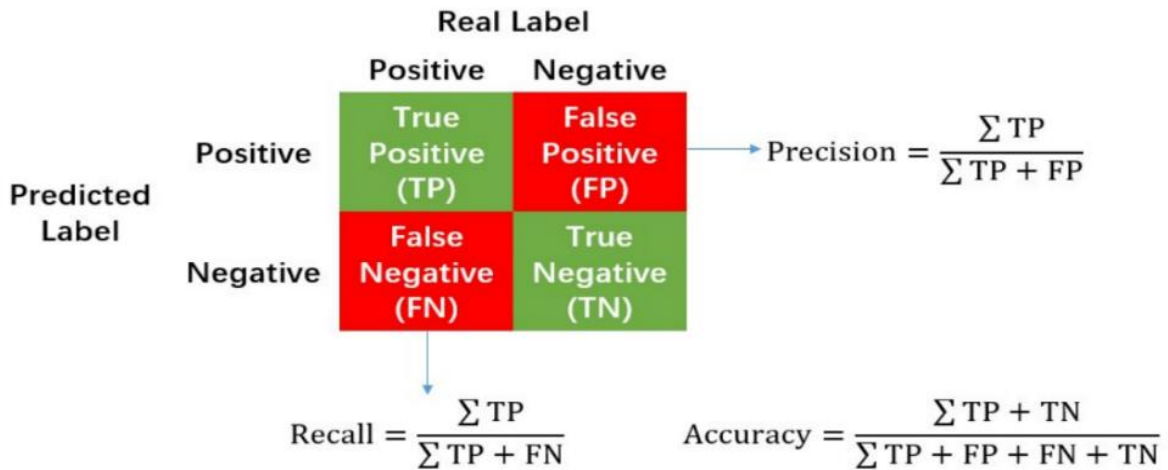
Figure 15. Confusion matrix structure and elements (Source: Ma et al., 2019).

erroneously classified as positive, and False Negative, "FN", indicates the count of positive instances classified as negative. Figure 15 represents the confusion matrix structure and its elements. The accuracy of a model is calculated using the formulas represented in Table 11. Accuracy measures the proportion of correct predictions out of all predictions made. However, this metric can be misleading when dealing with imbalanced datasets, prompting the need for alternative performance indicators derived from the confusion matrix (Kulkarni et al., 2020).

Precision or user's accuracy measures the accuracy of positive predictions made by the model. It focuses on the proportion of correctly identified positive instances out of all instances predicted as positive. A higher precision value indicates fewer false positives and better quality of positive predictions.

On the other hand, recall or producer's accuracy, also known as sensitivity, measures the proportion of actual positive instances that are correctly identified by the model. Table 9 provides an illustration of the labels of the confusion matrix in the present case.

The F1 score, a commonly employed metric, is the harmonic mean of precision and recall, offering a balanced evaluation of accuracy. This metric integrates precision and recall into a unified measure, effectively considering both positive and negative predictions in its assessment. Accuracy, F1 score, precision, and recall all have values that fall within the range of 0 to 1. A score of 0 indicates poor performance, while a score of 1 denotes perfect performance. (Ma et al., 2019). Hence, we have used F1, precision and recall to examine the model accuracies in the present study. The assessment of accuracy was conducted separately for PBC and OBC for each dataset, including S1, S2, and the merged product of both.

Table 10. Confusion matrix elements for accuracy assessment

Accuracy component	Specification in terms of crop damage classes
True Positive (TP)	The pixel/object correctly predicted as the respective damage class
False Positive (FP)	The pixel/object of other damage classes predicted as the current class
True Negative (TN)	The pixel/object correctly rejected for the respective damage class
False Negative (FN)	The pixel/object of the respective damage class predicted as other classes

Table 11.Formulas used to calculate different accuracy scores (Source: Ma et al., 2019)

Accuracy Score	Formula
Accuracy	(TP + TN) / (TP + TN + FP + FN)
Precision (User's Accuracy)	TP / (TP + FP)
Recall (Producer's Accuracy)	TP / (TP + FN)
F1 Score	2 * (Precision * Recall) / (Precision + Recall)

3.8.   Stakeholder Analysis

In relation to the current issue, the appropriate stakeholders were identified as outlined in Chapter 1, Section 1.7 The stakeholders were categorized into primary and secondary groups based on their importance in the project. A stakeholder classification matrix was created using Microsoft Excel. The interests and influence of each stakeholder were charted to help us decide the engagement strategy, whether to involve them extensively, engage when needed, keep them informed, or address their concerns. Moreover, the potential impact of interventions on all stakeholders was thoroughly assessed in Chapter 5.

4. EXPERIMENTS AND RESULTS

4.1. Random Forest Pixel based Classification

4.1.1. Experiment 1: Sentinel-1.

An experiment utilising the S1 with PBC approach explained in Section 3.5.1, was used to train and validate a RF model. We used a collection of 40,157 in situ data samples, which included 25,051 from the low damage class, 10,744 from the moderate damage class, and 4,362 from the high damage class. These samples were subdivided using a stratified random sampling technique to ensure a more accurate representation of each damage class. Specifically, 70% (28,111 samples) were allocated for training purposes, and the remaining 30% (12,046 samples) was dedicated to testing.

In its initial configuration, the RF model deployed default parameter values: *'ntree'*=500 and *'mtry'*=square root ($\sqrt{}$) of the total number of variables. The *'ntree'* parameter defined the forest's total trees, *'mtry'* indicated the number of variables randomly sampled at each node for tree construction, and *'sampsize'* prescribed the sample size for each tree's development. As outlined in Section 3.3.1, the 15 difference bands from the pre- and post-event data were calculated using eight Stokes parameters, four covariance matrix elements, two H/Alpha decomposition parameters, and the DpRVI index (see Table 9) As outlined in Section 3.5.1, these 15 difference bands were used for the present analysis.

The model underwent four iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation. The model's optimal configuration was ultimately determined as *'ntree'*=400 and *'mtry'*=3, *'sampsize'* = default value achieving optimal accuracies. Accuracy assessment was done using a confusion matrix, and F1 score, precision (user's accuracy) and recall (producer's accuracy) were extracted as detailed in section 3.7. The accuracies derived from these results are illustrated in Table 12. In evaluating the accuracy results of testing, we found that the low-damage class (0-33%) produces the highest F1 score (0.859), along with higher precision and recall values compared to other classes.

Table 12. Accuracy results of PBC using S1.

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (S1)	Low damage (0-33%)	0.842	0.899	0.869	0.785
	Moderate damage (33-50%)	0.633	0.604	0.618	
	High damage (50-75%)	0.749	0.569	0.646	
Testing (S1)	Low damage (0-33%)	0.899	0.823	**0.859**	0.763
	Moderate damage (33-50%)	0.605	0.632	**0.618**	
	High damage (50-75%)	0.535	0.694	**0.603**	

4.1.1.1. Qualitative analysis.

The qualitative evaluation for all the experiments was undertaken through visual interpretation. As depicted in Figure 16, the colours green, blue, and red correspond to low, moderate, and high damage classes, respectively.
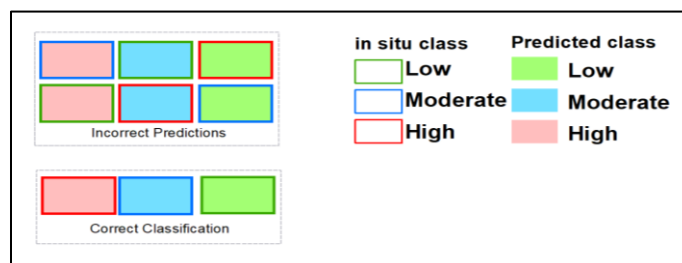
Figure 16. Qualitative assessment of correct and incorrect class predictions.

The solid colours with no border outlines represent the model's predicted outcomes for the three different damage categories, while the polygon outlines represent the damage classes based on in-situ data. Areas, where the filled-in colour (representing model prediction) matches with the colour of the polygon outline (representing in-situ data) signify instances where the damage classes were classified correctly. On the other hand, instances, where the polygon outline colour doesn't matches with the background colour indicate incorrect predictions (as demonstrated in Figure 16).

The qualitative analysis of all experiments is conducted on three tiles (Tile 1, Tile 2 and Tile 3), each from different locations within the study area.
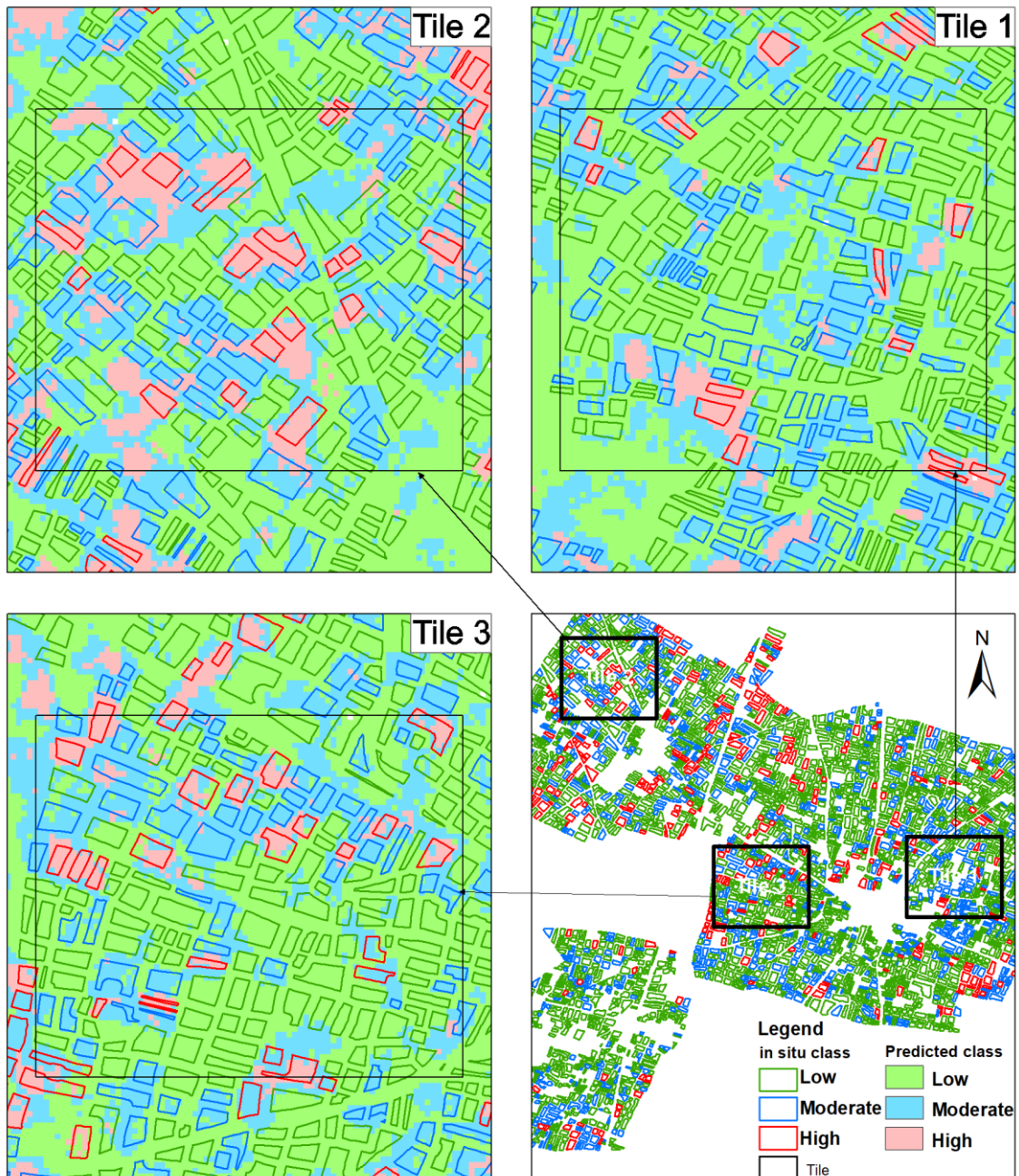
Figure 17. Sentinel-1 damage class prediction output of PBC (experiment 1).

Figure 17 displays the damage class prediction output from experiment 1. When comparing the quantitative and qualitative accuracies, the figure clearly illustrates that the low-damage classes, denoted by green, have been accurately classified with minimal intermixing of classes. On the other hand, there is a notable intermixing of classes evident within the moderate and high-damage classes. Here, class intermixing refers to pixels from different damage classes occurring within the same polygon area, leading to confusion in class attribution.

4.1.2. Experiment 2: Sentinel-2.

An experiment utilising the S2 with PBC approach explained in Section 3.5.2, was used to train and validate a RF model. As mentioned in section 4.1.1, we utilized a sample size of 40,157 in situ data, following the same approach as in experiment 1.
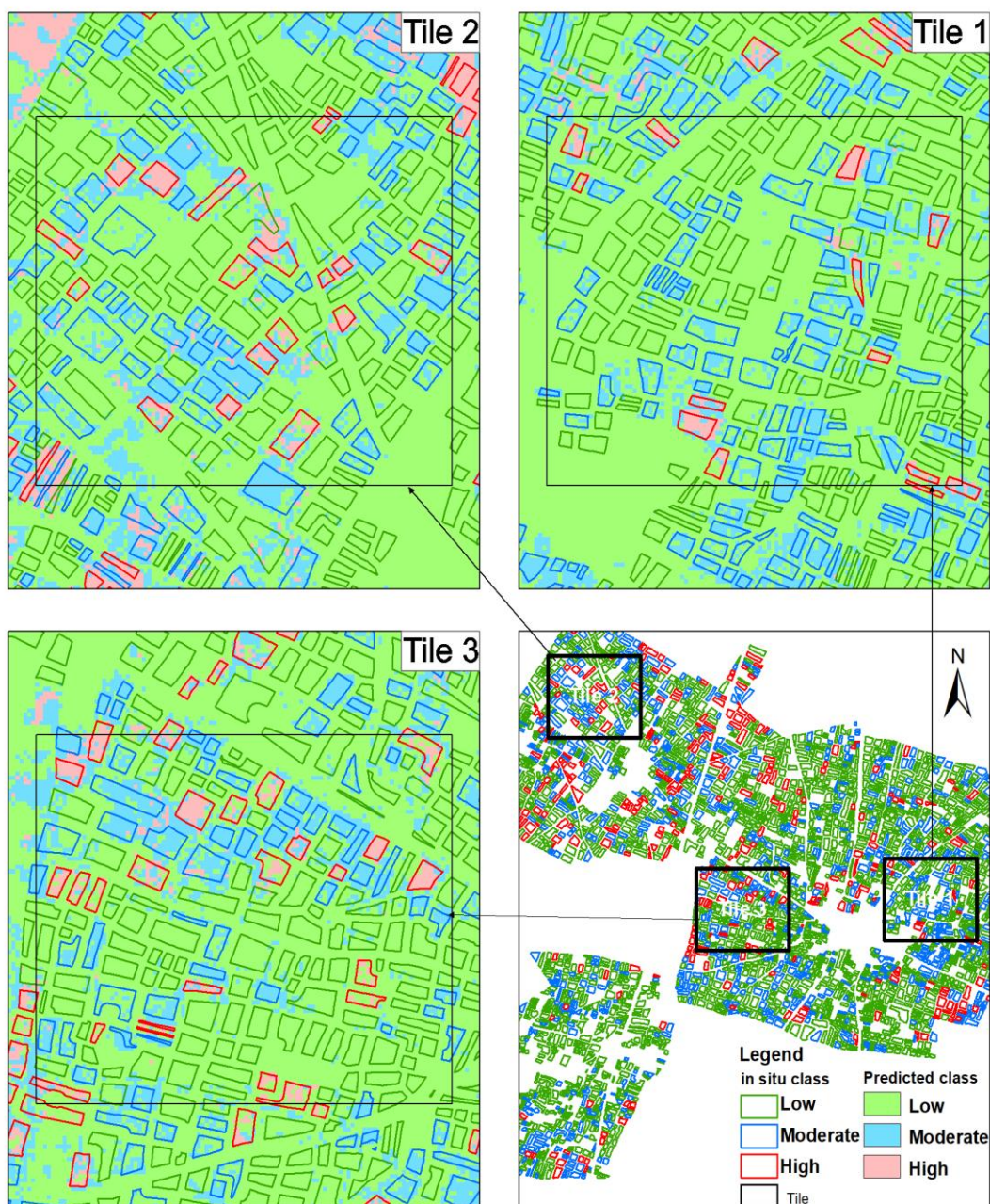
Figure 18. Sentinel-2 damage class prediction output of PBC (experiment 2)

In accordance with the methods described in Sections 3.3.2 and 3.5.2, we used difference bands comprised of 30 channels as input for the RF model in this experiment. These 30 channels were calculated by subtracting pre-event images from post-event images that included 11 original channels and 19 vegetation indices. The model underwent four iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation. The model's optimal configuration was ultimately determined as *'ntree'*=450 and *'mtry'*=12, *'sampsize'* = default value achieving optimal accuracies. The accuracies derived from these results are illustrated in Table 13.

When we assessed the accuracy results from our tests, we noted that the class representing low damage (0-33%) achieved the highest F1 score (0.775), along with superior precision and recall values compared to other classes. However, these figures fall short when compared to the results from the S1 experiment, which exhibited a higher F1 score (0.859). Notably, the accuracies for the moderate and high damage classes were significantly lower, 0.421 and 0.397, respectively, than those in the S1 experiment, which demonstrated higher accuracies, 0.618 and 0.603, respectively. The experiment's qualitative analysis was conducted through visual interpretation, as described in Figure 16 and Section 4.1.1.1.

Table 13. Accuracy results of PBC using S2.

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (S2)	Low damage (0-33%)	0.727	0.899	0.804	0.687
	Moderate damage (33-50%)	0.478	0.364	0.414	
	High damage (50-75%)	0.555	0.27	0.364	
Testing (S2)	Low damage (0-33%)	0.889	0.687	0.775	0.652
	Moderate damage (33-50%)	0.363	0.504	0.421	
	High damage (50-75%)	0.302	0.58	0.397	

Figure 18 displays the S2 experiment's damage class prediction output for 3 tiles from the study area. When evaluating both quantitative and qualitative accuracy assessments, Figure 18 shows a precise classification for low-damage classes, represented in green, with minimal class overlap. However, there is a considerable intermixing of classes within the moderate and high-damage levels. In comparison with the S1 experiment, there is a notable increase in category intermixing, which accounts for the reduced accuracy scores related to moderate and high damage classes.

4.1.3. Experiment 3: Combination of Sentinel-1 and Sentinel-2.

In this experiment, a combination of S1 and S2 datasets was used, as detailed in Section 3.5.3. We trained and validated a RF model. As mentioned in section 4.1.1, we utilized a sample size of 40,157 in situ data, following the same approach as in experiments 1 and 2. From both experiments model's most important variables or bands were identified by leveraging MDG and MDA. These were integral in achieving the highest accuracy in both experiments. From each experiment, 4 most important variables were selected, and a total of eight bands from S1 and S2 datasets were used for model training and testing, as tabulated in Table 14.

The model underwent five iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation. The model's optimal configuration was ultimately determined as *'ntree'*=350 and *'mtry'*=3, *'sampsize'* = default value achieving optimal accuracies. The accuracies derived from these results are illustrated in Table 15.
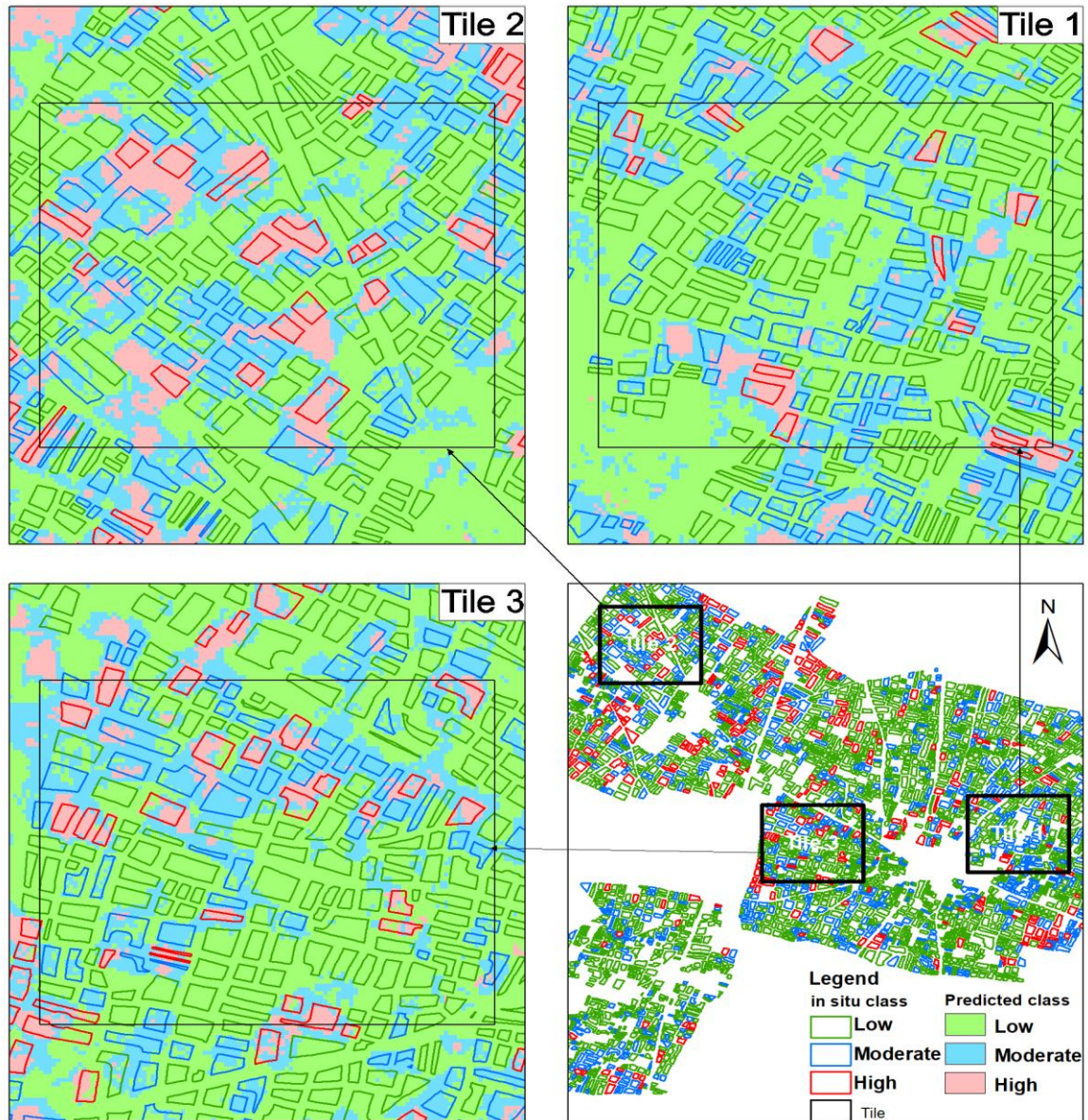
Figure 19. Combined (S1 and S2) damage class prediction output of PBC (experiment 3).

Table 14. Variables used for combined (S1 and S2).

Variables with highest MDA and MDG used for combined (S1 and S2)			
dataset	1	2	3	4
S1-difference Image (t2-t1)	|C11|	Stokes L1	Entropy	Stokes S0
S2-difference Image (t1-t2)	Band 11 - SWIR - 1610 nm	Band 12 - SWIR - 2190 nm	MCARI*	REIP

In evaluating the accuracy results of testing, we perceived that the low-damage class (0-33%) yields the highest F1 score (0.877) along with higher precision and recall values as compared to other classes. When these accuracy scores are contrasted with those obtained from the S1 experiment, they are also found to exceed the S1 experiment's F1 score (0.859). Notably, the accuracy scores for the moderate class (0.645) and high class (0.666) surpass those of the S1 experiment, which scored 0.618 and 0.603, respectively.

The qualitative analysis of this study was performed through visual examination, as evidenced by Figure 19, which displays the combined (S1 and S2) experiment's damage class predictions for three tiles within the study area. Figure 19 depicts high accuracy in classifying low-damage areas (highlighted in green) with minimal intermixing of classes. However, more class intermixing is observed within moderate and high-damage categories, leading to lower accuracies. Nonetheless, when comparing these classes from Experiment 3 with those from Experiments 1 and 2, there is a clear improvement with increased F1 scores and better precision and recall rates.

Table 15. Accuracy results of PBC using combined (S1 and S2).

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (S1 and S2)	Low damage (0-33%)	0.856	0.913	0.884	0.803
	Moderate damage (33-50%)	0.658	0.631	0.644	
	High damage (50-75%)	0.77	0.597	0.674	
Testing (S1 and S2)	Low damage (0-33%)	0.912	0.845	**0.877**	0.772
	Moderate damage (33-50%)	0.635	0.655	**0.645**	
	High damage (50-75%)	0.59	0.762	**0.666**	

To augment the precision of the classification results, a technique known as majority pixel classification was implemented. In this method, each parcel is classified as either low, moderate, or high damage, depending on the dominant pixel class within that specific parcel boundary. By adopting this method, potential misclassification errors are substantially reduced, thereby improving the accuracy of our final classification map. The accuracies derived from these results are illustrated in Table 16. This method could be applied to both S1 and S2 experiments as well, but for the sake of concision, it is only reported for the combined dataset.

In evaluating the accuracy results, we perceived that classification accuracies for all 3 classes significantly improved, and F1 scores are 0.97, 0.923, and 0.938 for low, moderate, and high classes, respectively. The experiment's qualitative analysis was conducted through visual interpretation. Figure 20 displays the combined (S1 and S2) experiment's damage class prediction output. It reveals a high accuracy classification result for all 3 classes.

Regions where the colour of the filled background (model prediction-majority class) aligns with the polygon outline (in situ data), indicate correctly classified damage classes (see Figure 16). Conversely, mismatches between the polygon outline colour and the background colour denote inaccurate predictions are marked in circles in Figure 20.

Table 16. Accuracy results of majority pixel classification using combined (S1 and S2).

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Majority Classification (S1 and S2)	Low damage (0-33%)	0.999	0.942	**0.97**	0.945
	Moderate damage (33-50%)	0.883	0.968	**0.923**	
	High damage (50-75%)	0.887	0.994	**0.938**	

Figure 20. Majority pixel classification of combined (S1 and S2) damage class prediction output.

4.2. Random Forest Object based Classification

4.2.1. Experiment 4: Sentinel-1.

Following the OBC methodology detailed in Section 3.6.1 and applying S1 data, an experiment was conducted to train and test a RF model. A sample size of 3394 polygons in situ data was used, comprising of 2138 polygons from low damage, 916 polygons from moderate damage, and 340 polygons from high damage categories.

A stratified random sampling method was employed to subdivide these samples, enhancing the representation of each damage category. Specifically, 70% of these samples (2377) were used for training, while the remaining 30% (1017 samples) were used for testing purposes.

In its initial configuration, the RF model deployed with default parameter values. As outlined in Section 3.5.1, 15 difference bands, including 8 stokes parameters, 4 matrix elements, H/α Decomposition parameter and DpRVI bands, were used for the present analysis to serve as input variables in the RF model. And for each training polygon, 11 statistical parameters (Mean, Standard Deviation, Median, Mode, Minimum, Maximum, Quantiles, Interquartile Range, Skewness, and Kurtosis) were computed. So total (11*15=165) variables were used in the model. Consequently, a data frame was structured, comprising 3394 rows (representing samples) and 165 columns (representing variables).

The model underwent five iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation.

Figure 21. Sentinel-1 damage class prediction output of OBC (experiment 4).

The model's optimal configuration was ultimately determined as *'ntree'*=450 and *'mtry'*=40, *'sampsize'* = default value achieving optimal accuracies. The accuracies derived from these results are illustrated in Table 17.

In evaluating the accuracy results of testing, we found that all 3 classes are classified with good accuracies with more than 0.85 F1 score. The low-damage class (0-33%) yields the highest F1 score (0.9564), along with higher precision and recall values as compared to other classes.

Table 17. Accuracy results of OBC using S1.

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (S1)	Low damage (0-33%)	0.9583	0.9686	0.9634	0.9583
	Moderate damage (33-50%)	0.8888	0.8941	0.8914	
	High damage (50-75%)	0.9295	0.8151	0.8685	
Testing (S1)	Low damage (0-33%)	0.9473	0.9658	**0.9564**	0.9235
	Moderate damage (33-50%)	0.8991	0.8989	**0.8990**	
	High damage (50-75%)	0.8373	0.8719	**0.8543**	

The experiment's qualitative analysis was conducted through visual interpretation. Upon comparison of qualitative accuracy assessments, Figure 21 clearly indicates the accurate classification for all 3 classes except for some misclassifications, which are marked by circles.

4.2.2. Experiment 5: Sentinel-2.

An experiment utilising the S2 with OBC methodology, explained in Section 3.6.2, was used to train and validate a RF model. As mentioned in section 4.2.1 we utilized a sample size of 3394 in situ data, following the same approach as in experiment 4.

In the initial setup, the RF model was implemented with default parameter settings. As established in experiment 2, 30 difference bands served as the input variables for the RF model. Furthermore, for each training polygon, 11 statistical parameters were computed, which included Mean, Standard Deviation, Median, Mode, Minimum, Maximum, Quantiles, Interquartile Range, Skewness, and Kurtosis. Therefore, a comprehensive set of 330 variables (11*30) was introduced into the model. Consequently, a data frame was structured, comprising 3394 rows (representing samples) and 330 columns (representing variables).

The model underwent four iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation. The model's optimal configuration was ultimately determined as *'ntree'*=400 and *'mtry'*=12, *'sampsize'* = default value achieving optimal accuracies. The accuracies derived from these results are illustrated in Table 18.

The quantitative assessment shows that the class representing low damage (0-33%) garnered the highest F1 score of 0.7703, along with higher precision and recall values compared to other classes. However, the F1 scores for the moderate and high damage classes were found to be inadequate, standing at only 0.3711 and 0.2131, respectively. The qualitative evaluation of the experiment was carried out using visual interpretation. By comparing these qualitative accuracy assessments with those from Experiment 4, it becomes evident, as depicted in Figure 22, that particularly the moderate and high damage classes have been incorrectly classified, as highlighted within the circled areas.

Table 18. Accuracy results of OBC using S2.

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (S2)	Low damage (0-33%)	0.7435	0.8931	0.8111	0.6668
	Moderate damage (33-50%)	0.4325	0.3380	0.3796	
	High damage (50-75%)	0.5556	0.1303	0.2125	
Testing (S2)	Low damage (0-33%)	0.8259	0.7218	**0.7703**	0.6427
	Moderate damage (33-50%)	0.4878	0.2971	**0.3711**	
	High damage (50-75%)	0.3889	0.1474	**0.2131**	
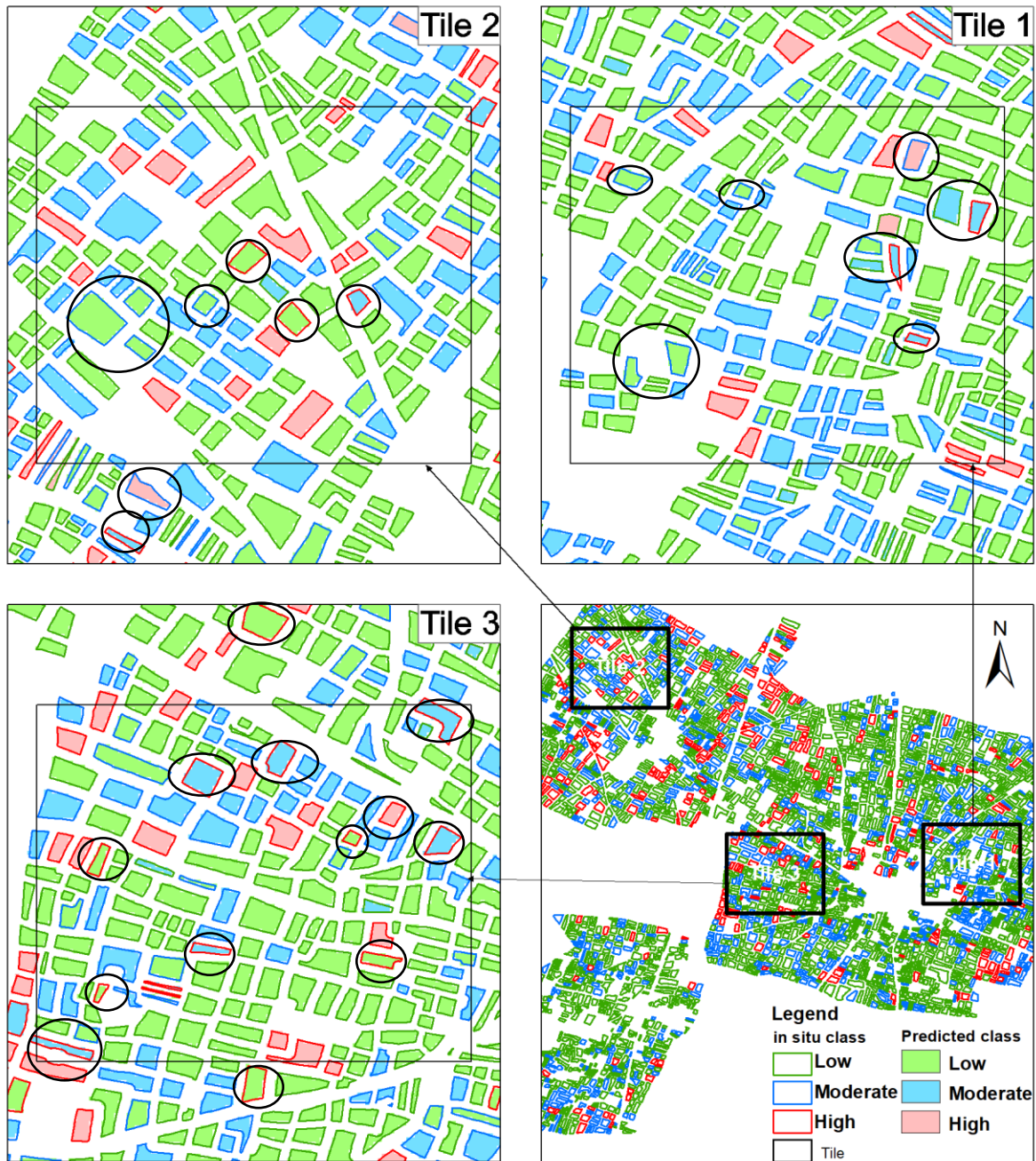
Figure 22. Sentinel-2 damage class prediction output of OBC (experiment 5).

4.2.3. Experiment 6: Combination of Sentinel-1 and Sentinel-2.

In this experiment, a combination of S1 and S2 datasets was used, as detailed in Section 3.6.3. An In-situ data sample size of 3394 was used, as described in experiment 4. From experiments 4 and 5 RF model's most important variables were identified by leveraging MDG and MDA values. From each experiment, 4 most important bands were selected, and a total of eight bands from S1 and S2 datasets were used for model training and testing. Table 19 presents the four most important bands according to MDA and MGA values of statistical variables derived from Experiments 4 and 5. We observed that Median, Mean, Q25, and Q75 played important role for OBC in experiment 4 and 5.

Furthermore, for every polygon, 11 statistical parameters were calculated, including Mean, Standard Deviation, Median, Mode, Minimum, Maximum, Quantiles (Q25, Q75), Interquartile Range, Skewness, and Kurtosis. Consequently, a total of 88 variables (11*8) were integrated into the model. The model underwent four iterations to investigate inherent uncertainty, optimising the *'mtry'* value based on the minimal OOB error. The *'ntree'* value was selected at the point of error rate stabilisation. The model's optimal configuration was ultimately determined as *'ntree'*=430 and *'mtry'*=19, *'sampsize'* = default value achieving optimal accuracies. The accuracies derived from these results are illustrated in Table 20.

Table 19. Variables used for combined (S1 and S2).

Data	Sr. no.	Difference bands	Statistical parameters with highest MDA and MDG
Sentinel 1	1	|C11|	Median, Q25, Mean, Q75, Min, Max, Skewness, SD
	2	Stokes L1	Q25, Median, Mean, Min, Q75
	3	SPAN	Median
	4	Stokes S0	Median
Sentinel 2	1	MCARI	Median, Q25, Mean, Q75
	2	B11	Q25, Median, Q75, Mean
	3	B12	Mean, Median, Q75, Q25
	4	REIP	Median, Mean

When examining the testing accuracy results, it was found that all three classes were well-classified, with F1 scores surpassing 0.80. The class indicating low damage (0-33%) produced the highest F1 score (0.8997), along with elevated precision and recall figures compared to other classes. However, we noticed that the accuracies resulting from the combined (S1 and S2) data were somewhat less than those from the experiment with S1. The combined experiment yielded F1 scores of 0.8997, 0.8111, and 0.8318, while experiment 4 achieved F1 scores of 0.9564, 0.8990, and 0.8543 for low, moderate, and high damage classes, respectively. Visually analysing the prediction output from the combined experiment establishes the connection with quantitative accuracies, Figure 23, marked with circles, indicates areas where class misclassification occurred when compared to Experiment 4.

Table 20. Accuracy results of OBC using combined (S1 and S2).

Type	Damage level	Precision	Recall	F1 Score	Overall Accuracy
Training (combined (S1 and S2))	Low damage (0-33%)	0.9269	0.9523	0.9394	0.8696
	Moderate damage (33-50%)	0.8469	0.8051	0.8255	
	High damage (50-75%)	0.8676	0.8148	0.8404	
Testing (combined (S1 and S2))	Low damage (0-33%)	0.8844	0.9154	**0.8997**	0.8332
	Moderate damage (33-50%)	0.817	0.8052	**0.8111**	
	High damage (50-75%)	0.835	0.8286	**0.8318**	

Figure 23. Combined (S1 and S2) damage class prediction output of OBC (experiment 6).

4.3. Result Comparison

4.3.1. Pixel-based classification vs Object-based classification.

Table 21 presents a summary of the accuracy results for all six experiments discussed in sections 4.1 and 4.2. Our observations indicate that the highest accuracy levels were obtained using PBC from the combined (S1 and S2) dataset, especially after assigning the majority pixel classification to each parcel. The F1 scores for the low, moderate, and high damage classes were 0.97, 0.92, and 0.94 respectively. Figure 24 illustrates the F1 score, precision (also referred to as user's accuracy), and recall (producer's accuracy) for all the experiments.

Table 21. Accuracy results from all 6 experiments.

RF model used	Experiment no	Satellite data used	Damage level	Precision	Recall	F1 Score
Pixel based RF model	1	Sentinel 1	Low	0.899	0.823	0.859
			Moderate	0.605	0.632	0.618
			High	0.535	0.694	0.603
	2	Sentinel 2	Low	0.889	0.687	0.775
			Moderate	0.363	0.504	0.421
			High	0.302	0.580	0.397
	3	Combined (S1 and S2)	Low	0.912	0.845	0.877
			Moderate	0.635	0.655	0.645
			High	0.590	0.762	0.666
		Combined (S1 and S2) with majority classification method	Low	0.99	0.942	**0.97**
			Moderate	0.883	0.968	**0.923**
			High	0.887	0.994	**0.938**
Object based RF model	4	Sentinel 1	Low	0.947	0.966	0.956
			Moderate	0.899	0.899	0.899
			High	0.837	0.872	0.854
	5	Sentinel 2	Low	0.826	0.722	0.770
			Moderate	0.488	0.297	0.371
			High	0.389	0.147	0.213
	6	Combined (S1 and S2)	Low	0.884	0.915	0.899
			Moderate	0.817	0.8052	0.8111
			High	0.835	0.8286	0.8318
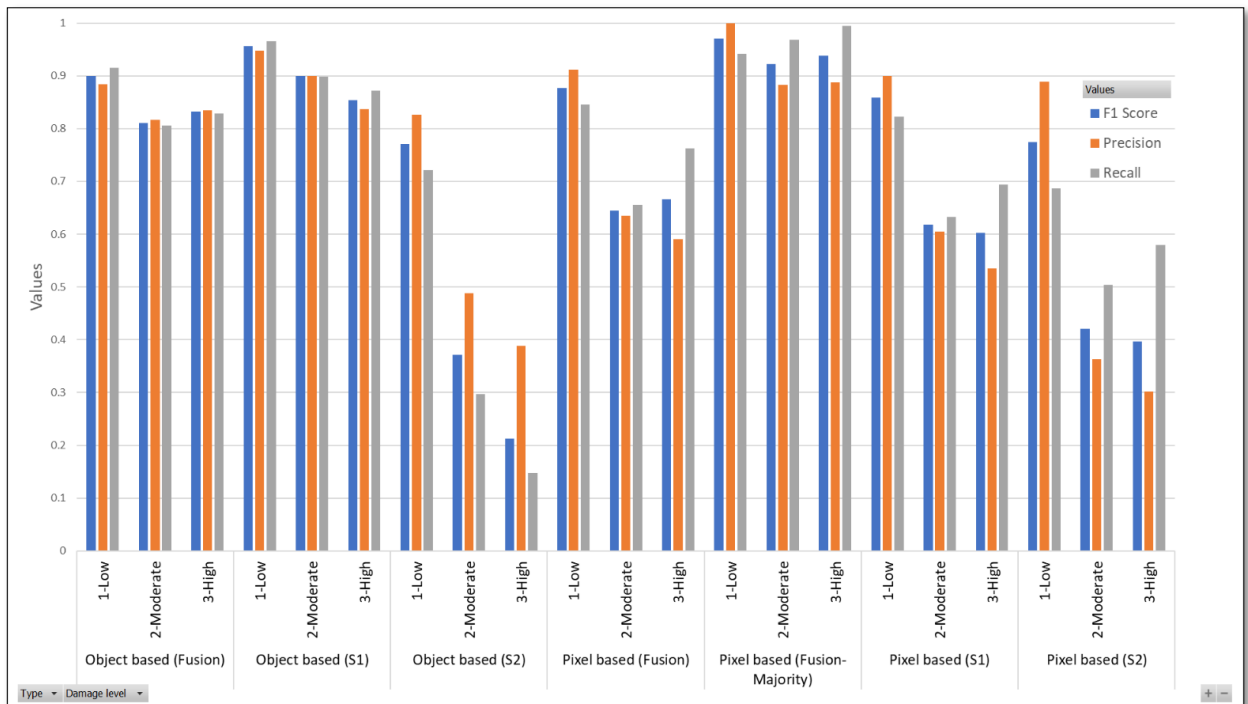
Figure 24. Accuracy comparison of PBC and OBC for S1, S2, and combined (S1 and S2).
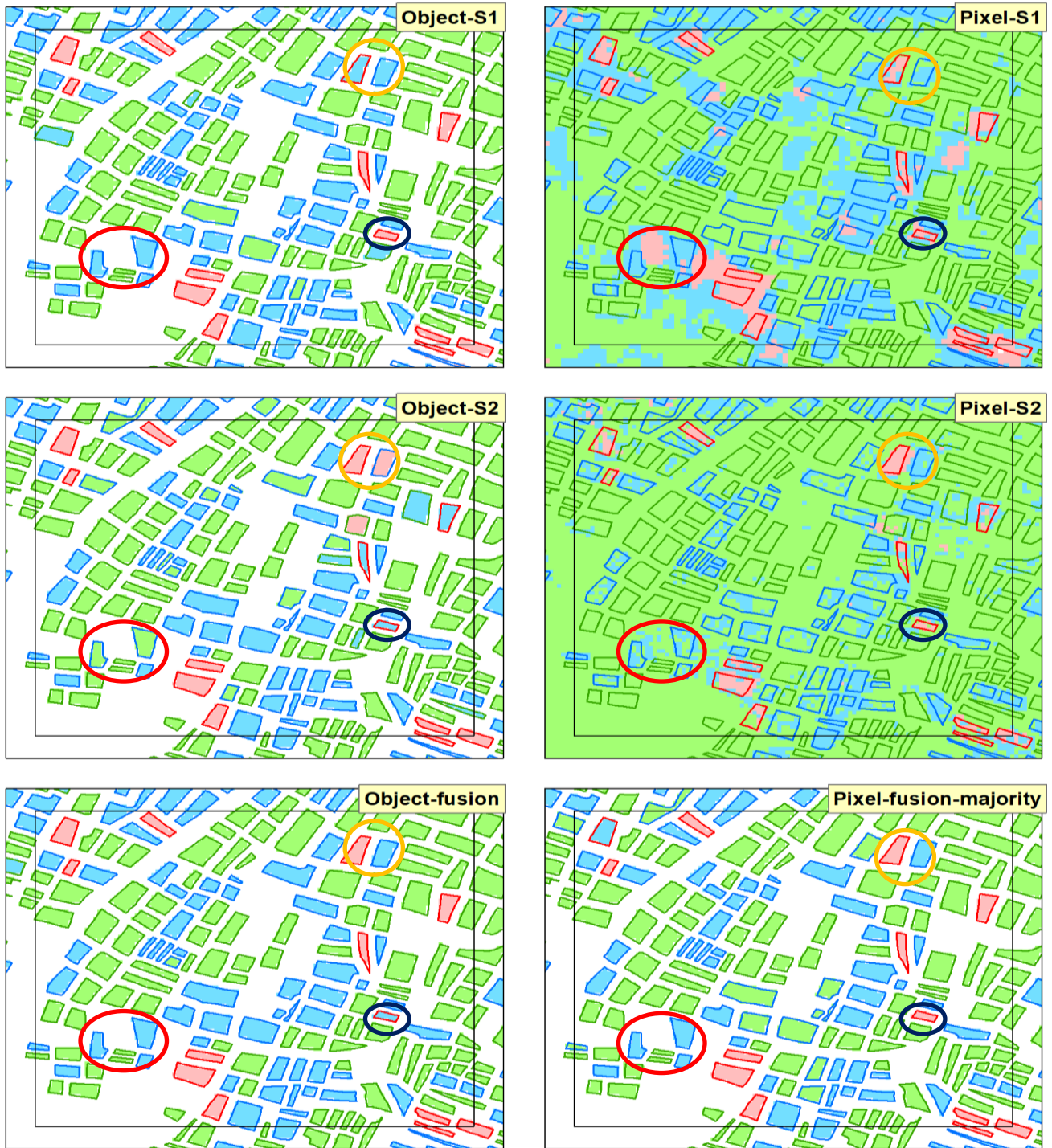
Figure 25. Qualitative variations of results from different experiments for tile 1.

We examined the qualitative accuracies using tile 1 for all experiments, which are displayed in Figure 25. Various coloured circles indicate the same location in tile 1. The yellow circle highlights that the classes within the circle are correctly classified in PBC, although there is intermixing of classes observed in the S2 data compared to the S1 data. Nevertheless, it is classified correctly in OBC with combined data. The blue circle reveals that PBC offers more accurate results compared to OBC. S1 with OBC also being accurate in this instance. The red circle illustrates correct classifications with the exception of OBC with S2. Ultimately, we found that PBC, when coupled with majority classification, yields the most accurate results.

5.   STAKEHOLDER ANALYSIS

5.1.   Stakeholder Analysis

As explained in Section 1.7, stakeholders were identified and classified into primary and secondary groups, depending on their importance. Their roles were subsequently outlined. This section delves into the exploration of the connections among stakeholders, including their interests, importance and their respective levels of influence or power. These groups were then evaluated using a scale from 1 to 10. As previously stated, this evaluation was carried out using stakeholder interview with farmer and insights were also obtained through a comprehensive review of the existing literature. Table 22 summarises this information, and Figure 26 visualizes the stakeholder classification matrix developed by the scale factor of importance and influence.

Table 22. Stakeholder with Interest, Importance, Influence and Strategy for current research problem (adapted from Olander & Landin, 2005).

Stakeholders	Type & Impact of Hailstorm	Source of Income	Interest	Importance with scale	Influence with scale	Strategy
Farmers	Primary-Direct Impact	Predominantly from the sale of crops and/or leasing of land	Quick and accurate damage assessment and fair compensation	9.5 (Directly affected by the lack of transparency, quality and time frame of damage assessment and compensation)	3 (Limited influence over the assessment and compensation process)	Address their concerns
Agricultural Labourers	Primary-Direct Impact	Wages earned from farming activities	Retention of job security and income stability despite crop damage	8 (Their livelihoods are directly affected, and they don't get any compensation due non ownership of land)	2 (No direct influence over the process)	Address their concerns
Rajasthan State Government	Primary-Indirect Impact	Various forms of state revenue such as taxes, royalties, central government grants, etc.	Efficient use of resources in disaster management, considering a fixed budget	9 (Ensures all processes work smoothly and farmers are compensated through the departments involved)	10 (Control over fund allocation and departmental directives)	Involve extensively
DMR, Rajasthan Government	Primary-Indirect Impact	State and central government funding through disaster response funds	Efficient centralised processing of damage reports and compensation disbursement within the constraints of a fixed budget	8 (Directly responsible for compensation management)	9 (Controls execution of the compensation process)	Involve extensively

Stakeholders	Type & Impact of Hailstorm	Source of Income	Interest	Importance with scale	Influence with scale	Strategy
Board of Revenue, Rajasthan Government	Primary-Indirect Impact	Funded by the Rajasthan state government	Streamlined and cost-effective assessment process within the constraints of a fixed disaster response budget	8 (Their actions directly affect the compensation received by farmers)	8 (Controls execution of the survey of the extent of damage)	Involve extensively
Local Communities	Secondary-Indirect Impact	Diverse sources, largely related to the agricultural sector, including agro-based jobs	Community resilience in the face of agricultural disasters	7 (Community livelihood is heavily dependent on agriculture and indirectly affected by crop damage)	2 (No direct influence over the process)	Address their concerns
Agriculture Department, Rajasthan Government	Secondary-Indirect Impact	Funded by the state government	Implementation of resilient farming practices to mitigate hailstorm impacts	5.5 (Their policies can lead to improved resilience but don't affect immediate post-disaster response)	6 (Not directly involved in the process but provide policy guidance and provide feedback for optimal agriculture practices)	Involve extensively
Central Government	Secondary-Indirect Impact	National revenue sources	Timely disbursement of NDRF funds for disaster management	4 (Fund disbursement affects overall damage mitigation)	8 (Disburses funds but not directly involved in local level execution, this is not the matter of central govt. directly)	Involve as needed
Print and Electronic Media	Secondary-Indirect Impact	Revenue from advertisements, subscriptions, etc.	Fair and comprehensive coverage of hailstorm events and subsequent damage	5 (Their coverage affects public perception, government decisions and guidelines, but they are not directly involved in disaster management)	7 (Can influence public perception and, indirectly, policy decisions)	Involve as needed
Crop Insurance Providers	Secondary-Indirect Impact	Premiums collected from crop insurance policies	Fair assessment and prompt payouts of insurance claims to the associated farmers	4 (Affects insured farmers' financial recovery)	5 (Determines terms of coverage and pay-outs)	Involve as needed

Stakeholders	Type & Impact of Hailstorm	Source of Income	Interest	Importance with scale	Influence with scale	Strategy
Research Institutions	Secondary-Indirect Impact	Various sources like grants, collaborations, etc.	Development of cutting-edge methodologies for damage assessment and compensation calculation	4.5 (Their research can lead to long-term improvements in disaster response)	6 (Potential influence by shaping future policies)	Involve as needed
NGOs	Secondary-Indirect Impact	Donations, grants, etc.	Providing support to affected communities and ensuring fair practices during disaster response	3.5 (Their efforts can augment government assistance and ensure fair practices)	6 (Can influence the narrative and assist in relief efforts)	Involve as needed
Local Supply Chain Entities	Secondary-Indirect Impact	Sales of agricultural products, services, and related items	Mitigation of supply chain disruptions due to crop damage	4.5 (Their functioning affects the local economy but is not directly related to damage assessment)	2.5 (Affected by but doesn't influence damage assessment process)	Keep informed

In the context of the present research, the term 'importance' signifies the level of attention that needs to be given to meeting the requirements and concerns of various stakeholders involved in the project (Olander & Landin, 2005). For instance, farmers and agricultural labourers are considered highly important stakeholders as they are directly impacted by hailstorms, and the effectiveness of damage assessment and compensation process can significantly affect their livelihoods and economic well-being. Conversely, 'influence' refers to the power or control that these stakeholders carry over the project's direction, actions, or outcomes. For example, the Rajasthan State Government, DMR, BoR. They control the policy decisions, fund allocation, and the execution of the assessment and compensation processes. Their decisions and actions directly impact how effectively the hailstorm damages are assessed and compensated, and therefore, they hold significant influence over the project's outcomes.

After analysing the interests, importance, and power dynamics of the stakeholders identified in Table 22, a stakeholder classification matrix (Figure 26) was designed to categorize them on the basis of influence and importance. This categorization helps to identify the appropriate engagement strategies for the research project, determining whose needs should be addressed, who should be extensively involved, who should be involved on an as-needed basis, and who should be kept informed.
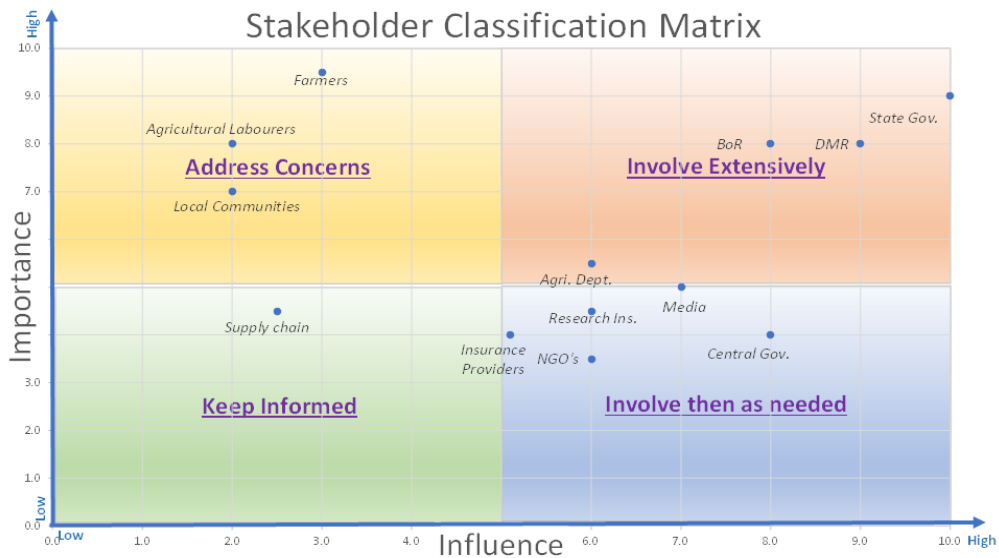
Figure 26. Stakeholder classification matrix (Source: Olander & Landin, 2005).

Primary stakeholders (directly impacted) stakeholders include farmers and agricultural labourers. These groups have a direct livelihood stake in agriculture and are adversely affected by hailstorm damages. Their interests lie in quick and accurate damage assessment and fair compensation. However, they have limited influence over the assessment and compensation processes.

Therefore, it is crucial to address their concerns effectively. Additionally, Local communities' secondary stakeholders (indirectly impacted) are largely linked to the agricultural sector through agro-based jobs. Hence any agricultural disaster indirectly affects them. The resilience of these communities, and their ability to adapt and recover in the face of such disasters, is a significant area of focus. On the other hand, primary stakeholders (indirectly impacted), such as the Rajasthan State Government, DMR, and BoR, are more engaged in disaster management's administrative and financial aspects. Their interests revolve around the efficient use of resources, accurate damage assessment, and smooth compensation disbursement. They hold significant importance and influence in the process and thus need to be extensively involved. Furthermore, the Agriculture Department, primarily tasked with implementing policies related to resilient farming practices, holds substantial importance and influence. Therefore, it necessitates extensive involvement.

Secondary stakeholders, which include entities like the Central Government, Media, Crop Insurance Providers, Research Institutions, NGOs, and Local Supply Chain Entities, play roles that indirectly contribute to the hailstorm damage assessment and compensation processes. While their interests and influences vary, they generally contribute to broader aspects of disaster management, policy direction, financial coverage, public perception, research development, community support, and local economic continuity. The matrix illustrates the intricate interaction between various stakeholders during a hailstorm, emphasizing the need for strategic engagement and coordination to ensure an efficient and fair response.

5.2. Stakeholder Impact

After stakeholder analysis, as described in section 5.1, the stakeholder relationships have been examined through the lenses of equity, equality, and empowerment concepts. (Espinoza, 2008) explains that the core principle of 'equity' theory is the belief that justice or fairness in social interactions is achieved when resources, rewards, and punishments are allocated in proportion to one's input or contributions. On the other hand, equality is defined as everyone having equal access to resources (Espinoza, 2008), whereas empowerment is to give power to someone to do something (Völker & Doneys, 2020). Presently the farmers who have damages less than 33% are not entitled to compensation (Interview, 2023). Further compounding the issue is the inconsistent categorization of similar types of damage due to a lack of transparency in the process. And the farmers utilising their influences to get compensation for undamaged or harvested crops creates inequality (Interview, 2023). Also, the amount of compensation is significantly lower in terms of investment in seeds, agriculture equipment, seeds fertilizers, workforce, and irrigation to grow the crops (Interview, 2023).

The broad ranges used to classify damage levels lead to disproportionate compensation, resulting in an inequitable distribution of resources. For instance, a farmer who has experienced 33%, 50%, or 75% damage is awarded the same compensation as another farmer who has underwent 50%, 75%, or 100% damage, respectively. This lack of proportionality in compensation with respect to the actual extent of damage experienced contributes to disparities in resource accessibility among farmers (Interview, 2023). There is an imbalance between powers as farmers don't have access to know how this assessment has been done. On the other hand, government officials must comply with financial allotment and resources available for such purposes, and they are biased in their decision-making due to this. The unavailability of any damage assessment reference system to check whether the assessment is done properly or not is a big concern.

Figure 27 shows the current placement of stakeholders within the stakeholder classification matrix, and Figure 28 illustrates the transition in stakeholder positions following the execution of the intervention. In this current research, the main focus is on farmers, who are anticipated to gain empowerment through this RS based damage assessment model. It will enhance transparency in the damage categorization process, fostering an environment of equality when dealing with shared resources.
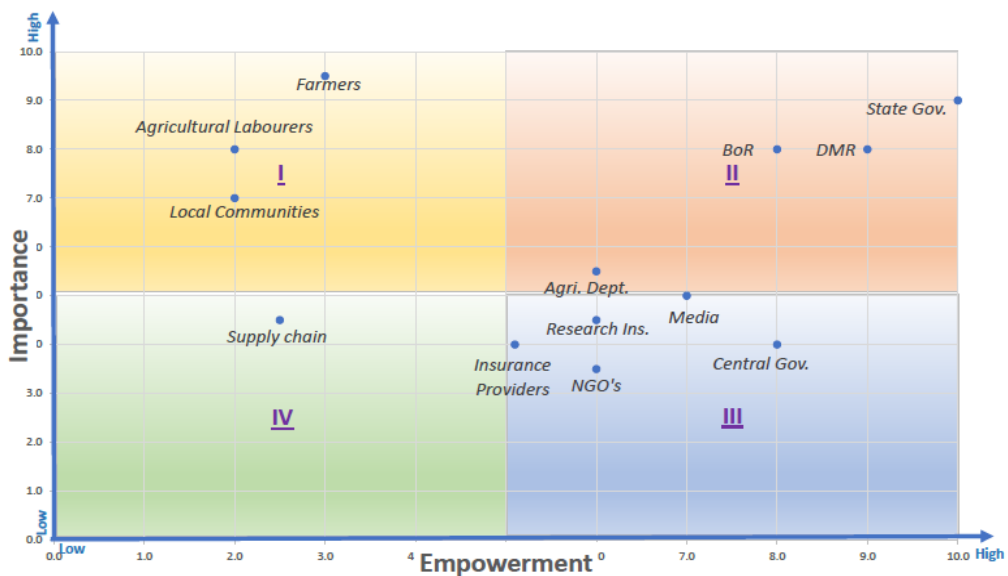
Figure 27. Current placement of stakeholders within the stakeholder classification matrix.

This shift in practice can be visualized as a shift to the second quadrant, as depicted in Figure 28, which symbolizes a transition towards a more equitable distribution of resources. Consideration must be given to ensuring the accessibility of this RS based damage assessment information to farmers. It is anticipated that the government's intention is to promote transparency in the damage assessment process, and thus, information derived from the RS method will be within reach for farmers.

To ensure the accuracy and reliability of the model, it is recommended that the damage categorization data derived from the model is validated through field checks on a specific percentage of cases, for instance, 5-10%. This cross-verification process not only bolsters the credibility of the assessment but also augments time efficiency and transparency. The acceleration in processing time has further implications. By expediting the damage compensation, farmers can secure their financial stability and promptly prepare for the next crop cycle. This enhanced productivity, in turn, can provide a significant boost to the local economy and augment food security within the region. The possible impact on other stakeholders is represented in Table 23.
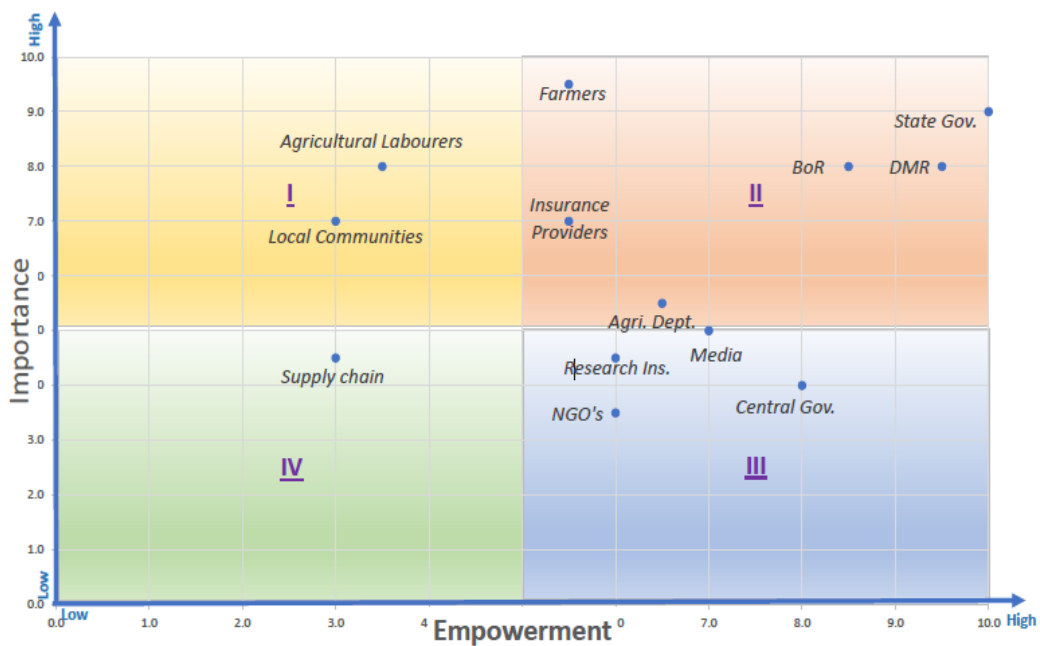
Figure 28. Illustrating the transition in stakeholder positions following the execution of the intervention.

Table 23. Impact of the intervention on other stakeholders.

Stakeholders	Description	Impact/Function
Agricultural Labourers	The financial security of farmers or landowners due to accurate crop loss assessment and timely compensation distribution can improve the condition of agricultural laborers. This could lead to increased employment opportunities and enhance their financial condition. They will experience a sense of empowerment as depicted in Figure 28.	Possible Impact
Local Communities	The proposed intervention stands to impact local communities positively, given their interdependence with the farming population. Improvement in farmers' conditions, driven by transparent and timely damage assessments, could potentially lead to an upswing in employment opportunities, not only in farming but also in sectors closely tied to agriculture. Consequently, these communities will experience a sense of empowerment, with a discernible increase in their empowerment, as depicted in Figure 28.	
Rajasthan State Government	The proposed process could enhance trust in governmental systems, although it might lead to a rise in compensation amounts. To counterbalance this effect, strategies could be established to involve crop insurance companies, encourage greater farmer participation, and use premium collections from unaffected farmers to offset the cost of crop damages. Conversely, if the state is currently allocating resources to conduct field inspections, employing Remote Sensing (RS) would prevent the need for such efforts or considerably lessen the time and number of instances requiring on-site presence. This approach could contribute to significant cost savings for the state.	
DMR, Rajasthan Government	The DMR will become more empowered (depicted in Figure 28) as they will be responsible for the quick disbursement of compensation. This intervention can help the DMR achieve its goals effectively.	
Board of Revenue, Rajasthan Government	The BoR will be more empowered (depicted in Figure 28) by having a platform to identify damages and minimise field efforts. Moreover, this methodology offers a means to oversee and prevent any unethical practices, thus fostering transparency in the assessment process - a solution to a significant challenge often faced in damage assessment. However, if the compensation amount increases due to accurate assessment, it requires policy discussions and engagement with stakeholders who have more influence.	

Stakeholders	Description	Impact/Function
Agriculture Department, Rajasthan Government	The Department of Agriculture is expected to take an active part in the application and orchestration of the RS-driven crop damage assessment model. They may be required to adjust their existing operations and work in unison with other stakeholders to ensure effective implementation. This undertaking will provide the department with a sense of empowerment (depicted in Figure 28).	Possible Impact
Crop Insurance Providers	Crop insurance providers can benefit from accurate crop damage assessment as it helps in determining the actual damages and payouts to affected farmers. They may need to collaborate with the government and other stakeholders to ensure seamless integration of the RS-based model into their insurance processes. They will be empowered as more farmers would be associated with insurance. Moreover, their significance within the system is set to rise due to the financial strength they acquire through intensified collaboration with the farming community (depicted in Figure 28).	
Local Supply Chain Entities	Local supply chains entities, such as agricultural input suppliers and traders, may be impacted by the accurate assessment of crop damage. It can help them plan their supply and distribution activities better, considering the actual crop losses. They will experience a sense of empowerment as depicted in Figure 28.	
Central Government	The central government may need to provide support and resources for implementing the RS-based model on a larger scale. They may also need to coordinate with state governments and other stakeholders to ensure smooth implementation.	Supportive Function
Print and Electronic Media	The media can play a role in creating awareness about the benefits of the RS-based model and its impact on farmers, labourers, and local communities. They can also provide updates on the implementation progress and communicate success stories	
Research Institutions	Research institutions can contribute to the development and improvement of the RS-based model. They can provide expertise, technical support, and research findings to enhance the accuracy and efficiency of the assessment process.	
NGOs	Non-governmental organizations can play a role in supporting the implementation of the RS-based model. They can provide assistance in raising awareness, capacity building, and advocating for the interests of farmers and local communities.	

6. DISCUSSION

This chapter discusses the research outcomes acquired for hailstorm damage evaluation using S1, S2, and combined datasets, as discussed in Chapter 4. In Section 6.1, "Performance Analysis of RF Model and Datasets," we discuss the role of the RF model and various datasets in identifying crop damage classes. In Section 6.2, "Addressing and Mitigating Wicked Problem," the discussion is dedicated to tackling wickedness within the scope of our research. Finally, in Section 6.3, "Limitations ", we discussed the constraints of the current study.

6.1. Performance Analysis of RF Model and Datasets

All the conducted experiments and their corresponding results have been summarized in Table 21. Implementing numerous combinations of input data on both PBC and OBC RF models to derive crop damage classifications discovered that the PBC analysis delivered superior results when used with the combined dataset, particularly after applying majority classification. It became evident that combining optical and SAR data bolstered the accuracy compared to their individual use in PBC. On the other hand, for OBC, where statistical parameters were employed, a slight decrement in accuracy was observed compared to using the S1 dataset alone, as detailed in section 4.2.3.

Our observations of F1 scores indicated that the S1 data was more successful in accurately predicting damage classes than S2 in both models. One potential reason for lower S2 accuracies could be the differences in acquisition dates of the post-event imagery; the S2 imagery was from 25th March 2020, while the S1 image was taken on 17th March 2020. Considering the hailstorm event occurred on 5th March, it is plausible that images captured closer to the event might yield more insightful information than those taken later. Additionally, there is a possibility that farmers might have commenced ploughing for the next crop season, influencing the observed outcomes. Secondly, the sensitivity of microwave signals to changes in physical structure and geometry renders them particularly advantageous for damage detection, thereby providing S1 with an edge over S2.

The impact of a hailstorm on the ground is not homogeneous, resulting in patchy crop damage with variability within the same field. Capturing this information is useful if farmers demand to increase the damage bin size (Interview, 2023). Our findings indicate that S2 with PBC delivered more realistic results, demonstrating a mix of pixels within the same field, compared to S1, which presented more homogeneous patches. This is mainly because multispectral signals are more responsive to changes in plant health, usually observed as changes in crop canopy colour. Thus, S2 outperforms S1 in detecting these plant health variations caused by a hailstorm, giving it a significant advantage. We have also done qualitative analysis for each experiment and found that qualitative accuracies aligned with quantitative accuracies.

The class representing low damage always yields the highest accuracies across all experiments compared to the other two classes. This is mainly because most of the samples in both the training and testing datasets fall into this class. As detailed in section 4.1.1, out of the 40,157 in situ data points used for PBC, 25,051 were categorized as low damage, 10,744 as moderate damage, and 4,362 as high damage. An attempt was made to ensure a proportional representation of the samples by employing stratified random sampling during training and testing.

We selected the RF model for this study due to its proficient handling of class imbalances, a situation observed in our case. Additionally, it demonstrated an excellent capability to manage high-dimensional data. This ensemble method inherently resists overfitting, making it ideal for generalization. Furthermore, it exhibits robustness in the presence of noise and outliers. Another advantage of the RF model is its ability to rank variables, facilitating dimension reduction. In our study, we utilized multiple layers of S1 and S2 datasets. Initial experimentation revealed that employing the difference band, as substantiated by the MDA and MDG values, provided superior outcomes compared to including all layers. This suggests that focusing on changes over time could offer more insightful information.

Furthermore, upon implementing a dimension reduction strategy, we observed a notable enhancement in the accuracy of our model. This suggests that simplifying the data inputs by reducing their complexity aided the model in making more precise predictions. Because of the RF model's innate resistance to overfitting, the testing accuracies closely corresponded to the training accuracies, as outlined in the accuracy tables of the experiments.

While the RF model has been used for over two decades, its application for hail-induced crop damage detection, especially in combination with SAR and Optical data, has not been explored, adding a novelty to our research.

6.2.   Addressing and Mitigating Wicked Problem

The farmers are the most affected stakeholder in crop damage due to hailstorms. This research aimed to develop an ML model using multitemporal RS data to assess crop damage caused by hailstorms. This research was motivated by the need to improve upon the existing method, which heavily relies on field surveys. There is a transparency deficit in damage categorization assessments and a prolonged period from damage assessment to the actual disbursement of compensation, which takes 5-6 months, as elaborated in Section 1.8 (Problem Analysis). Several farmers, leveraging their connections, obtain undue advantages and compensation, despite suffering the same or even lesser degrees of loss, or sometimes no loss. The existence of such issues creates inequality to access to equal resources among farmers.

Farmers have raised several disagreements with the government, detailed in Section 1.8.2 (Wicked Problem Analysis). These include their disapproval of the rule that only farmers sustaining more than 33% damage qualify for compensation, dissatisfaction with the amount of compensation per hectare, and a demand for more damage category bins to ensure compensation proportional to the actual damage.

This model may not directly address the wickedness of policy-related issues such as determining per hectare financial allocation or the government guidelines for more than 33% damage for compensation eligibility. These issues are linked to policy decisions and require extensive stakeholder consultations and participation. However, the present RF model does hold significant potential in tackling two critical areas that concern farmers: the obscurity of damage category assessment and the lengthy duration of damage assessment. Furthermore, the model could serve as a valuable tool in identifying and effectively handling additional challenges. For example, it could provide a satisfactory resolution to farmers' demand for more detailed damage categories. Additionally, identifying crop damage from other sources proves challenging with the current method, as surveys are conducted post-hailstorm. In contrast, the RS method employs both pre- and post-hail imagery, enabling the successful elimination of certain damage sources by verifying the presence of a healthy crop in the field before the hailstorm, using the pre-hail images. The utility of RS method could enable efficient monitoring of potentially dishonest practices by either department officials or farmers.

This research offers a proof of concept, illustrating the feasibility of this kind of damage assessment using RS and ML approaches. This ML tool can be utilized to create a reference dataset, and by integrating a sample survey from 10-15% of land parcels, it can provide a more precise damage assessment process. Implementing this model could significantly cut down the assessment duration and foster a sense of fairness and equality of access to resources among farmers.

While it is easy to criticize government officials for perceived delays, transparency deficits, and potential corruption, it is essential to recognize their efforts in managing a complex scenario, often under immense pressure from farmers, the media, and even the state government, all the while juggling with limited financial resources. They must navigate this tricky terrain with limited funds, ensuring the budget aligns with the funds earmarked for disaster scenarios like hailstorms. The damage category dataset, supplied by the department and utilized for both training and testing during the model development, aligns with the results generated by the ML model, and it appears that they are managing it in a manner that results in systematic errors of damage categorization to align with financial constraints. Therefore, the absence of financial resources and the lack of spatial knowledge render the task difficult and time-consuming.

Considering time limitations, an interview was conducted with only one farmer from the region under study. Additional analysis was performed through a comprehensive literature review. Table 22 details the examination of suitable engagement strategies for different stakeholders. When implementing an RS-based method, a similar approach should be taken for an in-depth stakeholder study, taking into consideration their interests, impact, and importance.

6.3.   Limitations

This research serves as a proof of concept, demonstrating the potential of RS and ML methods for assessing crop damage, explicitly focusing on the Rajasthan province in India. We only had four damage categories available in our in-situ data, and for our specific region, only three damage categories were identified. As such, if the number of damage categories increased, the model would require optimization with new in situ data of the respective categories for optimal classification.

The training and testing data for the RF model were sourced from the BoR department. Farmers have disagreements with BoR on the assessment process. However, our results validate that the damage classes are not assigned randomly, implying an unclear field method used by the department. Despite the absence of the highest damage category in the department's data, satellite data hints at abrupt physiological changes indicative of severe damage. This suggests a proportional shifting of damage categories, confirmed by farmers who mentioned that BoR officials tend to change damage categorization by classifying it into less severe damage categories than what is truly observed (Interview, 2023).

Moreover, it is worth mentioning that crop damage can occur due to various other factors, including, but not limited to, suboptimal farming practices, poor soil quality, insufficient water supply, crop diseases, defective seeds, improper use of fertilizers, and unsuitable cultivation materials. The current model is not designed to distinguish these types of damages. However, regular crop monitoring throughout the plant cycle is possible with RS time series data that detect minor changes not readily apparent to the human eye. Spectral signatures used to identify changes in crop health, structure, and hydration levels, thereby enabling the early identification of problems across wide areas. Hence, the model would require further modification, training and optimization to monitoring other damages.

Crop damage due to hailstorms is influenced by numerous factors, such as crop type, the crop's phenological stage at the time of the hailstorm, soil type, cultivation region, hail size and duration, and wind speed. These factors can lead to varied manifestations of damage across different crops. The model developed in this study utilized in situ data from fields primarily composed of nearly mature mustard crops. Hence, the model would require further training and optimization for different scenarios.

7. CONCLUSION AND RECOMMENDATIONS

7.1. Conclusion

The primary objective of this study was to develop an ML model utilizing multitemporal RS data to assess crop damage caused by hailstorms. The model was trained and tested using in-situ data gathered from the Kathumar block in the Alwar district of Rajasthan Province, India. We examined the effectiveness of PBC and OBC RF models, employing open-source Sentinel datasets, including S1 and S2. Later, we incorporated the combined S1 and S2 data into the model.

In our findings, PBC utilizing combined data exhibited compelling results, accurately identifying damage classes with an F1 score exceeding 0.90 after applying majority classification. Comparatively, OBC using S1 data offered the highest accuracy when considering a single dataset for damage assessment, which was also reflected in our qualitative analysis. This research also has societal significance as it attempts to reduce the wickedness in the challenging issue of present field-based crop damage assessment and its deep impact on farmers, which can, unfortunately, lead to financial instability and in some extreme cases, suicide. The current assessment process often fosters a sense of inequality and inequity. Thus, the broader goal of this research is to provide a proof of concept for addressing such situations in a better way.

This developed model could be further refined and employed to enhance the crop damage evaluation process on a larger scale, such as at the district or state level. It could function as a standard reference or a directed geoprocessing tool, effectively reducing the need for fieldwork while offering a solid answer to the present opacity issues in damage assessment categorization. Furthermore, it could act as a conduit to tackle other relevant issues. For example, it could provide a satisfactory resolution to farmers preference for a more detailed damage assessment scale, with narrower intervals such as 1-10%, 10-20%, and so on, up to 90-100%. Additionally, the utility of RS time-series data could enable efficient monitoring of potentially dishonest practices by either department officials or farmers.

In conclusion, this study has addressed and provided answers to the research questions outlined in Section 1.5, as follows:

RQ 1: How can using RF classifier on S1 data improve the present field survey-based crop damage assessment approach?

The limitations of existing techniques, as outlined in Section 1.8.1, can be addressed by implementing the model developed in this study. This could limit field inspections to 10-15%, focusing on cross verifying the damage categories predicted by the model with ground truth, or focusing on regions where disagreements persist. This approach could expedite the assessment process and serve as a baseline reference for allocating damage categories. As the department is launching an app to facilitate direct damage reporting by farmers, mitigating transparency issues. Integrating the current model with this upcoming app could result in a more robust tool, thus substantially reducing the complexity of the problem at hand.

The application of the RF model on S1 in OBC results in the second highest accuracies, following the pixel-based majority classification. Moreover, when applied to PBC, it achieves the third-highest accuracy. In situations of hailstorms, cloud cover is a common occurrence, which can limit the use of S2, as observed in our case. However, S1 is not confined by weather or time, allowing it to penetrate cloud cover

effectively. Additionally, the sensitivity of microwave signals to changes in physical structure and geometry renders them particularly advantageous for damage detection. Polarimetric features like covariance matrix elements, stokes parameter and H/α decomposition provide information in better understanding backscattering mechanism to identify crop damage. Our findings revealed that S1 outperformed S2 in PBC and OBC approaches.

However, SAR data inherently encompasses a high level of 'speckle' noise, which can impact the quality and usability of the data. Moreover, the complexities involved in processing SAR data, and its limited capability to offer visual and colour information present challenges.

RQ 2: How can using RF classifier on S2 data improve the present field survey-based crop damage assessment approach?

The benefits of RS over field-based methods, particularly with the use of S2 data, lie in its sensitivity to changes in plant health, usually observed as physiological changes appearing in canopy colour. In the present research MCARI and REIP vegetation indices and S1 SWIR bands b11 and b12 was the most important variable as per MDA and MDG values. MCARI and REIP vegetation indices, due to their sensitivity to chlorophyll content and b11 and b12, due to their sensitivity to the moisture content in crops, was able to identify crop stress due to hail. The impact of a hailstorm on the ground is not homogeneous, resulting in patchy crop damage with variability within the same field.

Our findings indicate that with PBC, S2 delivered more realistic results. also, combining with S1 in PBC increased the accuracies. Open source S1 brings considerable advantages with a spatial resolution (10-60m), spectral capabilities (13 bands from visible to SWIR), five days temporal resolution and radiometric quality (12-bit), ease of interpretation and less complex processing.

However, its usage is bounded by certain limitations, including dependence on weather and daylight conditions. It also requires rectifying sun glint effects, brought about by atmospheric disruptions and the identification of clouds and their shadows.

RQ 3: Which classification approach, PBC or OBC, results in a more accurate assessment of hailstorm-related crop damage when utilizing S1 and S2 data separately?

Our experiments of PBC and OBC models applied to individual datasets showed that the OBC approach with S1 data yields the highest F1, precision, and recall scores for all three damage classes. Specifically, we noted F1 scores of 0.9564, 0.8990, and 0.8543 for the low, moderate, and high damage classes. When considering the datasets individually, the OBC performed better with the S1 data, while S2 data yielded superior results with PBC.

RQ 4: How can using RF classifier on combined S1 and S2 data improve the present field survey-based crop damage assessment approach?

The advantages of RS over field methods and capabilities and limitations of individual datasets are described in answering RQ1 and RQ2. We employed both PBC and OBC approach on the combined dataset. Upon evaluation, we discovered that the PBC with the combined dataset outperformed all other experiments when the majority classification was applied. The corresponding F1 scores for the low, moderate, and high damage classes were 0.97, 0.923, and 0.938, respectively.

The combination of both optical and SAR data can counterbalance their limitations, providing a more comprehensive understanding of the damage correlation with different parameters. In our case PBC for combining S1 and S2 data enhanced the classification accuracies compared to separate S1 and S2 datasets However, in OBC, accuracies decreased after combining S1 and S2. Hence, it is not conclusive that the use of combined data always enhances classification. The decision to use separate or combined datasets depends largely on the nature of the application and the methods used.

RQ 5: Which classification approach, PBC or OBC, results in a more accurate assessment of hailstorm-related crop damage when utilizing a combination of S1 and S2 data?

In our current study, we employed combined data with PBC and OBC. Results from PBC showed that using combined datasets enhanced the classification accuracies. However, in OBC, accuracies decreased. The combined S1 and S2 experiment with PCB after applying majority classification provided F1 scores for the low, moderate, and high damage classes of 0.97, 0.923, and 0.938, respectively, which was the highest in all experiments.

RQ 6: How would the proposed remote sensing-based crop damage assessment method impact different stakeholders?

This current research focuses on farmers, who are expected to gain empowerment through this remote sensing (RS) based damage assessment model. It will enhance transparency in the damage categorization process, fostering an environment of equality when dealing with shared resources and symbolising a transition towards a fair distribution of resources. Agricultural labourers and local communities could experience enhanced financial security and employment opportunities due to more accurate and timely compensation for farmers. The state government, including the DMR and BoR, could increase public trust through a transparent, efficient process that also allows for cost savings. However, expected increased compensation may necessitate policy discussions and stakeholder engagement. Crop insurance providers could gain from more precise damage evaluations as they facilitate determining suitable payouts. Given the mutual benefits, this could boost their standing within the farming community and promote insurance practices among farmers.

7.2.   Recommendations

This research utilized dual-polarization data from S1 for the analysis. However, Quad-pol data, with its extensive information derived from HH, VV, HV, and VH polarization combinations, can offer a more detailed understanding of backscattering mechanisms of crop damage. Consequently, the inclusion of Quad-pol data in subsequent studies has the potential to generate diverse outcomes.

Further research into testing additional features generated from other polarimetric decompositions, notably polarimetric-interferometric decomposition, is still useful. The interferometric coherence image could reveal important information about crop damage. Monitoring the complete crop growth cycle might provide additional benefits, enabling us to distinguish between hail-induced crop damages and those caused by other factors. It also offers an opportunity to investigate the impact of hailstorms at different crop growth stages. The deployment of deep learning methods could further refine our understanding of the backscattering mechanism related to hail-induced crop damage. Classifying crops prior to hail damage analysis might provide valuable insights into how hail affects different crops based on their structural and biophysical properties.

In light of these findings, we recommend the government adopt RS based approach for damage assessment to expedite the evaluation process and enhance the accuracy of assessments. This strategy can foster trust and transparency and empower farmers, who are often marginalized despite their significant contributions. Even within the existing four damage categories and the current rules of non-compensation for damages less than 33%, the proposed method could be economically efficient. This approach could potentially reduce field checks to 10-15% of their current volume, leading to considerable cost savings.

REFERENCES

Ackermann, N. (2015). *Growing Stock Volume Estimation in Temperate Forested Areas Using a Fusion Approach with SAR Satellites Imagery*. https://doi.org/10.1007/978-3-319-13138-2

Al-Gaadi, K. A., Hassaballa, A. A., Tola, E., Kayad, A. G., Madugundu, R., Alblewi, B., & Assiri, F.(2016). Prediction of Potato Crop Yield Using Precision Agriculture Techniques. *PLOS ONE*, *11*(9), e0162219. https://doi.org/10.1371/JOURNAL.PONE.0162219

Angearu, C. V., Ontel, I., Irimescu, A., Sorin, B., & Dodd, E. (2022). Remote sensing methods for detecting and mapping hailstorm damage: a case study from the 20 July 2020 hailstorm, Baragan Plain, Romania. *Natural Hazards*, *114*(2), 2013–2040. https://doi.org/10.1007/S11069-022-05457-X/FIGURES/16

Annual Progress Report DMR. (2023). https://dmrelief.rajasthan.gov.in/index.php/irrigation-calender/annual-progress-report

Arciniegas, G. A., Bijker, W., Kerle, N., & Tolpekin, V. A. (2007). Coherence-and Amplitude-Based Analysis of Seismogenic Damage in Bam, Iran, Using ENVISAT ASAR Data. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, *45*(6). https://doi.org/10.1109/TGRS.2007.883149

Bal, S. K., Saha, S., Fand, B. B., Singh, N. P., Rane, J., & Minhas, P. S. (2014). *Hailstorms: Causes, Damage and Post-hail Management in Agriculture*. https://www.researchgate.net/publication/264667104_Hailstorms_Causes_Damage_and_Post-hail_Management_in_Agriculture

Barakat, R. (1977). Degree of polarization and the principal idempotents of the coherency matrix. *Optics Communications*, *23*(2), 147–150. https://doi.org/10.1016/0030-4018(77)90292-9

Barati, S., Rayegani, B., Saati, M., Sharifi, A., & Nasri, M. (2011). Comparison the accuracies of different spectral indices for estimation of vegetation cover fraction in sparse vegetated areas. *The Egyptian Journal of Remote Sensing and Space Science*, *14*(1), 49–56. https://doi.org/10.1016/J.EJRS.2011.06.001

BBC Future. (2022). https://www.bbc.com/future/article/20220314-how-big-can-hailstones-grow

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011

Bell, J. R., & Molthan, A. L. (2016). Article Evaluation of Approaches to Identifying Hail Damage to Crop Vegetation Using Satellite Imagery. *J. Operational Meteor*, *4*(11), 142–159. https://doi.org/10.15191/nwajom.2016.0411

Berger, K., Machwitz, M., Kycko, M., Kefauver, S. C., Van Wittenberghe, S., Gerhards, M., Verrelst, J., Atzberger, C., van der Tol, C., Damm, A., Rascher, U., Herrmann, I., Paz, V. S., Fahrner, S., Pieruschka, R., Prikaziuk, E., Buchaillot, Ma. L., Halabuk, A., Celesti, M., … Schlerf, M. (2022). Multi-sensor spectral synergies for crop stress detection and monitoring in the optical domain: A review. *Remote Sensing of Environment*, *280*, 113198. https://doi.org/10.1016/J.RSE.2022.113198

BoR. (2023). https://landrevenue.rajasthan.gov.in/content/landrevenuenew/en/board-of-revenue-for-raj-dep.html.html#

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1023/A:1018054314350/METRICS

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Bush, T. F., & Ulaby, F. T. (1978). An evaluation of radar as a crop classifier. *Remote Sensing of Environment*, *7*(1), 15–36. https://doi.org/10.1016/0034-4257(78)90004-4

Camps Valls, G. (2008). New machine-learning paradigm provides advantages for remote sensing. *SPIE Newsroom*. https://doi.org/10.1117/2.1200806.1100

Chandler, O., Young, F., & Apan, A. (2004). *ASSESSMENT OF HAIL DAMAGE TO CROPS USING SATELLITE IMAGERY AND HAND HELD HYPERSPECTRAL DATA*

Chattopadhyay, N., Devi, S. S., John, G., & Choudhari, V. R. (2017). *Occurrence of hail storms and strategies to minimize its effect on crops. 68*(1), 75–92.

Cloude, S. R., & Pottier, E. (1997). An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, *35*(1), 68–78. https://doi.org/10.1109/36.551935

CMDA. (2022). https://www.nrcan.gc.ca/earth-sciences/geomatics/satellite-imagery-air-photos/satellite-imagery-products/educational-resources/14652

Copernicus Open Access Hub. (2022). https://scihub.copernicus.eu/dhus/#/home

Corresponding, D. L., Mausel, P., Brondízio, E., Moran, E., Lu, D., Mausel{, P., Brondi´ziobrondi´brondi´zio, E., & Moran{, E. (2010). Change detection techniques. *International Journal of Remote Sensing*, *25*, 2365–2401. https://doi.org/10.1080/0143116031000139863

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment*, *74*(2), 229–239. https://doi.org/10.1016/S0034-4257(00)00113-9

Delegido, J., Verrelst, J., Alonso, L., & Moreno, J. (2011). Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors 2011, Vol. 11, Pages 7063-7081*, *11*(7), 7063–7081. https://doi.org/10.3390/S110707063

Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 1–13. https://doi.org/10.1186/1471-2105-7-3/FIGURES/1

DMRD. (2023). http://www.dmrelief.rajasthan.gov.in

DNA. (2019). https://www.dnaindia.com/jaipur/report-rajasthan-hailstorm-rain-damage-crop-in-hadouti-2720228

Emery, W., & Camps, A. (2017). Radar. *Introduction to Satellite Remote Sensing*, 291–453. https://doi.org/10.1016/B978-0-12-809254-5.00005-1

ESA - Sentinel-1. (2022). https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1

ESA Sentinel Mission. (2023). https://sentinel.esa.int/web/sentinel/missions

Espinoza, O. (2008). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Http://Dx.Doi.Org/10.1080/00131880701717198*, *49*(4), 343–363. https://doi.org/10.1080/00131880701717198

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64–87. https://doi.org/10.1016/J.RSE.2016.08.013

Fernandes, G. W., Oki, Y., Sanchez-Azofeifa, A., Faccion, G., & Amaro-Arruda, H. C. (2011a). Hail impact on leaves and endophytes of the endemic threatened Coccoloba cereifera (Polygonaceae). *Plant Ecology 2011 212:10*, *212*(10), 1687–1697. https://doi.org/10.1007/S11258-011-9941-Z

Fernandes, G. W., Oki, Y., Sanchez-Azofeifa, A., Faccion, G., & Amaro-Arruda, H. C. (2011b). Hail impact on leaves and endophytes of the endemic threatened Coccoloba cereifera (Polygonaceae). *Plant Ecology*, *212*(10), 1687–1697. https://doi.org/10.1007/S11258-011-9941-Z/TABLES/2

Gerhards, M., Schlerf, M., Mallick, K., & Udelhoven, T. (2019). Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review. *Remote Sensing 2019, Vol. 11, Page 1240*, *11*(10), 1240. https://doi.org/10.3390/RS11101240

Gitelson, A. A., & Merzlyak, M. N. (2010). Remote estimation of chlorophyll content in higher plant leaves. *Http://Dx.Doi.Org/10.1080/014311697217558*, *18*(12), 2691–2697. https://doi.org/10.1080/014311697217558

Ha, T., Shen, Y., Duddu, H., Johnson, E., & Shirtliffe, S. J. (2022). Quantifying Hail Damage in Crops Using Sentinel-2 Imagery. *Remote Sensing 2022, Vol. 14, Page 951*, *14*(4), 951. https://doi.org/10.3390/RS14040951

Hatfield, J. L., & Prueger, J. H. (2010). Value of Using Different Vegetative Indices to Quantify Agricultural Crop Characteristics at Different Growth Stages under Varying Management Practices. *Remote Sensing 2010, Vol. 2, Pages 562-578*, *2*(2), 562–578. https://doi.org/10.3390/RS2020562

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *80*, 91–106. https://doi.org/10.1016/J.ISPRSJPRS.2013.03.006

Immitzer, M., Vuolo, F., Atzberger, C., Sarathi Roy, P., & Thenkabail, P. S. (2016). First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing 2016, Vol. 8, Page 166*, *8*(3), 166. https://doi.org/10.3390/RS8030166

India Today. (2020). https://www.indiatoday.in/india/story/heavy-rainfall-hailstorm-lash-rajasthan-crop-damage-feared-1652889-2020-03-05

Izquierdo-Verdiguier, E., Zurita-Milla, R., & De By, R. A. (2017). On the use of guided regularized random forests to identify crops in smallholder farm fields. *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images, MultiTemp 2017*. https://doi.org/10.1109/MULTI-TEMP.2017.8035248

Jedlovec, G. J., Nair, U., & Haines, S. L. (2006). *Detection of Storm Damage Tracks with EOS Data.*

Karam, M. A., Fung, A. K., Lang, R. H., & Chauhan, N. S. (1992). Microwave Scattering Model for Layered Vegetation. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(4), 767–784. https://doi.org/10.1109/36.158872

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Arntz, L. R., Marinetti, C., van der Valk, D., Iannini, L., Molijn, R., Westerdijk, K., & van der Sande, C. (2019). Crop monitoring using Sentinel-1 data: A case study from The Netherlands. *Remote Sensing*, *11*(16). https://doi.org/10.3390/RS11161887

Krishi Vigyan Kendra, Alwar 1. (n.d.). Retrieved June 3, 2023, from https://alwar1.kvk2.in/district-profile.html

Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering*, 83–106. https://doi.org/10.1016/B978-0-12-818366-3.00005-8

Lee, J. Sen, & Pottier, E. (2017). Polarimetric radar imaging: From basics to applications. *Polarimetric Radar Imaging: From Basics to Applications*, 1–398. https://doi.org/10.1201/9781420054989/POLARIMETRIC-RADAR-IMAGING-JONG-SEN-LEE-ERIC-POTTIER

Leite, G. B., Petri, J. L., & Mondardo, M. (2002). Effect of anti-hail screen on some characteristics of apple fruits. Brazilian Journal of Fruit Growing *24*(3), 714–716. https://doi.org/10.1590/S0100-29452002000300037

Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest. 2*(3). http://www.stat.berkeley.edu/

Löw, F., Michel, U., Dech, S., & Conrad, C. (2013). Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines. *ISPRS Journal of Photogrammetry and Remote Sensing*, *85*, 102–119. https://doi.org/10.1016/J.ISPRSJPRS.2013.08.007

Ma, J., Ding, Y., Cheng, J. C. P., Tan, Y., Gan, V. J. L., & Zhang, J. (2019). Analyzing the Leading Causes of Traffic Fatalities Using XGBoost and Grid-Based Analysis: A City Management Perspective. *IEEE Access*, *7*, 148059–148072. https://doi.org/10.1109/ACCESS.2019.2946401

Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J. M., McNairn, H., & Rao, Y. S. (2020). Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sensing of Environment*, *247*. https://doi.org/10.1016/J.RSE.2020.111954

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

McNairn, H., & Shang, J. (2016). A review of multitemporal synthetic aperture radar (SAR) for crop monitoring. *Remote Sensing and Digital Image Processing*, *20*, 317–340. https://doi.org/10.1007/978-3-319-47037-5_15/FIGURES/5

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). *Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening.* https://doi.org/10.1034/j.1399-3054.1999.106119.x

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*(4), 358–371. https://doi.org/10.1016/J.BIOSYSTEMSENG.2012.08.009

Mutanga, O., Adam, E., & Cho, M. A. (2012). High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation*, *18*(1), 399–406. https://doi.org/10.1016/J.JAG.2012.03.012

Mutanga, O., & Skidmore, A. K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*(1), 34–42. https://doi.org/10.1016/J.ISPRSJPRS.2007.02.001

Naji, T. A. H. (2018). Study of vegetation cover distribution using DVI, PVI, WDVI indices with 2D-space plot. *Journal of Physics: Conference Series*, *1003*(1), 012083. https://doi.org/10.1088/1742-6596/1003/1/012083

NASA EO, N. (EO). (2000). *Measuring Vegetation (NDVI & EVI).*

NASA (ESDS), N. (ESDS). (2022). *What is Remote Sensing? | Earthdata.* https://www.earthdata.nasa.gov/learn/backgrounders/remote-sensing

NASA, N. (2023). *Global Temperature | Vital Signs – Climate Change: Vital Signs of the Planet.* https://climate.nasa.gov/vital-signs/global-temperature/

News NCR. (2022). https://www.newsncr.com/national/rajasthan-chief-minister-gehlot-ordered-special-girdawari-farmers-affected-by-hailstorm-will-get-compensation/

NSSL. (2022). https://www.nssl.noaa.gov/research/hail/

Olander, S., & Landin, A. (2005). Evaluation of stakeholder influence in the implementation of construction projects. *International Journal of Project Management*, *23*(4), 321–328. https://doi.org/10.1016/J.IJPROMAN.2005.02.002

Padalia, H., Sinha, S. K., Bhave, V., Trivedi, N. K., & Senthil Kumar, A. (2020). Estimating canopy LAI and chlorophyll of tropical forest plantation (North India) using Sentinel-2 data. *Advances in Space Research*, *65*(1), 458–469. https://doi.org/10.1016/J.ASR.2019.09.023

Pal, M. (2007). Random forest classifier for remote sensing classification. *Http://Dx.Doi.Org/10.1080/01431160412331269698*, *26*(1), 217–222. https://doi.org/10.1080/01431160412331269698

Park, M., Jung, D., Lee, S., & Park, S. (2020). Heatwave damage prediction using random forest model in Korea. *Applied Sciences (Switzerland)*, *10*(22), 1–12. https://doi.org/10.3390/APP10228237

Patrika News. (2023). https://www.patrika.com/udaipur-news/rajasthan-weather-update-hailstorm-with-rain-in-udaipur-8010857/

Periasamy, S. (2018). Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. *Remote Sensing of Environment*, *217*, 537–549. https://doi.org/10.1016/J.RSE.2018.09.003

Pinty, B., & Verstraete, M. M. (1992). GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetatio*, *101*(1), 15–20. https://doi.org/10.1007/BF00031911/METRICS

PMFBY Crop Insurance. (2023). https://pmfby.gov.in/

Portner, H. O., Roberts, D. C., Poloczanska, E. S., Mintenbeck, K., Tignor, M., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., & Okem, A. (2022). *IPCC AR6: Summary for Policymakers*. https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_SummaryForPolicymakers.pdf

Prabhakar, M., Gopinath, K. A., Reddy, A. G. K., Thirupathi, M., & Rao, C. S. (2019). Mapping hailstorm damaged crop area using multispectral satellite data. *Egyptian Journal of Remote Sensing and Space Science*, *22*(1), 73–79. https://doi.org/10.1016/J.EJRS.2018.09.001

Rain & hailstorm hit crops in desert districts in Rajasthan | Jaipur News - Times of India. (n.d.). Retrieved June 3, 2023, from https://timesofindia.indiatimes.com/city/jaipur/rain-hailstorm-hit-crops-in-desert-dists/articleshow/98504095.cms

Rao, V. U. M., Sikka, A. K., Rao, A. V. M. S., & Rao, B. B. (2014). *Hailstorm Technical bulletin 2014*.

Ray, S. S., Singh, S. K., Neetu, & Mamatha, S. (2016). Establishing an operational system for assessment and forecasting the impact of extreme weather events on crop production. *MAUSAM*, *67*(1), 289–296. https://doi.org/10.54302/MAUSAM.V67I1.1230

Rubel, O., Lukin, V., Rubel, A., & Egiazarian, K. (2021). Selection of Lee Filter Window Size Based on Despeckling Efficiency Prediction for Sentinel SAR Images. *Remote Sensing 2021, Vol. 13, Page 1887*, *13*(10), 1887. https://doi.org/10.3390/RS13101887

Sanchez, A. H., Picoli, M. C. A., Camara, G., Andrade, P. R., Chaves, M. E. D., Lechler, S., Soares, A. R., Marujo, R. F. B., Simões, R. E. O., Ferreira, K. R., & Queiroz, G. R. (2020). Comparison of cloud cover detection algorithms on sentinel-2 images of the Amazon tropical forest. *Remote Sensing*, *12*(8). https://doi.org/10.3390/RS12081284

Saxena, R. (2019). *Rajasthan: Hailstorm, rain almost 100 % damage crops in Hadouti. DNA news*. https://www.dnaindia.com/jaipur/report-rajasthan-hailstorm-rain-damage-crop-in-hadouti-2720228

Schubert, T. (1991). *HAIL DAMAGE TO PLANTS*.

Shrestha, R. M., & Rahman, Md. S. (2021). *Flood Monitoring and Crop Damage Assessment*. 321–349. https://doi.org/10.1007/978-3-030-66387-2_16

Singh, S. K., Saxena, R., Porwal, A., Neetu, & Ray, S. S. (2017). Assessment of hailstorm damage in wheat crop using remote sensing. *Current Science*, *112*(10), 2095–2100. https://doi.org/10.18520/CS/V112/I10/2095-2100

SNAP S2. (2023). https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-2

Sosa, L., Justel, A., & Molina, Í. (2021). Detection of Crop Hail Damage with a Machine Learning Algorithm Using Time Series of Remote Sensing Data. *Agronomy 2021, Vol. 11, Page 2078*, *11*(10), 2078. https://doi.org/10.3390/AGRONOMY11102078

State Portal Rajasthan. (2023). https://rajasthan.gov.in/

Szabó, A., Tamás, J., & Nagy, A. (2021). The influence of hail net on the water balance and leaf pigment content of apple orchards. *Scientia Horticulturae*, *283*, 110112. https://doi.org/10.1016/J.SCIENTA.2021.110112

Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sensing of Environment*, *71*(2), 158–182. https://doi.org/10.1016/S0034-4257(99)00067-X

Times of India. (2020). https://timesofindia.indiatimes.com/city/jaipur/farmers-upset-with-33-norm-for-compensation/articleshow/73341822.cms

Times of India. (2021). https://timesofindia.indiatimes.com/city/jaipur/rain-hailstorm-hit-farmers-hard/articleshow/80123248.cms

Times of India. (2022). https://timesofindia.indiatimes.com/city/jaipur/storm-damages-crops-power-lines-in-rajasthan/articleshow/88784772.cms

Times of India. (2023). https://timesofindia.indiatimes.com/city/jaipur/farmer-commits-suicide-in-bundi-due-to-crop-damage/articleshow/98791007.cms

Völker, M., & Doneys, P. (2020). Empowerment as one sees it: assessment of empowerment by women participants of development projects. *Https://Doi.Org/10.1080/09614524.2020.1828284*, *31*(1), 125–138. https://doi.org/10.1080/09614524.2020.1828284

Wang, E., Little, B. B., Williams, J. R., & Yu, Y. (2012). Simulation of hail effects on crop yield losses for corn-belt states in USA. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, *28*(21), 177–185. https://doi.org/10.3969/j.issn.1002-6819.2012.21.025

Wang, P., Ma, Q., Wang, J., Hong, W., Li, Y., & Chen, Z. (2013). An improved SAR radiometric terrain correction method and its application in polarimetric SAR terrain effect reduction. *Progress In Electromagnetics Research B*, *54*, 107–128. https://doi.org/10.2528/PIERB13052021

Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, *236*. https://doi.org/10.1016/J.RSE.2019.111402

Wu, W., Zhou, Z., Li, T., & Long, S. (2017). A Study of Sentinel-1 TOPS Mode Co-registration. *Cehui Xuebao/Acta Geodaetica et Cartographica Sinica*, *46*(9), 1156–1164. https://doi.org/10.11947/J.AGCS.2017.20160352

Zhang, H., Kang, J., Xu, X., & Zhang, L. (2020). Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. *Computers and Electronics in Agriculture*, *176*, 105618. https://doi.org/10.1016/J.COMPAG.2020.105618

Zhou, J., Pavek, M. J., Shelton, S. C., Holden, Z. J., & Sankaran, S. (2016). Aerial multispectral imaging for crop hail damage assessment in potato. Computers and Electronics in Agriculture, 127, 406–412. https://doi.org/10.1016/J.COMPAG.2016.06.019

APPENDICES

Transcribed Interview: Stakeholder's Perspective on Hailstorm Damage and Compensation.

Date: 16 June 2023 Time: 07:37 UTC
Interviewer: Rishi Pareek (M.SE.)
Interviewee: Farmar
Resident: District Alwar, Province Rajasthan, Country: India
Farming experience: more than 10 years

Experience with Hailstorms

Q1): How frequently do you encounter hailstorms in your vicinity?
Interviewee: Almost every year, we encounter hailstorms in our region. Over the past two years, they have been occurring more than once a year.

Q2): How often do you face crop damage due to hailstorms annually? In which season do hailstorms usually occur?
Interviewee: Given the yearly occurrence of hailstorms, our crops are affected almost every year. The intensity varies; sometimes, it is low, and other times it is high. Hailstorms typically occur in the Rabi season, most often in February and March, but occasionally in January and the end of December as well.

Q3): Could you share the types of crops you cultivate and which ones are the most vulnerable to hailstorms?
Interviewee: In the Rabi season, I grow wheat and mustard, and in the Kharif season, I generally cultivate millet and maize. The choice of the crop also depends on the expected market prices and experience with specific crops.

Q4): Can you elaborate on the impact of hailstorms on various crops in your region and the stages of crop growth at which the damages are most severe? Whether it matters if the hailstorm affects a younger or an older crop?
Interviewee: Hailstorms typically affect mature crops due to being harvested in about 15-20 days or those in mid-growth. If a storm occurs in the early growth stage, the crops may recover, but mature crops have a slim chance of recovery. This also depends on the severity and duration of the hailstorm. In the case of wheat, the grain size reduces and turns black. For mustard, the pods break, and the seeds become black.

Q5): Could you narrate your personal experiences dealing with hailstorm-inflicted damage on your crops?
Interviewee: In 2022, my crops suffered severe damage, for which I received compensation this year. It is a distressing situation. We invest financial resources, labour, and time for 4-5 months, and just when we expect returns, everything collapses. Sometimes, clearing the damaged crop requires additional manpower and resources, which is painful.

Q6): How do you manage crops damaged by hailstorms? Are you able to recover any yield from them, or is it a total loss?
Interviewee: Events like this are beyond our control; we are at the mercy of the weather. Sometimes we manage to recover 30-60 percent of the crops, but the quality significantly reduces.

Q7): If the damage is minimal, what is the usual yield loss you face?

Interviewee: Even with minimal damage, the crop quality reduces, and we don't receive prices as high as for fresh crops.

Q8): In your estimate, what is the financial impact of hailstorms on your income?

Interviewee: Two events - heavy precipitation and hail - are the most disastrous. We rely on credit from Kisan credit cards or bank loans for our farming activities. If we lose our crop, we sometimes feel it's better to engage in other work than farming due to the reduced uncertainty. We bear substantial financial losses and are burdened by loans.

Q9): Do hailstorms affect the quality and yield of crops, subsequently impacting the market price?

Interviewee: Yes, as mentioned before, the quality of crops significantly declines, impacting the market prices by 30-50 percent.

Existing Damage Assessment Process

Q10): Can you explain the current damage assessment process led by the department? Is it a comprehensive survey of each field, a partial survey, or is it based on past damage records or their knowledge?

Interviewee: The process depends on the patwari (revenue officer), who generally visits the village, gathers an overview from farmers and residents about the losses, then checks land parcels and marks the homogeneous patches on the revenue map. Sometimes they do this in the office, without visiting the sites, by analysing maps and using past experience of the area. Since they know every part of the village well, they can anticipate the topography, low-lying areas, and history of damages.

Q11): In the event of area-wide damage assessment, if only certain sections of crops are damaged, how does the department carry out the assessment?

Interviewee: We are not certain about the criteria they follow in such cases.

Q12): Is compensation uniform for fully-grown, partially-grown, and less-grown crops?

Interviewee: Yes, it is the same, but they generally consider fully grown crops.

Q13): If crops are damaged before a hailstorm due to other factors, does the department take this into account?

Interviewee: Records of such losses are not typically kept. On occasion, farmers might attribute crop damage to a hailstorm, and those with considerable influence and connections can secure compensation. For instance, since mustard is usually harvested earlier than wheat (about 15-20 days), if a hailstorm takes place during that period, farmers might report losses. They assert that they've discarded the affected sections of the crop and, thus, attempt to receive compensation.

Q14): Is the process subjective, or are there specific field measurements taken?

Interviewee: The process is purely subjective. They do it by their own we don't know how they do it.

Q15): Are you aware of the damage percentage bins (0-33%, 33-50%, 50-75%, 75-100%) set by the government for compensation? How satisfied are you with the current method of damage assessment? If unsatisfied, could you share your reasons?

Interviewee: I am not satisfied with the rules, particularly the one that excludes damages less than 33% from compensation. The categories for damage assessment are unclear, and officials often attempt to fit us

into the lower damage category. I believe compensation should be given in proportion to the loss. Moreover, the time taken for assessment and compensation is too long and should be reduced.

Q16): Would you advocate for further division of the damage bins (1-10%, 10-20%, ---- to 90-100%)? Why or why not? Should the compensation be proportionate to each damage category?
Interviewee: Yes, I would prefer the bins to be divided further, allowing for proportional compensation, which would be fairer.

Q17): What are your thoughts on the government's policy of not providing compensation for crop damage below 33%?
Interviewee: We disagree with this policy. It contradicts itself by acknowledging a loss but not compensating for it. I believe this policy aims to reduce the compensation amount and attempts to fit most farmers into this category.

Q18): Do you concur with the compensation rate per hectare set by the government? If not, why?
Interviewee: The compensation we receive, ranging from 6800/- Indian National Rupees (INR) (75 euro) to 18000 INR (200 euro) per hectare based on the damage bin, is substantially less than the expenses we bear. However, we appreciate the social initiatives like the PM-Kisan Samman Nidhi Yojana, which provides annual support of 6000 INR (75 euro), and PM KUSUM – MNRE, which grants various subsidies for the utilization of renewable energy resources. These schemes have indeed been beneficial for the farming community.

Q19): What are your options if you disagree with the damage assessment and compensation awarded? Is legal recourse possible?
Interviewee: Legal recourse is theoretically possible but practically unfeasible due to the time and cost involved. We would rather focus on the next crop instead.

Suggestions for Improvement

Q20): In your opinion, how can the current damage assessment process be improved?
Interviewee: The compensation scheme should include damages less than 33%. The amount per hectare should be increased. Farmers obtaining undue benefits should be scrutinized and penalized for unfair practices. Similarly, unfair government officials should be removed from their positions. Also, farmers should receive compensation in proportion to their losses. The current bins are too small and should be enlarged like 1-10%, 10-20% ----- to 90-100%.

Q21): What are your thoughts on integrating the existing damage assessment application with a Remote Sensing (RS) based model?
Interviewee: We are not well-informed about such technology. However, if it is something we can use on mobile to check damages, that would be helpful.

Crop Insurance

Q22): Do you have an insurance policy for your crops? If yes, what is the monthly premium?
Interviewee: We have PM Fasal Bima Yojana, which costs around 1500 INR (16 euro) per hectare per year. Generally, it is 2% of the sum insured. It is very with the type of crop sometimes.

Q23): How do you perceive the Pradhan Mantri Fasal Bima Yojana? How does it differ from private insurance policies?
Interviewee: The scheme is beneficial because the premium is very low. However, premiums for private insurance companies not aligned with PMFBY are ten times higher.

Q24): Have you benefited from any crop insurance schemes?
Interviewee: Yes, we benefited from the PM Fasal Bima Yojana last year when our crops were damaged by a hailstorm. We received compensation this year.

Q25): From your perspective, what role do insurance companies play in crop damage compensation?
Interviewee: I am not familiar with the specific role insurance companies play. Our knowledge is primarily centred around the Pradhan Mantri Fasal Bima Yojana. Farmers, including myself, are generally not inclined towards private insurance schemes due to the high premium rates they charge.

Q26): Are you familiar with the new mobile application launched by the department for damage assessment where farmers can report crop losses? If yes, have you used it? Could you share your experience?
Interviewee: We came across information about a new app in a newspaper article in March 2023, right after a hailstorm. However, we have not had the opportunity to use it so far. I believe it hasn't been officially launched yet.

Q27): Would you consider using such an application in the future?
Interviewee: Definitely, yes.

Q28): Are you aware of the application of the Remote Sensing approach for crop damage assessment?
Interviewee: We aren't familiar with such technology. However, we're interested in any tool that we can use on our mobile phones to check for crop damage.

Q29): What are your views on employing satellite imagery for hail damage assessment?
Interviewee: If it is capable of showing crop damage, we would definitely like to use it on our mobile devices.

Additional Concerns

Q30): Do you rely on loans for your agricultural activities?
Interviewee: Yes, we always rely on loans for our agricultural activities due to our limited financial resources.

Q31): What are your thoughts on the incidence of farmer suicides?
Interviewee: Farmer suicides are tragic and often result from crop loss and overwhelming loan burdens. It is distressing to think that those who grow our food sometimes don't have enough for themselves. The immense pressure they face is unbelievable.

Q32): In case of crop damage on rented farmland, who is eligible for compensation - the landowner or the farmer?

Interviewee: The compensation is given to the landowner. As tenants, they pay rent to the landowner at the beginning of the lease. Agreements vary, sometimes they are official, and sometimes they are based on mutual discussion. There isn't any discussion on crop losses generally. However, some agricultural labourers have started making agreements stating that the compensation should be transferred to their accounts if the owner receives it.

Final Remarks

Q33): Would you like to share any other experiences or insights related to hailstorm damage and the damage assessment process? Are there any gaps in the system that you believe need addressing?

Interviewee: I have shared most of my experiences. However, I want to emphasize that farming is becoming an increasingly risky business due to unpredicted weather and sudden crop losses. We don't have any other option for livelihood because this is all we have learned from early life. In case of crop losses, the government should act seriously and provide compensation in a timely manner. Furthermore, loans on farmers should be waived.

Field Data.

This section encompasses the field data utilized for training and testing in-situ data development, with subsequent sections showcasing land revenue maps. For ethical reasons, data pertaining to the RoR and the Hailstorm report haven't been shared as they contain personal information about farmers.

Land Revenue Maps:

This section includes the original maps of land revenue or parcels from the villages of Daroda, Badangarhi, Nagla Khooba, Nagla Madhopur, and Garoo. These maps were utilized for the creation of training and testing datasets.
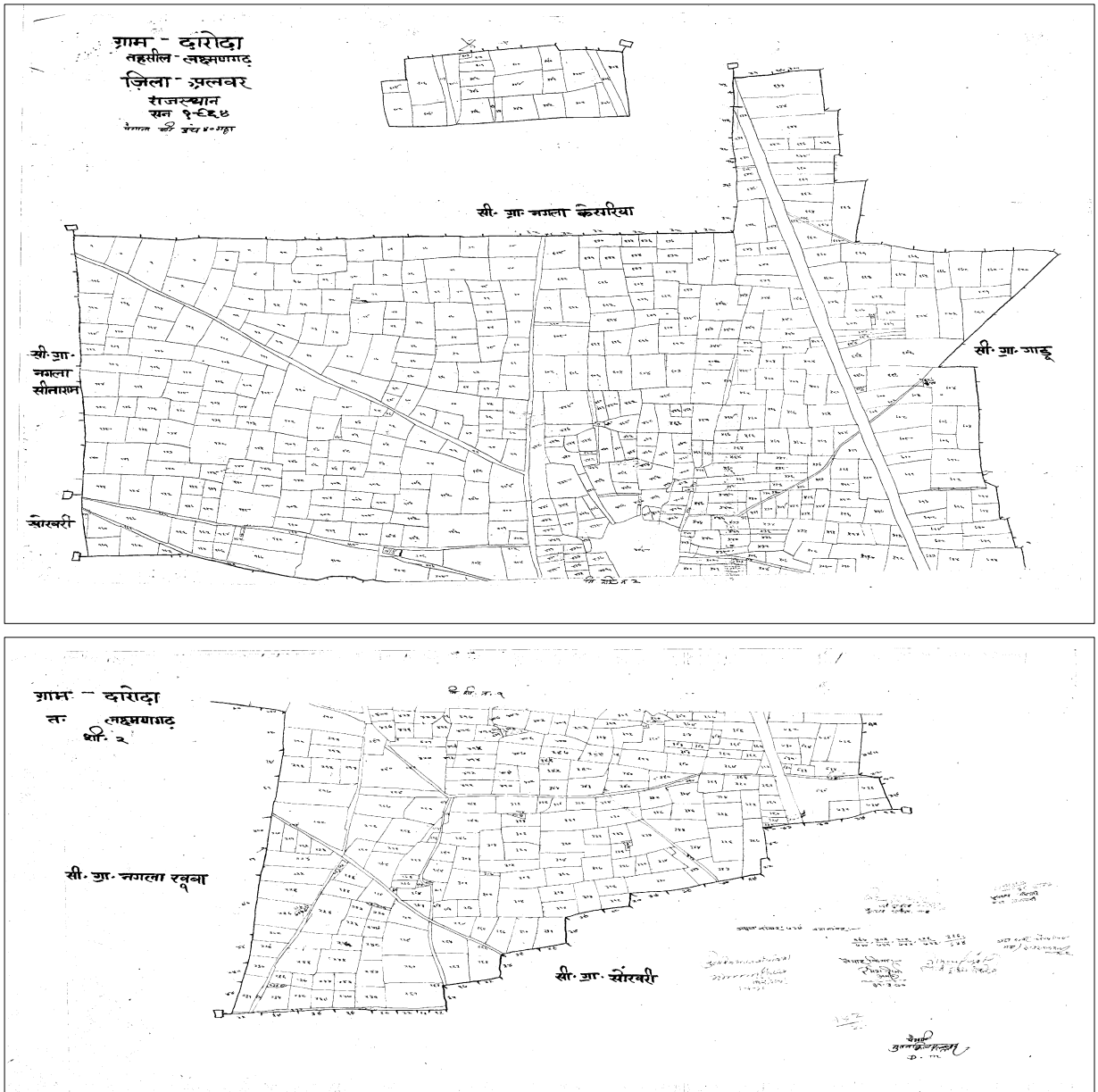
Figure 29 Parcel map village Daroda (2 sheets)
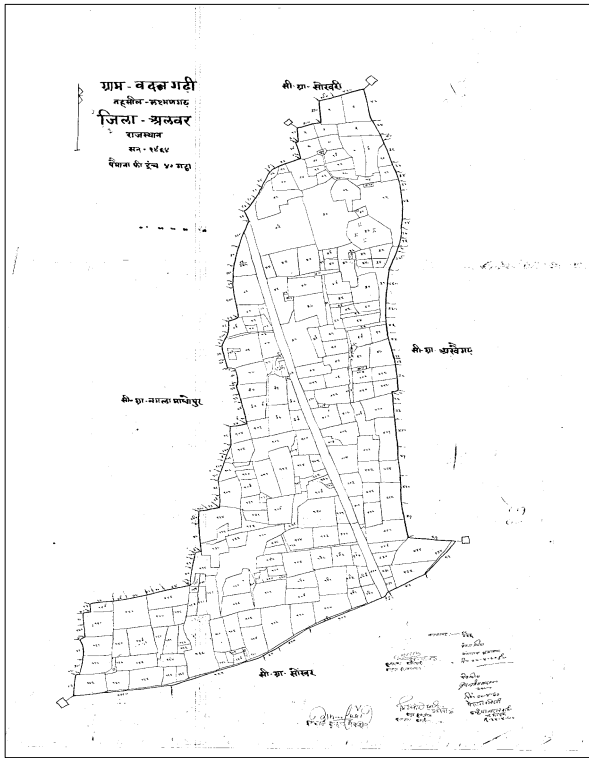
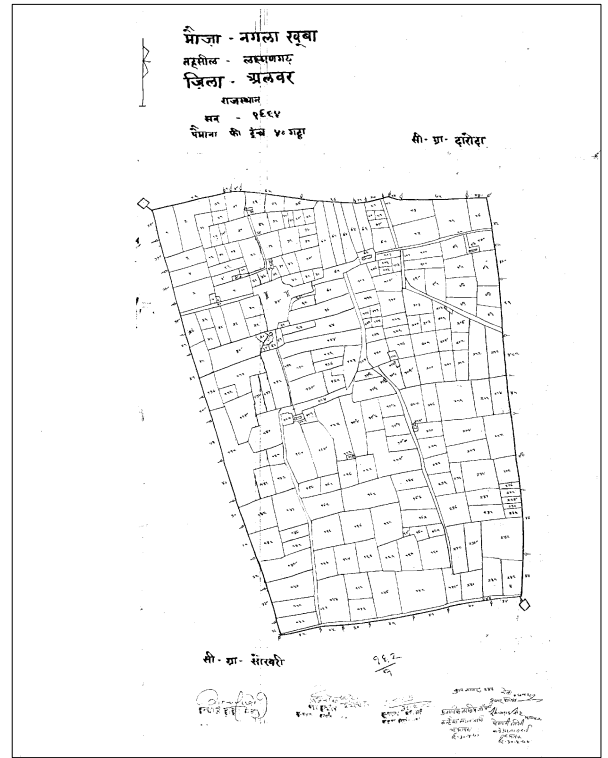Figure 31. Parcel map of village Badangarhi

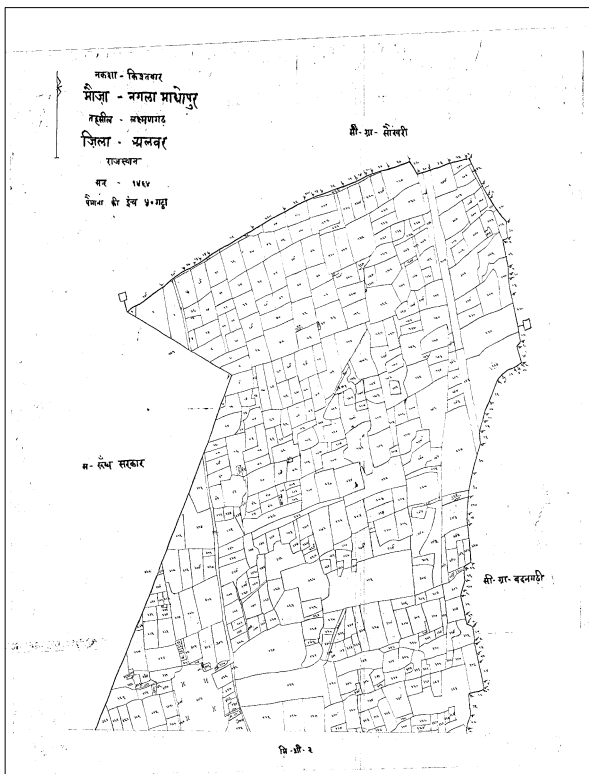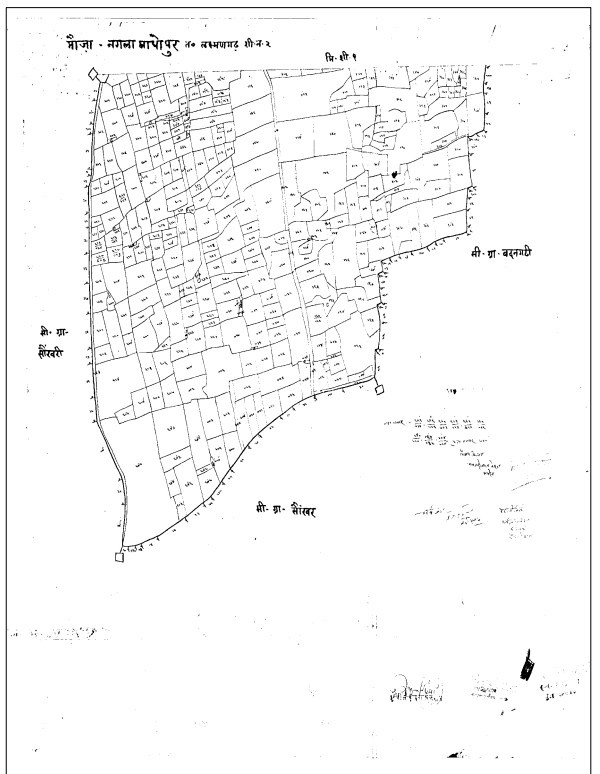Figure 30. Parcel map of village Nangla Khooba
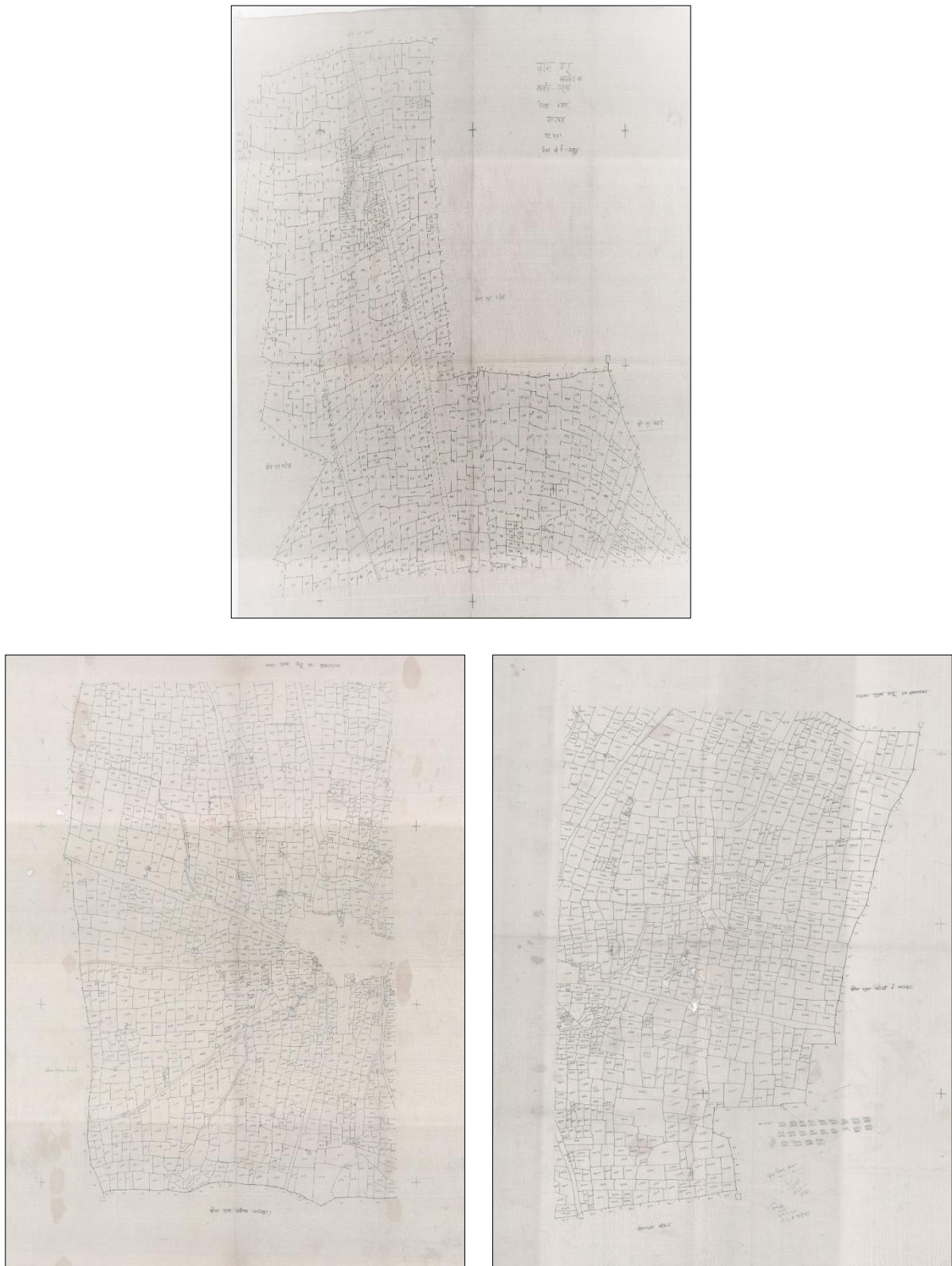

Figure 32 Parcel map of village Nangla Madhopur (2 sheets)

Figure 33. Parcel map of village Garoo (3 sheets)