DOWNSCALING, SOCIAL SENSING AND FUSION OF DIFFERENT TYPES

ClimEmpower Work Package 2, D2.3

Project ClimEmpower: User Driven Climate Applications Empowering Regional Resilience

Work Package 2, Deliverable D2.3

Date: 13.01.25

Please refer to this report as follows:

Chatzichristaki Ch., Karystinakis K., Moumtzidou A., Murano, I. (2024). Methodology of spatial data downscaling, social sensing and fusion of different data types. Deliverable D2.3 of the Horizon Europe project ClimEmpower.[1]

Project information	
Project name: **GA No.** **Start date:** **Duration:**	**ClimEmpower: User Driven Climate Applications Empowering Regional Resilience** **101112728** **01/09/2023** **36 months**
Coordinator:	Denis Havlik, Scientist AIT Austrian Institute of Technology Giefinggasse 4, 1210 Vienna, Austria
	Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.
Deliverable details	
Deliverable name	D2.3 Methodology of spatial data downscaling, social sensing and fusion of different types
Description:	Methodology of spatial data downscaling, social sensing and fusion of different data types. An overview of approaches to enhance Climate Resilience.
Version: **Due date:** **Submission:**	Final 23/12/2024 13/01/2025
Dissemination level:	PU Public
Lead:	Chatzichristaki Chrysa, CERTH
Author(s):	Chatzichristaki Ch., Karystinakis K., Moumtzidou A., (CERTH) Greece Murano I. (PLINIVS), Italy

[1] If you wish to reference this template, please reference the latest version thereof at zenodo

Revision history

Date	Version	Contributor/Reviewer	Description
29/10/2024	ToC	Chatzichristaki Ch., Karystinakis K., Moumtzidou A.,	Table of Contents shared with Consortium
07/11/2024	Preliminary draft	Chatzichristaki Ch., Karystinakis K., Moumtzidou A.,	1st version of deliverable with input from CERTH
29/11/2024	Preliminary draft	Ivan Murano	Input from PLINIUS
13/12/2024	Full Draft 1	Chatzichristaki Ch., Karystinakis K., Moumtzidou A.,	First complete version
18/12/2024	Internal review	Kaleta Patrick	Internal review from AIT
23/12/2024	Final report	Chatzichristaki Ch., Karystinakis K., Moumtzidou A.,	Revised version 1
13/01/2024	Final deliverable	Denis Havlik (AIT) → EC	Finalization and upload to Participants portal

Legal Disclaimer

All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user, therefore, uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and CINEA has no liability in respect of this document, which is merely representing the authors' view.

Final version of this document will be published under the terms of Creative Commons “CC BY” 4.0 license. That is, the Consortium explicitly permits redistribution, creating of derivatives, such as a translation, and even use of the publication for commercial activities, provided that appropriate credit is given to the author (BY) and that the user indicates whether the publication has been changed.

Please note that the scope of the license granted by Consortium is limited to own work and that the document can also contain materials attributed to a third party, such as tables, figures, or images. Users wishing to reuse such materials are responsible for determining whether permission is needed for that reuse and for obtaining permission from the copyright holder. The risk of claims resulting from infringement of any third-party owned component in the work rests solely with the user.

© 2024 by ClimEmpower Consortium

Table of Contents

List of Figures

List of Tables

List of Acronyms

AHP	Analytical Hierarchy Process
API	Application Programming Interface
BAI	Burned Area Index
BT	Brovey Transform
BSI	Bare Soil Index
CLC	Corine Land Cover
CNN	Convolutional Neural Network
CoP	Community of Practice
CS	Component substitution
CRT	Climate Risk Typology
CST	Climate Service Toolkit
CYGNSS	Cyclone Global Navigation Satellite System
DN	Digital Number
DDN	Delay-Doppler Maps
DEM	Digital Elevation Model
dNBR	Difference Normalized Burn Ratio
dNDVI	Difference Normalized Vegetation Index
DTM	Digital Terrain Model

ECMWF	European Centre for Medium-Range Weather Forecasts
EDO	European Drought observatory
EFAS	European Flood Awareness System
EFFIS	European Forest Fire Information System
EMS	Emergency Monitoring System
ESA	European Space Agency
ETC-CCA	European Topic Centre on Climate Change Impacts, Vulnerability, and Adaptation
EU-DEM	Digital Elevation Model over Europe
EUROSTAT	European Statistics
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation Anomaly
FFPI	Flash Flood Potential Index
FWI	Fire-weather Index
GDP	Gross Domestic Products
GEMI	Global Environment Monitoring Index
GFCS	Global Framework for Climate Services
GIS	Geographic Information Systems
GISCO	Geographic Information system of Commission
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HR	High Resolution
IPCC	Intergovernmental Panel on Climate Change
JRC	Joint Research Centre
JSON	JavaScript Object Notation
KBDI	Keetch-Byram Drought Index
LSTM	Long Short-Term Memory
LSWI	Land Surface Water Index
LULC	Land use / Land cover
MCA	Multicriteria Analysis
MCDA	Multi-Criteria Decision Analysis
MIRBI	Mid-InfraRed Bispectral Index
MRA	Multi-resolution analysis
MS	Multispectral

MSI	Multispectral Instrument
NASA	National Aeronautics and Space Administration
NBR	Normalized Burn Ratio
NDBI	Normalized Difference Built-up Index
NDDI	Normalized Difference Drought Index
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NER	Named Entity Recognition
NetCDF	Network Common Data Form
NIR	Near-Infrared
NLP	Natural Language Processing
NUTS	Nomenclature d' Unités Territoriales Statistiques
OSM	Open Street Map
PAN	Panchromatic
PDSI	Palmer Drought Severity Index
PET	Precipitation and Potential evapotranspiration
PSNR	Peak Signal-to-Noise Ratio
PNN	Pansharpening Neural Network
RCM	Regional Climate Models
RCP	Regional Climate Projections
RDI	Reconnaissance Drought Index
RdNBR	Revitalized difference Normalized Burn Ratio
RGB	Red Green Blue
SAR	Synthetic Aperture Radar
SAVI	Soil-Adjusted Vegetation Index
SDG's	Sustainable Development Goals
SMA	Soil moisture Anomaly
SM	Spectral Metrics
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture and Ocean Salinity
SPEI	Standardized Precipitation Evapotranspiration Index
SPI	Standardized Precipitation Index
SRCNN	Super-Resolution Convolutional Neural Network

ST	Statistics-based methods
SWIR	Short-wave infrared
TCI	Temperature Condition Index
UIQI	Universal Image Quality Index
UNCCD	United Nations Convention to Combat Desertification
VCI	Vegetation Condition Index
VPD	Vapor Pressure Deficit
VNIR	Visible and near-infrared
WFAS	Wildland Fire Assessment System
WHO	World Meteorological Organization
WPC	Wavelet Transform

Glossary

Climate impacts	The consequences of realized risks on natural and human systems, where risks result from the interactions of climate-related hazards (including extreme weather and climate events), exposure, and vulnerability. Impacts generally refer to effects on lives; livelihoods; health and well-being; ecosystems and species; economic, social and cultural assets; services (including ecosystem services); and infrastructure (based on IPCC, 2018)
Agricultural Drought	Agricultural drought by definition refers to conditions that result in adverse plant responses, which can range from reduced crop and forage yields to total crop or forage failure.
Drought	Drought is a prolonged dry period in the natural climate cycle that can occur anywhere in the world. It is a slow-onset disaster characterized by the lack of precipitation, resulting in a water shortage. Drought can have a serious impact on health, agriculture, economies, energy and the environment.
Flood	An overflow of water onto normally dry land. The inundation of a normally dry area caused by rising water in an existing waterway, such as a river, stream, or drainage ditch. Ponding of water at or near the point where the rain fell. Flooding is a longer term event than flash flooding: it may last days or weeks.
Heatwaves	Heatwaves, or heat and hot weather that can last for several days, can have a significant impact on society, including a rise in heat-related deaths. Heatwaves are among the most dangerous of natural hazards, but rarely receive adequate attention because their death tolls and destruction are not always immediately obvious.
Flash Flood	A flood caused by heavy or excessive rainfall in a short period of time, generally less than 6 hours. Flash floods are usually characterized by raging torrents after heavy rains that rip through river beds, urban streets, or mountain canyons sweeping everything before them. They can occur within minutes or a few hours of excessive rainfall. They can also occur even if no rain has fallen, for instance after a levee or dam has failed, or after a sudden release of water by a debris or ice jam.

Meteorological Drought	Meteorological drought is defined usually on the basis of the degree of dryness (in comparison to some “normal” or average amount) and the duration of the dry period. Definitions of meteorological drought must be considered as region specific since the atmospheric conditions that result in deficiencies of precipitation are highly variable from region to region.
Wildfire	Wildfires (also known as bushfires, brush fires or forest fires) are large, uncontrolled and potentially destructive fires that can affect both rural and urban areas. They can spread quickly, change direction and even 'jump' across large distances when embers and sparks are carried by the wind. They are caused by a range of natural causes (such as lightning) or by human carelessness (such as a discarded cigarette). The spread of a wildfire depends on the arrangement of land, available fuel (vegetation or dead wood) and weather conditions (wind and heat).

Executive summary

This report is a D2.3 deliverable of the ClimEmpower project. It presents a **methodology to address critical regional gaps related to climate hazards** and to improve risk assessment and resilience building. These gaps include flood and fire hazard mapping, access to meteorological data, drought monitoring for agricultural areas, high-resolution spatial data, and social sensing data—key elements for effective disaster management and resilience planning.

ClimEmpower aims to meet regional needs by proposing tailored datasets for impact assessment on environmental, economic, and social levels, as well as for comprehensive risk analysis. The methodology incorporates **overlaying exposed infrastructures, population data, and other relevant datasets to assess risk holistically**. This framework enables impact assessments using satellite-derived indicators.

For disasters such as wildfires and floods, **satellite indices** are used to estimate burned areas and fire severity rapidly, facilitating timely implementation of restoration actions. For floods, the estimation of inundated areas, combined with data on property values, crop types, affected infrastructure, and population, supports accurate and integrated planning. Similarly, for droughts, combining satellite indices with meteorological data enhances monitoring capabilities and enables strategic water resource management. This integrated methodology fosters informed decision-making and appropriate resource allocation, contributing to an effective and proactive climate adaptation framework.

To address the need for high-resolution multispectral satellite data, this document evaluates **geographical downscaling methods to improve the spatial resolution of Sentinel-2 data**. Various approaches for spatial downscaling have been tested, aiming to enhance the resolution of low-resolution spectral bands in Sentinel-2 data. These improvements increase the accuracy and usability of satellite data for environmental monitoring and decision-making processes.

Finally, the innovative use of **social media is highlighted as an important tool for real-time disaster monitoring**. By analysing user-generated content from platforms like X (formerly Twitter), timely updates on extreme weather events can be provided. Additionally, advanced machine learning and natural language processing techniques address the challenge of filtering relevant content. This approach strengthens disaster monitoring and ensures resources are allocated effectively when needed.

1. ClimEmpower summary

ClimEmpower is a Horizon Europe collaborative research project dedicated to addressing the ongoing Climate Crisis in Europe by empowering the regional stakeholders in some of the most vulnerable European regions (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

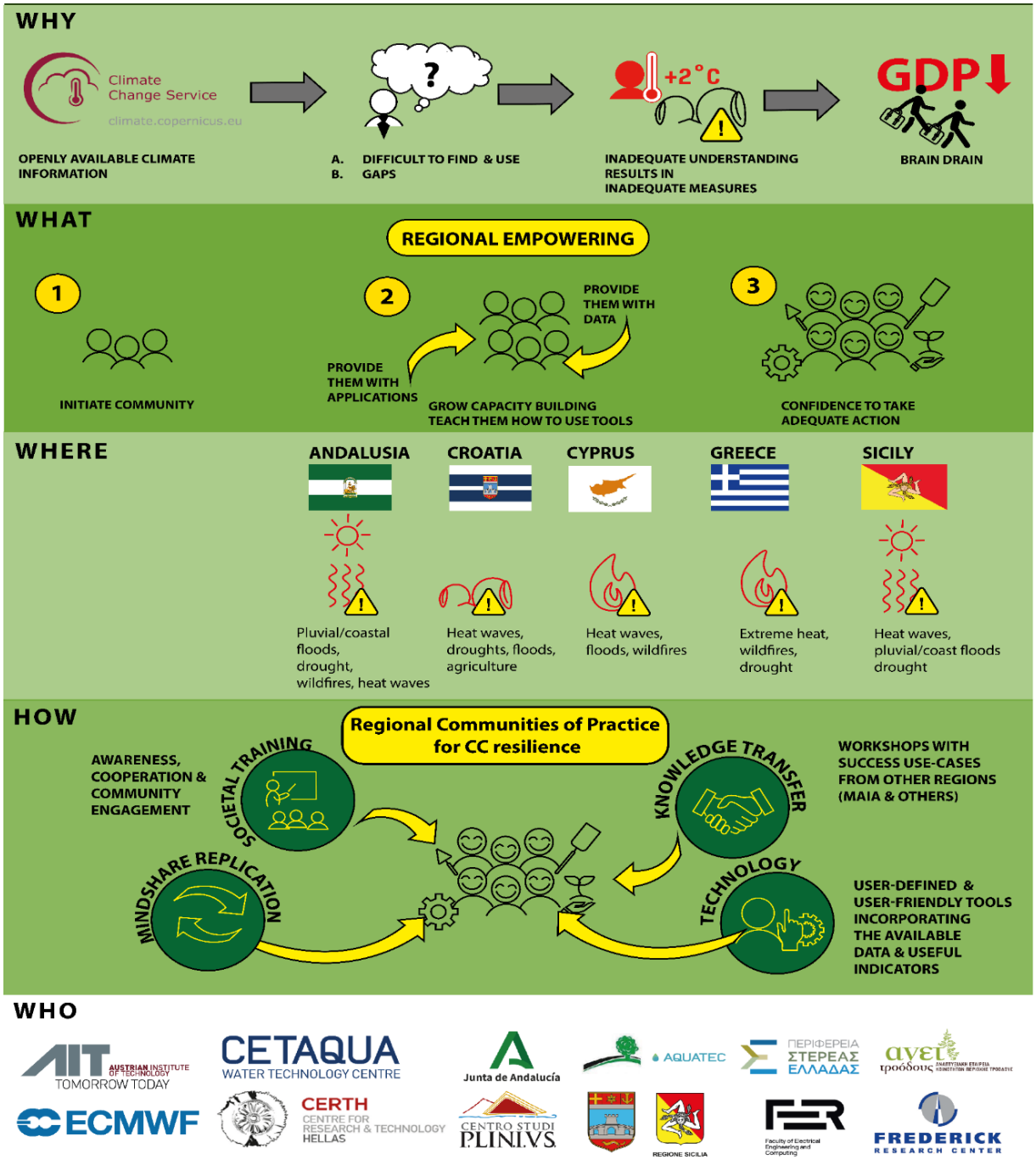


Figure 1: ClimEmpower at a glance: why, what, where, how, and who.

1.1 Project Context

Climate risks results from a combination of hazard, exposure, and vulnerability (IPCC, 2022). Addressing all three aspects is crucial for effective increase of regional resilience. However, exposure, vulnerability, and related aspects, such as adaptive capacity, strongly depend on available knowledge and climate literacy. Consequently, global climate crisis frequently has a higher impact on socioeconomically vulnerable regions, thanks to a higher human and economic potential for addressing the issue in more affluent regions. To maximize its impact, ClimEmpower has therefore chosen to address the EU regions featuring a combination of high potential CC impacts and low and/or stagnant regional GDP/capita. This is mainly the case for regions in South and Southeast Europe (Figure 2).

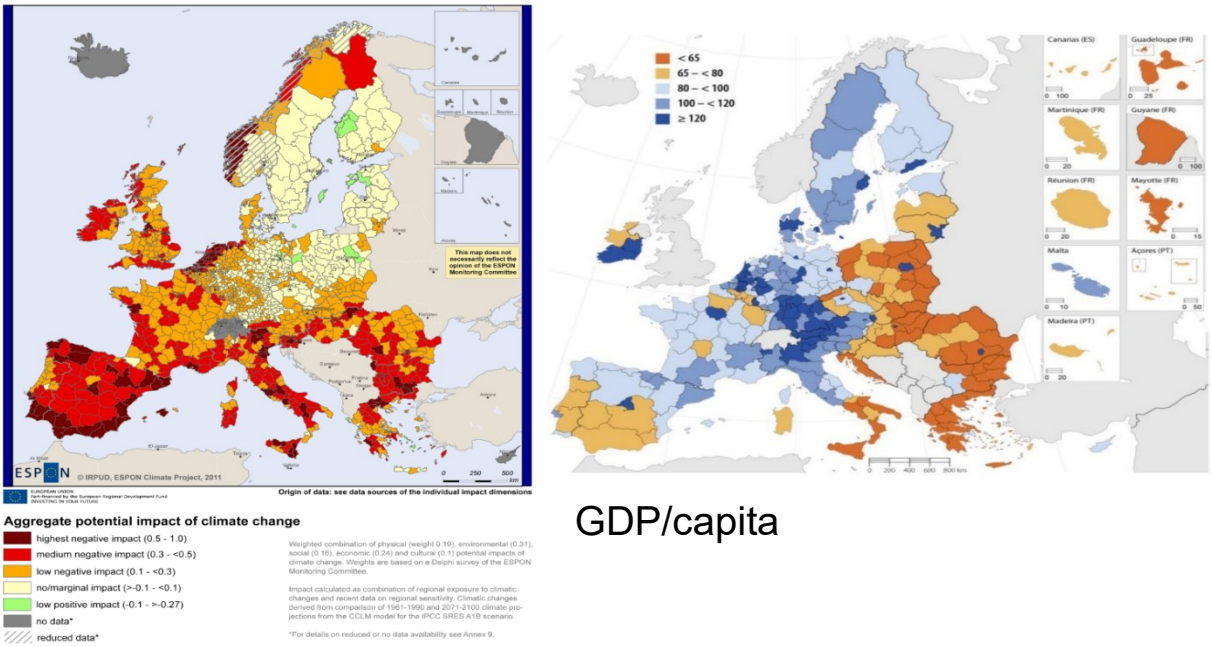


Figure 2: Aggregated potential impact of climate change and GDP/capita; Sources: (Espon, 2012) and (Eurostat, 2021).

The context the project addresses is thus one of an ongoing global warming, high regional vulnerability and low coping capacity of the participating regions, and the **overarching strategic objective of ClimEmpower** is to empower the Regional Authorities (RAs) and other Mission Users (MUs) in five EU-regions featuring a combination of exceptionally high climate hazards and exceptionally low coping capacity by improving their collective understanding of the Climate Change (CC) hazards, risks and resilient development pathways and supporting their knowledge-based regional planning and development through provision of relevant data, knowledge and user-defined and user-friendly decision support applications.

1.2 Project Objectives

To achieve this overarching goal, ClimEmpower has identified six SMART2 Strategic Objectives (SO), each one related to one or several work packages. The SOs have also been classified according to different categories: societal, contributing to improved dialogue, awareness, cooperation and community engagement as highlighted by the European Climate Pact (SO1, SO5); scientific, corresponding to research activities for advances beyond the state of the art (SO2, SO3); technological, suggesting and/or developing novel solutions, integrating state-of-the art and digital advances (SO4); and outreach, aimed at sharing ClimEmpower results to a broader scientific and non-scientific audience, including additional regions and communities, to maximize project impact (SO6).

SO1	Understand regional background, challenges, and expectation (WP1, societal)
SO2	Addressing the gaps in availability and usability of CC data and services (WP2 and WP4, scientific)
SO3	Identification, definition, estimating, and communication of climate impact/resilience indicators suitable for local end-users (WP2 and WP4, scientific)
SO4	Simplify access to CC data and development of end user applications (WP3, technological)
SO5	Empower the regions to activate and enhance their potential for addressing the climate change challenge. (WP4, societal)
SO6	Ensure the use and impact of the ClimEmpower outputs (WP4 and WP5, scientific and societal)

ClimEmpower's key ambition is to prove beyond doubt that CC-resilience should, and can, be an integral part of regional development everywhere in EU and beyond it. That is, we anticipate that the regional stakeholders will recognise that CC-resilient development pathways offer multiple benefits to them, including but not limited to higher quality of life and reviving economy, and that these can be understood using available data, tools, and services. Second key ambition of the project is to help the regions address the CC resilience in key community systems addressed in five ClimEmpower trials.

Underlying philosophy of the project is to **"help the regions to help themselves"**. This will be achieved through various mechanisms, including co-creation and mediation of the regional **"Communities of Practice"**, provision of the **Climate Change -resilience training materials**, as well as in provision and training in use of the user-centric data and services –

[2]Specific (related to WPs), Measurable (by relevant KPIs), Achievable (the WPs in which they will be achieved are listed), Realistic (since they are referred and explained in the methodology section), and Timebound (each KPI is related to a deliverable and a month of achievement).

2. Introduction

2.1 Deliverable summary

The objective of this report is to **outline methodologies to address regional gaps in climate hazard resilience**. This goal can be achieved **through the integration of spatial data, social sensing, and multi-hazard datasets**.

To enhance climate hazard resilience, this report presents a comprehensive approach that integrates hazard mapping, exposure analysis, vulnerability assessment, and satellite-derived indicators. Satellite applications play a crucial role across various hazard events. For instance, they are used for the rapid estimation of burned areas and fire severity after wildfire events to inform restoration planning. Similarly, satellites aid in inundation mapping combined with socio-economic data for flood events and employ high-resolution satellite indices for drought monitoring. Additionally, the integration of satellite data with other meteorological indicators further strengthens disaster adaptation planning.

Particularly, in **chapter 3** of this deliverable a description of selected climate services and datasets is given due to their ability to directly address the identified gaps and provide critical information in spatio-temporal monitoring of wildfire susceptibility, flood exposure, drought risk and heatwave prediction which are also parts of this chapter.

In **chapter 4** the deliverable emphasizes in downscaling methods to improve the spatial resolution of Sentinel-2 spectral bands, enhancing their utility for environmental applications, focusing on pansharpening technics which aligns in environmental monitoring and hazard assessment.

Finally, **chapter 5** highlights the innovative use of social media platforms, particularly X (formerly Twitter), for real-time disaster monitoring. Location extraction, event detection and context analysis enhance awareness. By integrating these platforms with GIS and NLP techniques, emergency response efforts can overcome challenges and support effective decision-making.

Improvement on spatial resolution of Sentinel-2 data in combination with targeted datasets and social media offer a multidisciplinary approach and strengthens the adaptive capacity of regions and showcasing the comprehensive framework of the project.

2.2 Methodology to close the identified gaps

This deliverable employs a multifaceted methodology to address critical regional gaps and needs in climate hazard resilience by integrating various datasets and services with innovative analysis techniques. The approach is specifically designed to support climate hazard monitoring, impact assessment, and resilience building, providing a robust framework for regional adaptation and decision-making.

A comprehensive **review of open datasets and services** was conducted to address regional data gaps. These datasets and climate services, which are relevant to hazard analysis, risk mapping, and socio-economic assessment, form the foundation for identifying vulnerabilities and building effective resilience strategies.

The methodology integrates hazard mapping with exposure elements to identify vulnerable areas and assess the impacts of various hazards. For **drought** assessment, the application of remote sensing indices is highlighted, particularly through the development and use of drought indices combined with meteorological data to evaluate moisture deficits. Incorporating soil moisture data further enhances the accuracy of drought monitoring and impact evaluation.

Given the significance of **heatwaves** for the ClimEmpower regions, the deliverable also identifies the most applicable indices for heatwave monitoring and assesses the intensity and impacts of heatwaves, which are critical for understanding their implications on socio-economic stability and regional ecosystems.

For **wildfire and flood hazards**, a methodology is presented that incorporates susceptibility mapping to predict areas at risk and estimate the potential impact of hazards. This is achieved by combining these areas with exposure elements and socio-economic data. Satellite-derived indices are utilized for burned area detection and inundation mapping, supporting restoration strategies. Additionally, tools like LISFLOOD are employed for flood extent modelling, with detailed descriptions provided on their purpose, required input parameters, and the outputs generated from real-time and forecast data.

Socio-economic vulnerability mapping through the EUROSTAT grid is also addressed. This involves analysing and integrating data for targeted use cases, such as risk assessments. Considering the main gaps in high-resolution data for ClimEmpower regions, the deliverable includes methods for spatial downscaling of Copernicus Satellite Images. These methods focus on addressing challenges in spatial resolution and enhancing the utility of low-resolution spectral bands for environmental applications, particularly climate hazard monitoring.

In addition to **spatial data integration**, this deliverable incorporates **social media analysis** and real-time data exploration are incorporated for disaster monitoring, event detection, and continuous monitoring. Natural Language Processing (NLP) and Large Language Models (LLMs) are leveraged to integrate these data sources with Geographic Information Systems (GIS), providing actionable insights with spatial context.

In conclusion, this deliverable outlines methodologies to address identified data gaps, improve hazard monitoring, and support decision-making with real-time data for disaster resilience and adaptation planning.

The significance of these methodologies becomes even more apparent when considering the broader context of climate-related hazards such as floods, wildfires, droughts, and heatwaves, which pose significant risks at the EU level. Their impacts on infrastructure, ecosystems, agriculture, and socio-economic stability need to be thoroughly assessed. Addressing these challenges requires a focus on data availability and gaps, spatial resolution, and decision-making tools, all of which are crucial for building resilience.

The methodologies proposed in this deliverable align closely with the Intergovernmental Panel on Climate Change (IPCC) recommendations In its Sixth Assessment Report (AR6), the IPCC emphasizes global climate risk assessment and highlights the need for regional adaptation strategies. The report underscores the importance of data integration and predictive modelling to mitigate risks and improve management effectiveness. Additionally, AR6 Working Group II stresses the need for locally tailored adaptation measures to anticipate and address specific climate risks. The IPCC also highlights the necessity of integrating diverse datasets, including

satellite imagery and socio-economic data, to better understand and respond to climate challenges.

These insights are also consistent with the United Nations Sustainable Development Goals (SDGs). Specifically, SDG 11 focuses on urban resilience by addressing risks from all climate hazards. It emphasizes assessing vulnerabilities, mapping exposed infrastructure and populations, and prioritizing interventions. Similarly, SDG 13 underscores the importance of climate action, improved hazard monitoring, early warning systems, and strategic adaptation planning to mitigate the impacts of climate-related hazards effectively.

Summarizing the key points of the methodology are:

Table 1. Key points of the methodology

Key points	Description
Review of important open datasets and Climate services	Analysis of selected datasets and services for hazard and risk analysis and mapping and socio – economic assessment
Integration of hazard mapping and exposure elements	Identification of vulnerable areas and combination with socio-economic data
Drought monitoring and assessment	Use of remote sensing indices, soil moisture data to monitor drought and evaluate impacts
Heatwave monitoring	Relevant heatwave indices and evaluate socio-economic and ecological impact
Wildfire susceptibility and burned area mapping	Susceptibility mapping for identification areas in risk and burned area mapping for assess the impact
Flood susceptibility and impact	Flood hazard mapping, combination with models (LISFLOOD) and exposure and socio-economic data for more accurate impact assessment
Socio-Economic Vulnerability	EUROSTAT grid and additional data for risk assessment
Spatial Downscaling of Copernicus Sentinel2 Images	Address challenges in spatial resolution
Social Media Analysis	Leveraging of NLP and LLMs methods for real-time event detection and disaster monitoring – crowdsourced impact assessment

2.3 Relation to other work

The main gaps of the ClimEmpower regions are identified in **Deliverable 1.2 "ClimEmpower Scenarios"** and the resilience assessment and in Deliverable 2.1 "Climate change resilience: identified data, services and gaps" in which each region outlines its specific needs and gaps according to the climate hazard that need to be addressed.

In Deliverable 2.1 many available and useful open datasets and climate services have been collected in key categories of risk management (hazard, exposure ad vulnerability. Identifying regional gaps and needs was a critical component of this deliverable, particularly in terms of map availability, higher spatial resolution data, and crowdsourced information which provide valuable inputs for deliverable 2.3.

Some of identified datasets also utilized in this methodology in order to combine with other data and offer a holistic approach in regional resilience framework. Furthermore, there is a connection with **Deliverable 2.2** in which socio-economic, and governance indicators are going to be analysed to address the vulnerability as mentioned from most regions as need. The combination of Deliverables 2.2 and 2.3 is expected to address the needs of the regions as each one provides critical elements in risk management and resilience enhancement.

The methodologies that outlined in Deliverable 2.3, particularly the downscaling of spatial data, and the various datasets have a strong contribution to tasks in WP3, specifically in the development of advanced data services and tools. There is a strong connection with Task 3.2 as for the data centric services as they rely on the methodologies for hazard mapping, satellite data and social sensing. These methods enable the development of region-specific indicators for resilience assessment. Additionally, **Task 3.3,** focused on the GUI development, benefits from this deliverable by ensuring compatibility with regional needs and accessibility for end-users.

In relation to **WP4**, the approaches described in this deliverable directly support Community of Practice (CoP) activities by providing educational material. These materials aim to enhance the stakeholder's knowledge for the climate hazards and enhance their understanding of these disasters and the methodologies applied. Furthermore, the methodologies will be validated through regional trials ensuring they meet the usability and requirements of local needs.

Beyond the ClimEmpower project, the developed methodologies have strong connections to **other EU Horizon projects**. There is a significant link to ICARIA project, which focuses on enhancing the climate risk analysis with a focus on decision-making tools and policy support. Additionally, there is connection with methodology aligns with the CLIMAAX project, as both projects prioritize hazards and impact assessment, offering actionable insights through cross-disciplinary collaborations and climate adaptation strategies.

3. Collection, analysis and combination of data from different sources

This section discusses the collection and analysis of open datasets and climate services as well as the integration with data from various sources to address climate hazards and enhance resilience. It emphasizes the importance of combining different data sources with advanced indices to generate insights that strengthen climate resilience and disaster risk management.

The identified gaps summarized in deliverable 2.1 “Climate change resilience: identified data, services and gaps” are different, and the adaptive capacity of each region is also diverse. As for the gaps in drought prediction, heatwave forecasting, flood and fire hazard maps, and training-educational material, these are needs and gaps discussed by most ClimEmpower regions. While open datasets and services offer insightful information on hazards, they often don’t meet the regional needs or don’t cover all regions. To address this, all available data should be integrated into a single application, making it accessible to administrators and stakeholders.

In order to provide comprehensive and practical insights, effective climate risk assessment and management must integrate a wide range of datasets and approaches. Chapter 3 discusses the significance of the use of open datasets and climate services to address climate related hazards. Integration of climatic indicators, especially from Copernicus Data Store (CDS), data from climate services, indices derived from satellite images and meteorological past observations and future projections can provide a comprehensive understanding of different risks. Additionally, by gathering all these data in an appropriate way offers a valuable tool in climate resilience strategies.

The section begins by introducing open datasets and climate services, emphasizing their potential to inform climate monitoring and risk assessments. It then examines the use of satellite-derived indices for drought monitoring, such as NDVI, EVI NDWI, LSWI, NDDI and NMI, as well as meteorological drought indices like SPEI, VPD and Keetch Byram drought index which provide complementary viewpoints on water stress and availability.

Beyond drought monitoring, the section discusses critical indicators for other climate hazards. It describes methods for estimating soil moisture and heatwave intensity, both essential for understanding regional risks. For wildfire and flood hazards, advanced susceptibility and vulnerability mapping methodologies are described, including the use of satellite-derived indices for burned area detection, flood extent estimation, and impact assessments.

Finally, this section builds a strong framework for identifying vulnerabilities, assessing risks, and influencing decision-making by integrating spatial and temporal data from diverse sources. It highlights the value of incorporating varied datasets into comprehensive hazard assessments, enabling stakeholders to better understand, monitor, and mitigate the effects of climate change.

3.1 Overview of open datasets and climate services useful to close the gaps

Climate services and datasets are critical for understanding, assessing, and addressing the diverse impacts of climate change. These tools provide critical data for monitoring climate variables, forecasting extreme occurrences, and assessing vulnerabilities and risks across various industries and regions. Their value lies in their ability to transform complex climate data into actionable insights, assisting in decision-making processes at the local, regional, and global levels.

Climate services, such as risk dashboards, typology maps, and hazard-specific tools, enable stakeholders to understand and assess climate hazards efficiently. They facilitate informed planning, risk mitigation, and adaptation strategies by offering access to high-quality, tailored data. Similarly, datasets containing historical and predicted climatic indicators, river discharge data, fire hazard indices, and hydrological models form the scientific foundation for complete climate risk assessments.

By integrating these resources into processes, authorities, researchers, and planners can detect hazards, analyse possible impacts, and prioritize solutions in a more effective manner. Finally, climate services and databases are useful as they can bridge the gap between scientific data and practical applications, fostering resilience and long-term decision-making in the face of climate change.

3.1.1 Climate Services

Climate services and datasets are critical in understanding and resolving the different effects of climate change in ecosystems, societies and also economies. Services like the Global Framework for Climate Services (GFCS), Climate Risk Dashboards, the European Risk Typology Map, and the Pan-European Fuel Map Service, give actionable insights by combining climate data into user-friendly solutions designed for a wide range of stakeholders. These platforms provide decision-makers, researchers, and industry professionals with region-specific information needed for planning and resilience building.

Datasets focusing on key indicators like river discharge, hydrology-related climate impact metrics, fire hazard measurements, and broader climate indicators across Europe complement these services, forming a robust foundation for informed decision-making. The use of these datasets by stakeholders to forecast climate hazards and enhance adaptation and mitigation strategies is vital. Additionally, climate services and datasets provide tools to a wide range of stakeholders, helping to increase preparedness and support sustainable development in the face of climate change.

Global Framework for Climate Services (GFCS)

The Global Framework for Climate Services (GFCS)[3] provides climate data and services and supporting decision – makers about climate adaptation. The GFCS led by Word Meteorological Organization and its climate data and information are comprehensive and applicable in

[3] GFCS, https://wmo.int/site/global-framework-climate-services-gfcs

different sectors like agriculture, water, disaster risk reduction and focus at regional and local level. Additionally, supports delivery and use of climate services at different levels.

The Climate Service Toolkit (CST)[4] developed through the Global Framework for Climate Services (GFCS), provides data, software tools, guidance and training for climate services development. The toolkit focuses on key areas such as climate data rescue, climate analysis and monitoring, climate prediction and projection. The CST offers access to a variety of climate datasets, including historical observations, reanalysis data and climate projections. Additionally, it contains an extensive collection of technical documents that provide detailed information about the data and resources available within the toolkit. Specifically, the following are included:

- **Climate Data Rescue Management and Mining**: This involves a list of key services and datasets for data rescue, managing, and mining global observed data from various meteorological parameters (air temperature, sea surface temperature, precipitation), as well as observed and reanalysis data at global and regional levels. Additionally, it includes guidance on climate data observations and training materials to teach users how to utilize the various applications of global satellites and data applicable to climate datasets.
- **Climate Analysis and Monitoring**: The report provides historical context and insights into climate variability and long-term trends. It presents methodologies and tools specifically designed for climate monitoring and understanding climate variability. The guide emphasizes key analytical techniques, highlights the importance of climate indices, and outlines best practices for data integration.
- **Climate Prediction**: This document provides detailed guidance on climate data, prediction tools, and the most suitable methods for effective climate adaptation. It emphasizes climate prediction models, reanalysis data, and datasets for forecasting climate impacts. Additionally, the document includes training modules that offer users hands-on experience with interactive examples. It also covers how to communicate climate products and forecasts to stakeholders or the public. Case studies and frameworks illustrate how these tools support stakeholders in making informed decisions regarding climate adaptation strategies in response to climate variability.
- **Climate Projection:** This report presents long-term projections based on different scenarios of future societal development, with a specific focus on building societal resilience to climate change and variability. It outlines methodologies for using projection data to analyses the impacts of climate change in key sectors such as agriculture, water resources, and health. The report also investigates best practices for downscaling global projections to the regional level to guide adaptation planning.

Climate Risk Dashboard

The Climate Risk Dashboard[5] is an interactive web tool which was developed by PROVIDE horizon Europe funded program and provides information for different future global scenarios. The indicators cover sectors related to climate variability, extreme events, urban heat stress

[4] CST, https://gfcs.wmo.int/site/global-framework-climate-services-gfcs/climate-services-toolkit/about-cst
[5] Climate Risk Dashboard, https://climate-risk-dashboard.climateanalytics.org/

and oceanic habilitation. It is important to mention that not all indicators apply to every sector. The dashboard was developed in coordination with stakeholders and adaptation practitioners. It is designed as an interactive application that enables users to explore future climate scenarios. The PROVIDE approach allows users to establish impact thresholds as the starting point for analysis and adaptation.

The dashboard provides options for selecting countries, cities and indicators. Terrestrial climate, biodiversity, urban heat stress, marine climate, global carbon cycle and glaciers are among the targeted sectors. However, none of these sectors currently have data for all countries or cities. For example, Greece has 10 terrestrial climate indices, 1 macroeconomic indicator and 4 for biodiversity indices.

The core functionalities of this service revolve around geographical and indicator selection. Thus, users can choose specific regions and indicators to analyse, gaining insights into climate risks such as temperature changes, precipitation patterns, biodiversity loss and heat stress in urban environment. The dashboard enables users to explore future climate impacts and provide insights into different emission paths, helping them understand the consequences of various climate action. The Climate Risk Dashboard is a valuable resource for climate scientists, climate risk experts, policy makers and urban planners. Additionally, users can access tutorials and interviews with the developers to better understand how to utilize the dashboard successfully and understand the underlying scientific concepts.

At the national level the dashboard focuses on terrestrial and biodiversity indices (i.e., extremely cold/hot year, FWI during the fire season, and soil moisture content) and limited number of macroeconomic indicators (e.g., climate change impact on GDP). This facilitates assessments of environmental health and economic sustainability, offering valuable approach for policies aiming at long-term ecosystem protection and conservation.

At the city level, the urban heat stress indicators (i.e., number of heatwave days per year, mean daily (maximum) temperature, population exposed to extreme heat stress, and lost annual working hours used for outdoor activities) provide critical insights for major cities in each country. In urban areas, population density and concentrated infrastructure exacerbate heat stress phenomena, affecting public health and quality of life. This information supports the design of resilient and climate adapted urban environments.

The indicators available at country and city level are essential for assessing climate needs, guiding decisions and implementing adaptation strategies at various geographic scales.

European Risk Typology Map

The European Risk Typology Map (CRT)[6] is an interactive tool designed to explore the typology of risks and supports adaptation and resilience strategies. It allows users to visualize, compare and assess climate risk in Europe. CRT is designed to enhance the understanding of climate risk in European cities and regions through visualization, description, comparison and analysis, thus improving adaptation and resilience planning.

[6] European Risk Typology Map, https://european-crt.org/index.html

The CRT platform is structured around the four risk domains (hazard, exposure, sensitivity and adaptive capacity) outlined in Intergovernmental Panel on Climate Change (IPCC) guidelines. Most of the indicators included in the platform were developed as part of the RESIN project to define the typology. Each region's climate risk profile is created using a set of indicators that assess:

- **Hazards**, assessed based on likelihood and impact
- **Vulnerability**, determined by socioeconomic and demographic characteristics which affect sensitivity
- **Adaptive capacity,** which accounts for infrastructure, resources and resilience.

The tool integrates climate projections, historical climate data and socioeconomic datasets to examine different climate scenarios. Users can select geographic areas, climate variables and timeframe to analyse. The platform provides data visualisation through maps and graphs, allowing users to interpret the information. A risk assessment for the vulnerability on different sectors of climate impacts offers insights on potential future risks according to various climate scenarios. After all the potential impact on the region might contribute to strategic planning.

The CRT according to the IPCC framework, includes hazard, exposure, sensitivity and adaptive capacity, reflecting how various factors contribute to overall risk. The indicator-based analysis includes a detailed set of indicators, creating a profile of climate risk for each region. Additionally, each hazard is assessed based on its likelihood and impact, with indicators focusing on the social and economic characteristics that influence the regional vulnerability.

An interactive map is the primary output, allowing users to explore climate risks at the selected regions, with the NUTS3 level being the most granular. The typology's classes and sub-classes are based on the position of cities or regions within Europe's diverse climate risk landscape. This classification groups cities and NUTS3 regions into eight classes that share similar climate risks, often aligning geographically (e.g., Mediterranean, Northern Europe, or inland areas). Sub-classes, ranging from 3 to 5 per class, distinguish regions further based on specific climate risk characteristics. More details about these classes and sub- classes can be found in the final report of typology[7].

Data visualization integrates all considered indicators into map layers across domains. A key feature of this interactive tool is its ability to provide downloadable data for detailed research or policy planning. Moreover, this tool supports local governments in adaptation planning, enabling comparative analyses between regions with similar climate risk profiles to identify optimal strategies. Furthermore, as adaptation plans are implemented and new data becomes available, the CRT will facilitate tracking changes over time.

In summary, the European Climate Risk Typology is a critical resource for understanding climate risk, enhancing regional adaptation efforts, and enabling comparative analysis across Europe. Within the ClimEmpower project, it serves as a structured platform for integrating data across regions, catering to diverse stakeholders, including policymakers, urban planners, and researchers.

[7] https://european-crt.org/files/typology-final-report.pdf

Pan-European Fuel Map server

The Pan-European Fuel Map-server [8]is an open geodata portal that supports fire risk assessment. It is an output of the European Union's project for fire-resilient territories in Europe, "Fire-Res". The tool is hosted on the Google Earth Engine platform and combines datasets and methodologies to aid fire management strategies. In the past, significant efforts have been made to collect data on fuel conditions and assess potential fire behaviour under various parameters. "Fire-Res" allows the integration of key factors related to fire ignition, progression, and management is consolidated into a unified platform. Its primary aim is to provide critical information, enhance management and planning of protective measures, and ensure accessibility and usability for diverse stakeholders.

The tool receives as input parameters the topographic data, canopy and surface fuels, vegetation indices derived from satellite imagery, and canopy features. The output is a pan-European Fuel Map delivered at a high spatial resolution of 100 meters, covering all of Europe. Fuel type classification follows Anderson's Fire Behaviour Model. It is important to mention that the platform includes selection tools for variables, countries, and EU NUTS levels (0 to 3). After selecting parameters, users can visualize the results on the map and export them in NETCDF format for further assessment (Kutchartt, E., et al., 2024). The outputs can be used for conducting fire risk analysis at various administrative levels, integrating with weather and climate datasets for risk evaluation, and supporting climate adaptation and fire risk mitigation efforts.

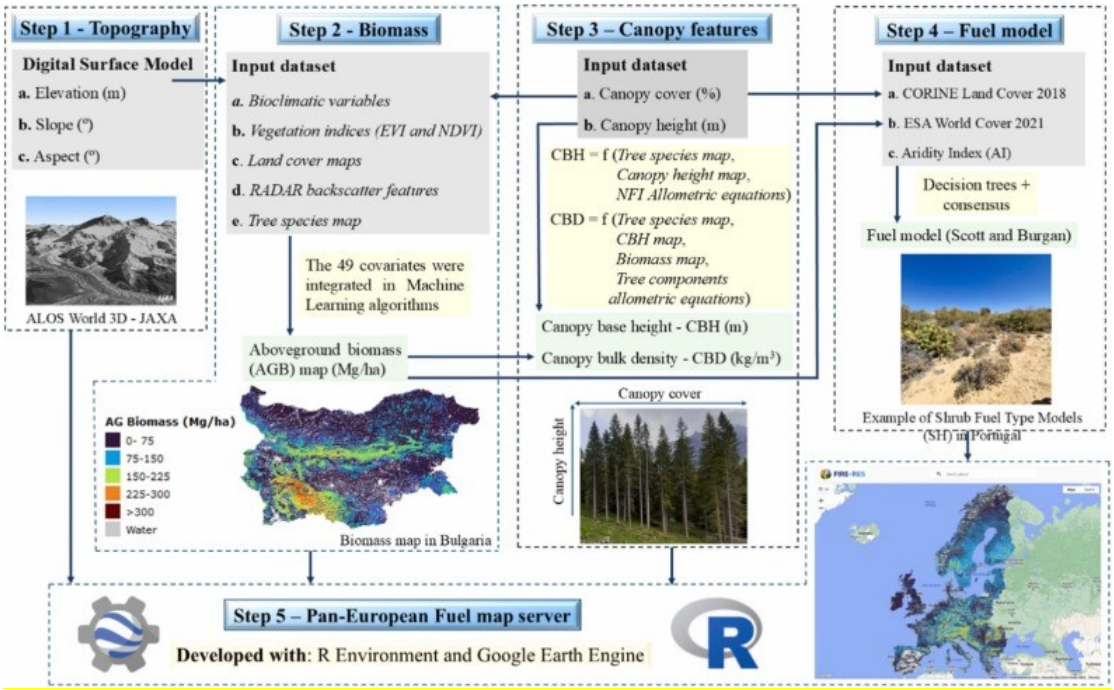


Figure 3: The flow chart of steps and inputs needed for Pan-European Fuel Map (Source: Kutchartt, E., et al., 2024).

[8] https://www.cirgeo.unipd.it/fire-res/app/

3.1.2 Open Datasets

There are many open datasets available for hazard mapping, as well as for providing historical data and projections. One of the most widely used sources for climate datasets is the Copernicus Data Store, a cloud-based platform designed to simplify access to high-quality data for users from diverse backgrounds. It offers a wide range of datasets, including observations, reanalysis, forecasts, and climate model outputs, all presented in a common data model.

River Discharge and Forecasted Data by the European Flood Awareness System

The European Flood Awareness System (EFAS) aids in preparation for major flood events and is the first operational European system for monitoring and forecasting floods. It provides real-time and forecasted river discharge data across Europe, which are critical for flood risk assessment and mapping through gridded modelled hydrological forecasts. The dataset includes both observed and modelled river flow data, along with flood probability information. By predicting potential flood events, it supports early warning systems and enables prompt responses.

This data is valuable for flood risk mapping, as it allows for the identification of high-risk zones, designating areas likely to experience flooding. This is crucial for prioritizing emergency responses. Additionally, it enables accurate flood mapping and infrastructure design to mitigate flood impacts. The dataset also provides data into the LISFLOOD hydrological model. Data is available from October 10, 2018, to the present, with a 30-day delay. Real-time data is only accessible to EFAS partners.

The EFAS-CEMS Early Warning Data Store also provides seasonal river discharge forecasts[9], offering insights into long-term hydrological conditions. These data are generated by ensemble hydrological and meteorological models and provide predictions for rivel flow anomalies, droughts and high flow. Furthermore, this data is useful for climate impact assessments by providing a basis for tracking the effect of seasonal variability on hydrological patterns.

Hydrology-related climate impact indicators from 1970 to 2100 derived from bias adjusted European climate projections

The Copernicus Climate Data Store provides various datasets that are useful for analysing historical events and making projections. While deliverable 2.1 has already mentioned the dataset of bioclimatic indicators additional datasets could be combined with other data to provide a more comprehensive approach to mitigation planning. One such dataset is the Hydrology – related climate impact indicators[10] spans from 1970 to 2100 and integrates past, present and future scenarios.

The dataset is useful for trend analysis, as it provides flood frequencies and supports scenario planning for potential future risks. Furthermore, it enhances flood mapping by integrating dynamic, climate driven variables with static hazard mapping, thus offering an understanding

[9] https://ewds.climate.copernicus.eu/datasets/efas-seasonal?tab=overview
[10] https://cds.climate.copernicus.eu/datasets/sis-hydrology-variables-derived-projections?tab=overview

of long-term risks. Adaptation strategies can be supported by the development of resilient flood management policies.

Fire danger indicators for Europe from 1970 to 2098 derived from climate projections

In terms of fire danger, another dataset in the Copernicus Climate Data Store contains fire danger indicators[11] spanning from 1970 to 2098. This allows for the analysis of past trends, the current situation, and future scenarios under changing climatic conditions (different climate scenarios, RCP 4.5 and RCP 8.5). The indicators include the Fire Weather Index (FWI), the number of days with very high, high, and moderate fire danger, as well as seasonal fire danger.

This dataset is crucial for assessing historical fires and supporting potential fire risk evaluations. It provides critical inputs for fire management strategies and, when integrated with other data (e.g., fuel type, land-use, socio-economic datasets), it enhances regional hazard assessments. These indicators are essential tools for developing tailored mitigation strategies.

Climate indicators for Europe from 1940 to 2100 derived from reanalysis and climate projections

There are several datasets that address various climate change-related situations. The dataset "Climate Indicators for Europe from 1940 to 2100)[12] is a comprehensive collection of climate indicators that combines historical data with future projections, applicable to risks such as floods, fires, heatwaves, and droughts.

This dataset includes flood indicators, seasonal and extreme river discharge, precipitation intensity, and locations prone to flooding from heavy rainfall. It also features fire danger indices, such as the Fire Weather Index (FWI) and the number of high-risk days, identifying regions at increasing fire risk. Drought indicators, such as the Standardized Precipitation-Evapotranspiration Index (SPEI) and soil moisture deficits, offer long-term insights into prolonged conditions that affect agriculture and ecosystems. Additionally, heatwave frequency, intensity, and duration reveal trends critical for public health, energy demand, and planning. All these indices cover the hazard categories established by the IPCC and the European Topic Centre on Climate Change Impacts, Vulnerability, and Adaptation (ETC-CCA).

The hazard insights provided by this dataset facilitate hazard mapping and vulnerability assessments. Furthermore, it demonstrates strategies and provides data to guide mitigation and adaptation policies. All in all, it is a valuable resource for stakeholders in disaster risk reduction and significantly contributes to climate change mitigation efforts.

Flood Extent Enhancement and Water Depth Estimation Tool (FLEXTH)

The Flood Extent Enhancement and Water Depth Estimation Tool (FLEXTH)[13] was recently developed by JRC scientists to estimate flood water depth by combining satellite and topographic data.

---

[11] https://cds.climate.copernicus.eu/datasets/sis-tourism-fire-danger-indicators?tab=overview
[12] https://cds.climate.copernicus.eu/datasets/sis-ecde-climate-indicators?tab=overview
[13] FLEXTH, https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/europe-floods-new-tool-estimate-water-depth-and-extent-2024-09-20_en

Flash floods induce inland flooding, and with limited adaptation to this risk, flood water depth is a crucial component in assessing the impact. JRC scientists developed this tool to address the limitations of satellite imagery, such as cloud cover and the inability of satellites to sense water depth. By combining the advantages of satellite data with topographic information, FLEXTH produces accurate flood maps and introduces depth information, enhancing flood risk management and response.

According to Batterle & Salamon (2024), the input parameters of FLEXTH include a binary raster map of flooded areas and a Digital Elevation Model (DTM). The tool's outputs are water level and depth maps. A key feature of FLEXTH is its scalability and flexibility, which makes it adaptable to regions with or without available data. Additionally, it is a cost-effective tool that provides rapid data in near real-time.

In terms of impact assessment following a flood event, FLEXTH helps identify high-risk areas and allocate resources accordingly. It supports hazard assessments by providing insights into flood-prone regions and informing the design of resilient infrastructure. FLEXTH represents a significant advancement in satellite-based flood mapping and offers important solutions for flood monitoring. As an open-source Python tool, it is still in the testing and validation phase, but it holds great potential.

3.2 Multi-source drought monitoring indices

3.2.1 Drought indices derived from remote sensing

According to WHO drought is a prolonged dry period within the natural climate cycle that can occur anywhere in the world. It is a slow-onset disaster characterized by the lack of precipitation, leading to water shortages. Drought can have a significant impact on health, agriculture, economies, energy and the environment (WHO). Agricultural drought is a specific type of drought disaster that occurs globally and affects a large population (Xiao C., et al., 2023).

Drought is the most serious hazard to livestock and crops in nearly every part of the world. The impact of drought is expected to gradually increase under global warming prompting numerous in-depth studies on drought conducted in recent years (Yujia et al., 2023). Given its global significance, research institutes and universities at both national and international level are actively studying this phenomenon to enhance understanding and improve mitigation strategies and measures (EDO[14]). Compared to other natural hazards, droughts often result in larger associated costs due to their severe economic, environmental and societal consequences which frequently stem from prolonged and widespread drought events (Brown et al., 2013). Drought is a slow-onset hazard, becoming apparent only when its impacts on society and environment are felt (Singh, C. et al, 2021; Wilhite, D. A. 2000). Furthermore, drought impacts are non-structural and extend across large geographical regions (Naumann, G., et al., 2015). Finally, it is worth noting that the World Meteorological Organization (WMO) classifies drought based on the affected domain as meteorological, agricultural, hydrological, and socioeconomic (WMO).

[14] EDO https://drought.emergency.copernicus.eu/

Research on drought, its severity and its impact across various fields such as agriculture, forestry, urban trees and natural environment has become increasingly important due to climate change. Additionally, the use of advanced technology, particularly satellite remote sensing, has become indispensable for drought monitoring and management. Remote sensing provides critical data on vegetation conditions enabling the tracking of changes, especially in crop development. Satellite data also help identify areas facing water shortages, facilitating timely interventions and integrated strategic solutions and the development adaptation plans.

According to literature, there are two basic approaches in monitoring drought, meteorological and vegetation indices (Liu, X., et al., 2016). Meteorological indices are developed using precipitation and evapotranspiration as key inputs. The most well-known and widely used are Palmer Drought Severity Index (PDSI) (Palmer W.C. 1965), Standardized Precipitation Index (SPI) (Mckee T.B., et al., 1993), Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente et al., 2010), and Reconnaissance Drought Index (RDI) (Tsakiris et al., 2007). All these indices rely on ground–based data collected from meteorological stations.

Besides meteorological indices, vegetation indices derived from remote sensing provide a reliable alternative in drought assessment. Specifically, Sentinel-2 is widely used for land monitoring, maritime monitoring, emergency management, and security. Its spatial resolution varies from 10 m to 60 m and its multispectral instrument (MSI) has a radiometric resolution of 12 bits with a combined revisit time of 5 days (Varghese D., et al., 2021). Sentinel 2's, high resolution multispectral imagery is a valuable tool for monitoring drought, offering real time assessment on vegetation health, soil moisture and water resources. Leveraging satellite data is crucial for identifying regions experiencing severe drought impacts enabling timely interventions and resource management.

In recent years, there has been an increased use of remote sensing technology for drought monitoring in Europe, particularly following the introduction of European Space Agency (ESA[15]) Sentinel satellites. These satellite images enable the observation and monitoring of drought-related variables across large temporal and spatial scales. Furthermore, it is worth noting that the European Drought observatory (EDO) utilizes both meteorological and satellite data to monitor and predict drought. Key indicators include the Standard Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), Soil moisture Anomaly (SMA), Fraction of Absorbed Photosynthetically Active Radiation Anomaly (FAPAR) and Normalized Difference Vegetation Index (NDVI) (Sepulcre-Canto et al. 2012).

The spatial resolution and coverage of remote sensing data, along with drought-related variables, have enabled effective environmental management. Additionally, their temporal resolution varies, with some satellites offering daily data and others providing data on a weekly basis.

According to the literature, various vegetation indices serve as crucial tools for monitoring drought conditions. Most of these indices are used to assess vegetation health, water stress, and soil moisture. This research aims to evaluate the most suitable indices for monitoring drought in forested and agricultural lands. Assessing the potential of Sentinel-2 for drought

[15] ESA https://www.esa.int/

monitoring requires consideration of multiple parameters, including vegetation characteristics, soil moisture, evapotranspiration, surface water, and land use/land cover change, as illustrated in the following image (Varghese et al., 2021).

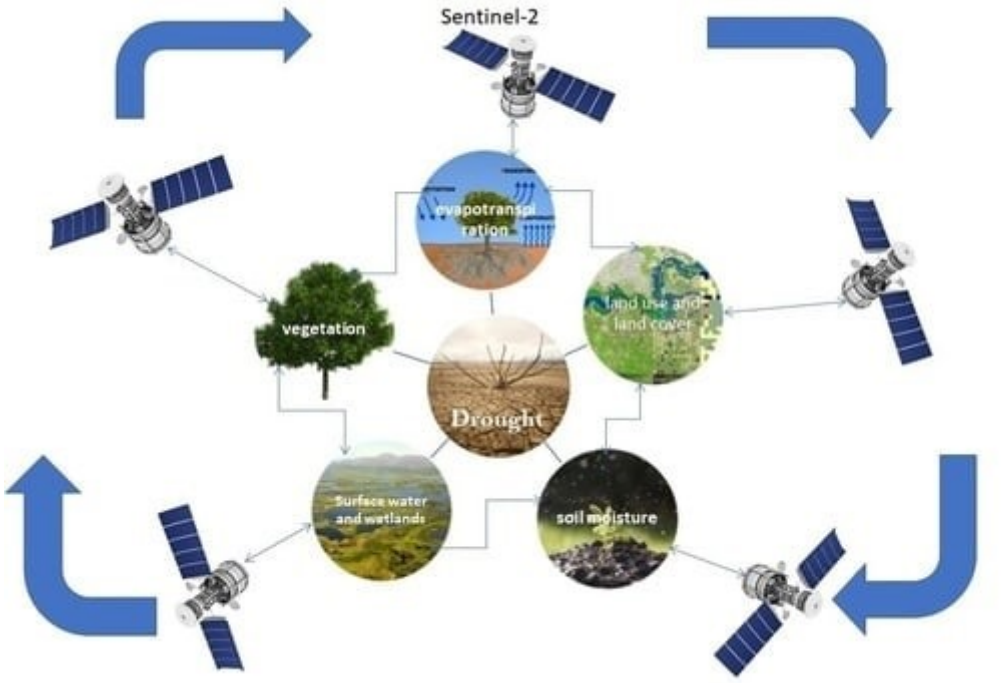


Figure 4: Assessing drought from Sentinel 2 (Source: Varghese, D., et al., (2021)).

Agricultural drought - Remote sensing BASED INDICES

3.2.1.1 Vegetation indices (NDVI, EVI, VCI)

The imprints of drought are often reflected in vegetation conditions (Breshears et al., 2005; Ji L., et al., 2003, Adams HD., et al., 2009). The Normalized Difference Vegetation Index (NDVI) is the most commonly used vegetation index derived from satellite imagery and serves as a primary data source for assessing vegetation health. NDVI was among the first remote sensing-based metrics developed to monitor drought by evaluating vegetation growth and health.

Vegetation responses to drought are estimated through spatio-temporal analysis of NDVI. Variations in NDVI values before and during a drought period are used to quantify the intensity of drought (Varghese, D. et al., 2021).

The mathematical formula for calculating NDVI is the following:

$$NDVI = (NIR + RED)/(NIR - RED\ )$$

where NIR and RED are the bands 8 and 4 respectively for Sentinel-2 data.

The main purpose of the index is to measure the greenness of vegetation. NDVI is used to separate vegetation from soil background and provide measurements for vegetation health (AghaKouchak, A. Et al., 2015). Additionally, time series plots of NDVI have proven valuable in detecting and tracking drought (Almouctar, M. A. S et al., 2024).

Additionally, NDVI is the most common index which is used for estimation of vegetation health and monitoring of climate change impact. As it has already mentioned NDVI is used also in categorizing drought severity, and this is according to comparison of current NDVI values with historical data or long-term averages of the index.

According to F.N Kogan (1995), the thresholds for drought (Table 2), as derived from NDVI index, are outlined below (Pati P. et al., 2024):

Table 2: NDVI drought thresholds

NDVI values	Drought category
-1.0 to 0.3	Severe drought
-0.3 to 0	Moderate drought
0 to 0.3	Mild drought
0.3 to 1	No drought

In summary NDVI is a widely used vegetation index for evaluating vegetation health. However, modifications such as the Soil-Adjusted Vegetation Index (SAVI) and the Enhanced Vegetation Index (EVI) have been developed to address some of the limitations of NDVI, particularly those caused by soil background interference.

Thus, the Enhanced Vegetation Index (EVI) is another vegetation index commonly used as tool for monitoring drought conditions. While NDVI can be influenced by atmospheric conditions, EVI was developed to mitigate these limitations (Yujia et al., 2023). Specifically, EVI is particularly effective at correcting atmospheric interference and soil background signals, making it most suitable for areas with dense vegetation (A. Huete et al., 2002).

The mathematic formula for EVI is:

$$EVI = 2{,}5 * \frac{(NIR - RED)}{(NIR + C1 * RED - C2 * BLUE + L)}$$

, where *L* represents the canopy background adjustment, *C1* & *C2* are coefficients of the aerosol resistance term used to correct aerosol influences in the red band through the use of blue band.

Many scientific studies have noted that NDVI can indicate stressed vegetation but does not explicitly define it as drought or degradation. To address this limitation, the Vegetation Condition Index (VCI) was developed. VCI utilizes historical NDVI sequences of the monitoring target to detect relative changes in NDVI over time (Kogan and Sullivan, 1993). VCI is suggested as a tool, based on NDVI, to detect changes in weather elements that influence vegetation (Joo Heon Lee 2016).

VCI (Vegetation Condition Index) focuses on the impact of weather conditions on vegetation changes, assuming that vegetation achieves its maximum vitality under optimal conditions and its minimum during extreme conditions, such as drought. It compares the current NDVI to the range of values observed in the same period in previous years (Richaud Bertrand, 2019). Moreover, VCI is expressed as a percentage, providing insights into where the observed value falls relative to the extreme values recorded historically. This comparison is conducted for

each specific pixel, evaluating the current NDVI or EVI values against historical data from the same period (UNCCD-led Drought toolbox[16], Agriwatch[17]).

VCI is calculated as:

$$VCI = \left[\frac{NDVI - NDVImin}{NDVImax - NDVImin}\right] * 100$$

High VCI values indicate that the NDVI value is very close to historic maximum while low VCI values mean the NDVI value is near to historic minimum. In summary, VCI reflects how much vegetation conditions deviate from normal. This index is particularly useful in identifying the onset of drought, its duration, and its intensity.

It is also worth noting that the VCI index is often used alongside other indices such as Temperature Condition Index (TCI) () to provide a comprehensive view of vegetation health. The combination of VCI and TCI is a very effective way to measure drought stress on vegetation.

3.2.1.2 Water stress indices for drought monitoring (NDWI, LSWI, NDDI, NMI)

The Normalized Difference Water Index NDWI is a vegetation and water stress index. It utilizes the Near-Infrared (NIR) and Shortwave Infrared (SWIR) bands for its calculation, as the SWIR region has a higher ability to absorb liquid water compared to NIR. Consequently, crops with high water content exhibit higher NDWI values, while stressed vegetation displays lower NDWI values (Jayawardhana & Chathurange, 2020). NDVI and NDWI should be considered complementary indices, as NDWI provides additional information rather than serving as a substitute for NDVI (Bocai Gao 1996).

The mathematic formula of NDWI is:

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR}$$

The NDWI equation produces positive values for water features and negative values (or zero) for soil and terrestrial vegetation. Regarding the interpretation of NDWI values, water bodies typically have values greater than 0.5, while vegetation yields smaller values, facilitating accurate distinction. Built-up areas generally fall within the range of 0 to 0.2[18] .

A classification of NDWI values for drought estimation according to V. Shashikant, 2021 can be found in Table 3.

Table 3: NDWI thresholds for drought estimation.

NDWI values	Drought category
-1.0 to 0.2	Extreme drought
0.2 to 0.3	Moderate drought

[16] https://www.unccd.int/sites/default/files/inline-files/04_Vegetation_index.pdf
[17] Agriwatch, https://www.agriwatch.nl/knowledge-center/vegetation-condition-index-vci
[18] https://eos.com/make-an-analysis/ndwi/

0.3 to 0.4	Mild drought
0.4 to 0.5	Moderate Moisture content
0.5 to 1.0	Very High Moisture content

According to literature NDWI values respond more quickly to drought conditions than NDVI and the correlation between the two indices highly depends on soil and vegetation cover type (Patil P., et al., 2024).

A combination of NDVI and NDWI has led to the development of another index, the Normalized Difference Drought Index (NDDI). This index measures the drought severity using the NDVI and NDWI of the observation period.

The mathematic formula for NDDI is:

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI}$$

NDDI index effectively captures both vegetation health and moisture levels. Recent scientific research has focused on categorizing NDDI values into drought categories as shown in the Table 4. (P. Patil et al., 2024):

Table 4: NDDI values in drought categories.

NDDI values	Drought category
Less than -1.0	Water Body
-1 to 0.2	No Drought
0.2 to 0.3	Mild drought
0.3 to 0.4	Moderate Drought
0.4 to 0.5	Severe Drought
0.5 to 1.0	Extreme Drought
More than 1.0	Unclassified

Another combined remote sensing-based index for drought monitoring is the Normalized Moisture Index (NMI), derived from the combination of NDVI and NDWI (Jae- Dong Jang et al., 2006). The NMI index is negatively correlated with surface temperature. The mathematic formula for NMI index is:

$$NMI = NDVI + NDWI$$

Land Surface Water Index, LSWI, is another water - related vegetation index which is used to assess vegetation condition and monitor/ drought. Like NDWI, LSWI is an effective indicator for monitoring the water content representing the total water content in vegetation (Christian et al., 2022). Many studies reported that water related vegetation indices are more sensitive to drought than greenness-based indices (Bajgain R. et al., 2015, Chandrasekar et al., 2010).

The mathematic formula for LSWI is:

$$LSWI = \frac{(NIR - SWIR)}{(NIR + SWIR)}$$

An important feature of the LSWI index is its ability to capture the temporal and spatial evolution of flash droughts, providing early warnings for drought development (Christian et al., 2022). Furthermore, LSWI can be used to extract cropland irrigation information by comparing the LSWI of cropland pixels to adjacent forest pixels with similar vegetation index values. This method has shown good accuracy in mapping irrigated regions and monitoring yearly changes (Kunlun Xiang et al., 2020).

In conclusion, LSWI is a valuable tool in monitoring drought and providing early warnings of drought events. Its high sensitivity makes it particularly accurate in drought assessment.

3.2.2 Meteorological drought indices (SPEI, VPD, KEETCH-BYRAM)

Meteorological drought indices are essential tools for assessing and monitoring drought conditions. The most commonly used indices include Standardized Precipitation Evapotranspiration Index (SPEI) which provide drought severity across different time scales, the Vapor Pressure Deficit (VPD) which reflects plant stress and water availability and the Keetch-Byram Drought Index (KBDI) which primarily measures soil dryness. these indices are vital for understanding and monitoring drought impacts and for supporting adaptation strategies.

3.2.2.1 SPEI Standardized Precipitation Evapotranspiration Index

The Standardized Precipitation Evapotranspiration Index (SPEI) is a meteorologic index for monitoring drought. Its importance lies in the data required for its calculation. Specific climatic parameters such as precipitation and evapotranspiration are key components for drought assessment. SPEI can be calculated for different time scales or seasons depending on the period of drought being analysed. According to literature, SPEI combines the sensitivity of PDSI to changes on evaporation demand with the simplicity of calculation and the multitemporal flexibility of the SPI (Standardized Precipitation Index) which only focuses on precipitation anomalies. As a result, SPEI is more comprehensive and better reflects the impacts of temperature and climate change on water availability. While the Standardized Precipitation Index (SPI) is recommended by the World Meteorological Organization (WMO) for monitoring meteorological drought, other indices such as the Palmer Drought Severity Index (PDSI) (Palmer 1965), the Standardized Precipitation Evapotranspiration Index (SPEI) are also widely used (Vicente-Serrano et al., 2010).

SPEI is calculated based on the difference between precipitation and potential evapotranspiration (PET) also known as the Climatic Water Balance and defined below:

$$Di = Pi - PETi$$

, where *i* is the selected period for which is calculated.

By applying various mathematical methods based on the requirements of calculation of SPEI, the final formula is:

$$SPEI = W - [(Co + C_1W + C_2W^2)/\ (1 + d_1W + d_2\ W^2 + \ d_3W^2]$$

Where $W = \sqrt{-2\ln(P)}$ for $P \leq 0.5$.

A detailed analysis of the SPEI calculation is provided by Vicente S. and Sergio M., 2024.

SPEI can be calculated for a wide range of timescales from 1 to 48 months, depending on the need. Additionally, SPEI adopts the same dry and wet categories used by SPI. The corresponding drought categories are provided in the Table 5 (Mishra A. K. et al.,2009).

Table 5: SPEI values for drought categories.

Categories	SPEI values
Extreme drought	< -2.00
Severe drought	-1.99 to -1.50
Moderate drought	-1.49 to -1.00
Near normal	-0.99 to 0.99
Moderately wet	1.00 to 1.49
Severely wet	1.50 to 1.99
Extremely wet	> 2.00

Additionally, SPEI can be calculated for different timescales and various types of drought. Several studies indicate that timescales 1 to 3 months ($SPEI_1$ to $SPEI_3$) correspond to meteorological drought, 3 to 6 months to agricultural drought and 6 to 12 to hydrological drought and periods longer than 12 months to socio-economic drought.

On the key advantages of SPEI is its ability to be calculated for both historical and projected climate periods depending on the input datasets. As for historical data, reanalysis datasets such as ERA5can be used to calculate historic and near real time drought conditions. Vicente-Serrano et al. 2022 generate a global drought dataset based on the SPEI using ERA5 reanalysis dataset which maintained in near real time and updated weekly. This dataset, ERA5 SPEI dataset, is available online [19]and offer the opportunity to the user to get data only in the area of interest.

SPEI can be calculated for future climate conditions leveraging up-to-dated Regional Climate Models (RCMs). These models incorporate precipitation and evapotranspiration data under different Representative Concentration Pathways (RCPs) (Politi N., et al. (2022); Kostopoulou and Giannakopoulos (2023)).

3.2.2.2 VPD Vapor Pressure Deficit

The Vapor Pressure Deficit (VPD)[20] plays a critical role in controlling the land-atmosphere exchange of water and CO2. As drought monitoring becomes increasingly important in the context of climate change VPD has emerged as a vital metric. It measures the difference between the actual moisture content in the air and the amount it could hold at full saturation (100% humidity). During the last years VPD has often been referred to as air dryness. Therefore, the VPD is a crucial meteorological parameter with significant implications for agriculture as it notable affects plant growth and can lead to reduced yields and lower quality production (Restaino et al., 2016). It also influences forest mortality and wildfire occurrence as

[19] https://global-drought-crops.csic.es/
[20] VPD, https://vpdrought.wsl.ch/en/what-is-vpd/

high VPD values are associated with dry conditions and increased risk of wildfires (Park W. et al., 2013, Seager et al., 2015, Balch et al., 2022). This makes VPDa a valuable tool for wildfire prediction and management. Beyond these impacts, VPD affects biodiversity and ecosystem functioning underscoring its broader ecological significance.

The computation of VPD involves the difference between the saturation vapor pressure (e_s) and the actual vapor pressure (e_a).The mathematical formulas are as follows:

$$VPD = e_s - e_a \qquad \text{or} \qquad VPD = e_s - (RH * e_s/100)$$

It is important to note that VPD is a direct measure of the atmospheric desiccation strength.

The Swiss project named "VPDrought"[21], running from 2022 to 2028, is applying a scale planning approach to disentangle the processes affected by atmospheric and soil drought in a mature Scots pine forest.

VPD can be calculated using various sources of atmospheric data. These include Reanalysis datasets like ERA5, a state-of-the-art global reanalysis dataset and offers historical meteorological variables with important temporal and spatial resolution. Using ERA5 data, VPD can be calculated for the past and current climatic conditions. Additionally, weather forecast models can be employed for the short-term predictions offering hourly or daily estimations. This approach is particularly valuable for agriculture, irrigation scheduling and ecosystem management. Finally future climate projections can provide estimations of temperature and humidity under different emission pathways (RCPs) supporting efforts to enhance climate change adaptation strategies.

3.2.2.3 KBDI Keetch – Byram drought index

The Keetch-Byram Drought Index (KBDI) is a continuous reference scale used to estimate the dryness of the soil and duff layers. It represents the net effect of evapotranspiration and precipitation in producing a cumulative moisture deficiency in deep duff or upper soil layers. The index value increases each day without rain with the rate of increase depending on the daily high temperature and decreases when it rains. Thus, the KBDI primary reflects recent rainfall patterns and serves as a measure of meteorological drought indicating water gain or loss within the soil.

Keetch-Byram drought (KBDI) index is the most known and widely used drought index for fire potential assessment and fire management, agricultural drought monitoring and climate research. Originally designed by Keetch and Byram in 1968 specifically for assessing fire potential, it remains a critical tool for drought monitoring in national weather forecast and wildfire prevention. The Wildland Fire Assessment System (WFAS) of United States Forest Service[22] provides daily fire danger maps using KBDI as one of its indices. Since 2019 EFFIS has also incorporated KBDI along with other relevant indices such as FWI and computed using numerical weather predictions from the ECMWF model.

Interpreting KBDI values shows that higher values are associated with conditions favourable for wildfire occurrence, although wildfire risk depends on a combination of multiple weather

[21] https://www.wsl.ch/en/projects/vpdrought/
[22] https://www.wfas.net/index.php?Itemid=487&id=86&option=com_content&view=article

factors. The KBDI scale ranges from **0** (no moisture deficit) to **800** (extreme drought), with the range determined by assuming that 20 cm of moisture is present in saturated soil, readily available to vegetation (Keetch and Byram, 1968). KBDI is widely used globally for drought monitoring in national weather forecasting, wildfire prevention, and agriculture, particularly in regions with rain-fed crops.

Table 6: KBDI drought classes

Indices values	KBDI drought classes
0 – 200	Low fire danger, soil moisture content is high, early spring season
200 – 400	Moderate fire danger; early steps of drought, late spring season
400 – 600	High fire danger; significant drying of organic materials and vegetation, summer and early fall
600 - 800	Extreme fire danger; severe drought conditions with the highest potential for fire spread

For the calculation of the KBDI daily precipitation, both the daily maximum temperature and the average annual total precipitation are required as inputs. Thus, the mathematical formula for KBDI is as follow:

$$Qt = Q_{t-1} + dQ - dP$$

As for *dQ,* it is estimated via

$$dQ = \frac{10^{-3}(800 - Q_{t-1})(0.968e^{0.0486T} - 8.3)dt}{1 + 10.88e^{-0.0441R}}$$

, where *Qt* is KBDi value for the current day, Q_{t-1} is the KBDI value of the previous day, *dQ* is the daily adjustment and *dP* is the daily precipitation (Gannon and Steinberg, 2021).

From the above, it is evident that weather forecast data is essential for short-term prediction. Additionally, KBDI is well-suited for evaluating future wildfire potential, as it can be calculated using RCM projections of temperature and precipitation. Finally, KBDI is important for assessing fire danger and monitoring drought in agricultural areas.

3.2.3 Soil Moisture

Soil moisture is an important drought indicator since it reflects the amount of water available to plants and ecosystems. Long-term declines in soil moisture impair agricultural output, natural vegetation, and water resource management, making it an important criterion for measuring and managing drought effects. There are different ways to estimate the soil moisture from satellite observations depending on the instruments onboard the satellite. The sensors can be broadly distinguished into three categories: a) passive, b) active, and c) active-passive.

Passive microwave radiometer instruments observe the long-wave electromagnetic radiation emitted by the Earth's surface. Radiometers measure the brightness temperature which is a function of the energy emitted by Earth's surface. The advantage of microwave radiometers is that they are not affected by day light, weather or vegetation, which means that they are characterized by high sensitivity to soil moisture. This explains the increased accuracy compared to other methods. On the other hand, passive radiometers have low resolution (30-40 Km). Two products are the most widely used for soil moisture monitoring, the ESA's Soil

Moisture and Ocean Salinity (SMOS), and the NASA's Soil Moisture Active Passive (SMAP). SMOS uses a microwave interferometric radiometer with three deployable arms (at an 120° angle from each other) that carry 69 receivers in total. Each receiver observes the brightness temperature at a different angle and an algorithm combines the 69 estimates into one soil moisture product at 35-50 km horizontal resolution. On the other hand, SMAP measures the brightness temperature at only one angle with a resolution of 36 km. The advantage of this instrument is that it rotates several times per minute, acquiring several images at 36 Km resolution but for slightly different positions (shifted by 9 Km in the X and Y direction). An algorithm combines these images and generates soil moisture products at 9 km pixel sampling. Both, SMOS and SMAP, have a temporal sampling of 2-3 days depending on the geographical location.

On the other hand, **active radar (SAR)** instruments emit radiation towards the Earth's surface and receive the reflected echo (backscatter), and more specifically the co-polarized backscattering coefficient (σ_{VV}^{0}) has been shown to carry the soil moisture signal. Unlike the microwave radiometers, they can operate at a much higher resolution (1-100 meters), however, with a reduced accuracy. This is because, unlike the microwave radiometers, the radar reflected beam is more sensitive to surface characteristics, roughness and vegetation (plant water content, density, size and crop row orientation), which are highly variable, and less to soil moisture. The challenge with SAR instruments is the development of algorithms to isolate the SSM signal embedded in the σ_{VV}^{0} measurements. This sensitivity to other than soil moisture surface characteristics requires a preprocessing of the SAR image to augment the soil moisture signal. This is achieved by using classification methods that distinguish agricultural lands from any other type of land (forests, urban etc.), and resampling the SAR resolution to a lower resolution (500-1000m).

It should be noted, that initially SMAP was designed to implement such technology by combining a passive microwave radiometer and a SAR active radar in order to generate high resolution and accurate soil moisture maps, however the SAR instrument experienced failure right after launch.

Active-passive (signal of opportunity) methods rely on GPS signals that are emitted by a satellite and are received by another satellite as it is reflected by Earth's surface (GNSS reflectometry). The GNSS signal is recorded in the form of Delay-Doppler Maps (DDMs). The characteristics of the DDMs are affected by soil moisture as well as other surface characteristics (roughness and vegetation) as in the case of the SAR radar signal.

The most common product in this case is the Cyclone Global Navigation Satellite System (CYGNSS), with a spatial resolution of a few kilometres. The only disadvantage of this product is its geographical constraint, it only provides data only for the tropics, from 37°S to 37°N. This disadvantage, however, will be set off by the next generation ESA mission (HydroGNSS), which will provide data for the whole globe.

3.3 Heatwave Indices

A heatwave[23], according to WHO, can be defined as a period where local excess heat accumulates over a sequence of unusually hot days and nights. The IPCC[24] states that with further global warming, we can expect an increase in the intensity, frequency and duration of heatwaves. Heatwaves aggravate a variety of dangers, including higher human mortality, drought and water quality, wildfires and smoke, power outages, and crop losses. Moreover, vulnerability to heat is influenced by both physiological parameters, such as age and health state, and exposure factors such as occupation and socio-economic conditions.

Several indices have been defined to monitor heat waves. These indices are usually based on meteorological variables and are designed to represent the comfort level for different sectors of society (health, agriculture, and electricity supply) during heat waves. Because of this, it is difficult to find a universal index, that will represent the comfort level during heat waves across domains. Some of the most common meteorological variables used are the minimum temperature (T_{min}), maximum temperature (T_{max}), or apparent temperature (T_a) also known as the "heat index".

The maximum temperature (T_{max}) is one of the most common choices to define heat waves. To calculate it, we need daily T_{max} values spanning several decades, ideally at least 30 years. First, it is necessary to estimate the 90th percentile of T_{max} for each calendar day of the year. These 90th percentile values form a curve with a strong seasonal cycle (higher values in the summer and lower in the winter). The number of years used will influence the smoothness of this seasonal cycle. Inevitably, the resulting seasonal cycle will have some noise, which is then smoothed using a 15-day moving window. A heatwave is detected when T_{max} exceeds the 90th percentile for at least 3 consecutive days (Perkins and Alexander (2013); Perkins-Kirkpatrick and Lewis (2020)). A similar analysis can be applied to T_{min}, where a day is defined as a heat wave day if T_{min} exceeds the 90th percentile of the minimum temperature for a number of consecutive days.

Additionally, there are several heat wave characteristics, i.e., intensity, duration, frequency, cumulative heat. Intensity is the average T_{max} for all heat wave days, frequency is defined as the sum of all heatwave days, and duration is defined as the duration of the longest heat wave event. To compute the cumulative heat first, we estimate the temperature anomaly (defined as the deviation of the daily temperature from the calendar-day 90th percentile for all days experiencing a heat wave during a season), and next we add these anomalies.

Another heat wave definition was introduced by Perkins and Alexander (2013), the Excess Heat Factor (EHF). It is based on two excess heat indices EHI (accl.) and EHI (sig.):

$$EHI\ (accl.) = \frac{T_i + T_{i-1} + T_{i-2}}{3} - \frac{T_{i-3} + \cdots + T_{i-32})}{30}$$

$$EHI\ (sig.) = \frac{T_i + T_{i-1} + T_{i-2}}{3} - T_{95}$$

[23] https://wmo.int/topics/heatwave
[24] https://www.ipcc.ch/report/ar6/syr/

Where T_i is the average daily temperature (average of T_{max} and T_{min}) for day *i* and T_{95} is the calendar day 95th percentile for the period under investigation. The first index represents anomaly of a 3-day window average temperature against the average temperature over the last 30 days. The second index represents the anomaly of the same window against the 95th percentile. The EHF index is defined as

$$EHF = \max[1, \quad EHI\ (accl.)] \times EHI\ (sig.)$$

A heat wave is detected if this index acquires positive values for at least 3 consecutive days.

Forecasting heatwaves is an important part of climate resilience since it allows for early warnings and mitigation steps for protecting sensitive populations and ecosystems. To project future heatwave activity, we use General Circulation Models (GCMs). The ocean and the atmosphere are fluids, and their behaviour is governed by physical laws, which are represented by a system of mathematical equations. A GCM is a computer program that solves these equations and generates meteorological variables as it moves the system forward in time. To forecast heat waves, we need to acquire the necessary simulated data (T_{max} or T_{min}) and apply exactly the same methodology as with observed data.

3.4 Wildfire susceptibility and vulnerability mapping

3.4.1 Fire susceptibility Mapping

Climate change is increasing the frequency and severity of fire events due to changing climate conditions. Rising temperatures in recent years have prolonged drought periods, which in turn impact the frequency and the severity of fire. According to the Joint Research Centre (JRC), higher temperatures and extended droughts lead to longer fire danger periods. This trend is especially concerning for populations in wildland-urban interface areas, who are exposed to fire risks for more days. Not only does this increase the duration of exposure, but it also raises the number of people at risk.

Fire is a natural hazard that occurs globally and especially in the Mediterranean region due its forest, land, and climatic conditions. Unlike drought hazard mapping which is focused mainly on indices derived from meteorological or remote sensing data, fire hazard mapping is much more complex. Fire hazard or fire risk modelling is a multifaceted process and requires integration of diverse datasets beyond meteorological indices.

Many studies, especially in the Mediterranean area, focus on fire hazard, considering a wide range of parameters. These include topography (elevation, slope, aspect), meteorological factors, human activities or general human presence, proximity to artificial structures, vegetation type, fuel type, as well as biomass density. Additionally, other research focuses on using remote sensing data and techniques in creating susceptibility maps. According to the literature, integrating fire events with factors like NDVI, topographic features, meteorological variables, and human presence is the most effective approach for developing fire hazard maps.

An integrated analysis of fire hazard mapping using spatiotemporal and geostatistical methods is crucial. According to Sakellariou et al. (2020), combining multiple data sources allows for a better understanding of fire hazard dynamics under climate change and helps develop appropriate measures for the most vulnerable areas. In recent years, a considerable amount of research has been conducted on this topic, with fire risk and susceptibility maps being produced using spatial techniques and statistical models (Gheshlaghi et al., 2020; Busico et al., 2019).

Key Parameters for Fire Hazard Mapping:

- **Topography:** Elevation, slope, and aspect are critical factors for fire hazard mapping. To calculate these, a Digital Elevation Model (DEM) or Digital Terrain Model (DTM) is used. Various types of DEMs are available for topographic analysis, which help produce statistics on area characteristics, such as the percentage of lowland versus upland. Once DEM is used to calculate slope and aspect, these factors are considered stable and require only one-time calculation.
- **Vegetation:** Vegetation is a key parameter in fire hazard assessment, especially in forested areas. The Corine Land Cover (CLC) dataset, provided by Copernicus, is essential for identifying fire-prone areas by categorizing land cover types. CLC offers open access data and is an essential tool for fire hazard assessment as it helps identify areas with high susceptibility to fires based on vegetation types and their flammability. The distribution and density of vegetation are crucial for understanding fire behaviour. Vulnerability to fire in an area is determined by climate and ecological factors and from land use changes (de Torres Curth et al., 2012). As for vegetation, additional parameters which take into consideration the current situation of vegetation is the NDVI (Normalized Difference Vegetation Index) index, which can be calculated from remote sensing data, and allows a more near real time estimation of vegetation health. Moreover, NDMI (Normalized Difference Moisture Index) can also be calculated from remote sensing data and provides information about moisture levels of vegetation. These indices are critical for understanding how vegetation moisture influences fire spread.
- **Human Presence:** Human activity also affects fire events. Proximity to road networks is one of the key factors, as transport infrastructures such as paths, trails, and railways may also be important in fire risk assessment (ANDALUSIA). The development of areas adjacent to roads and buffer zones around these infrastructures is vital in wildfire hazard mapping. Moreover, the proximity of settlements or other human-made structures to forests increases the likelihood of human-caused ignitions. Increasing tourism for recreational activities further elevates fire risks. Development of zones adjacent to human settlements network is another important parameter and especially at Wildland Urban Interface areas. As for road network and settlements data, Open Street map (OSM) is the most usable open dataset (Stefanidou A. et al., 2019; Sakellariou et al., 2020).
- **Buffer Zones:** Buffer zones are important for delineating areas of risk near roads and human settlements. In Mediterranean ecosystems, buffer zones are often classified in intervals of 100 meters, resulting in five classes, while some studies use a broader 500-meter threshold. Experts use local knowledge to define these zones effectively (Stefanidou A. et al., 2019; Sakellariou et al., 2020).

Other factors, such as distance from drainage networks, topographic wetness index, rainfall, temperature, and wind speed, are also valuable for fire risk assessment (Abdo et al., 2022). Other variables, including Land Surface Temperature and proximity to crops, further contribute to understanding fire susceptibility (Ghosh et al., 2018). These parameters, when combined, offer a comprehensive view of fire risk across regions. In Bentekhici N., et al. (2020), the authors propose a fire risk mapping risk approach that clusters parameters in the following way: a geomorphology index (slope, aspect and elevation), human index (proximity to forest, roads and urbanization), and vegetation index (NDVI application as biomass density).

After selecting relevant parameters, Multicriteria Analysis (MCA) is applied to evaluate and integrate different, often conflicting, criteria in hazard assessment. MCA assigns weights to each parameter based on its significance in determining fire risk. The Analytical Hierarchy Process (AHP) is a commonly used method in MCA, where a decision-maker uses pairwise comparisons to prioritize criteria and determine relative weights. Through per pair analysis a decision maker must prioritize all criteria by defining relative position of one in relation to all other. According to literature, the relative weights of each individual criterion are determined by calibrating the intensity of each criterion in relation to others.

Fire risk estimation is especially crucial in Mediterranean regions, where the combination of environmental, topographic, climatic, and anthropogenic factors plays a significant role in fire risk. By synthesizing diverse data sources through spatial and temporal analysis, authorities can identify high-risk areas and develop effective adaptation and mitigation strategies to protect vulnerable regions.

This methodology for fire hazard mapping at a regional level integrates open-access geospatial and satellite data, incorporating both static and dynamic factors into risk assessments. This approach is valuable as enables update of dynamic data, monthly or weekly, depending on region's priorities. This frequent update offers the ability to response in changing conditions and enhance mitigation strategies. All the factors can be integrated with other open data, such as population data, tourism infrastructure, and critical facilities (e.g., hospitals, schools), to identify vulnerable areas and prioritize protective measures.

3.4.2 Satellite-Derived Indices for Burned Area detection and Impact Assessment

Satellite remote sensing has significant advantages and is an essential tool for burn scars. The most significant advantage is the extraction of burned area and the assessment of ecological changes in a short imaging period and over large spatial extents. This capability is critical for evaluating burn severity and for conducting post-fire assessments, which are essential for implementing effective mitigation strategies.

Wildfires are among the most devastating natural disasters, and satellite imaging and Earth observation play critical roles in their investigation and management. High-resolution data from platforms like Sentinel-2 and Landsat 8 allow researchers to evaluate pre- and post-fire landscapes with amazing accuracy, making it easier to identify and map burned regions. According to Zhang et al. (2023), the index threshold method is commonly utilized for detecting burned areas. This method involves assessing the spectral reflectance curves of burned areas, selecting appropriate wavebands to establish indices, and applying thresholds to efficiently differentiate burned areas.

There are several burned area mapping indices including the Burned Area Index (BAI) (Chuvieco & Palacios, 2010), Normalized Burn Ratio (NBR) (Cocke, et al., 2005), and Mid-InfraRed Bispectral Index (MIRBI) (Smith et al., 2007). Furthermore, remote sensing vegetation indices such as the Normalized Difference Vegetation Index (NDVI), Soil-Adjusted Vegetation Index (SAVI), Global Environment Monitoring Index (GEMI), and Enhanced Vegetation Index (EVI) (Sun et al., 2019) are commonly used to map burned areas, providing critical insights into the effects of fire on vegetation.

All of these studies demonstrate that spectral indices, particularly those utilizing NIR, RED, and SWIR bands from satellite photos, are the most effective and powerful tools for detecting and monitoring burned areas. As a result, the remote sensing approach improves our

understanding of fire impact and promotes better management. According to literature, the most useful indices for mapping fire-affected areas are NDVI and dNDVI, NBR and dNBR, RdNBR and NDBI.

Normalized Difference Vegetation Index (NDVI)

NDVI's (Normalized Difference Vegetation Index) mathematical formula has already been mentioned in relation to drought indices. However, the range of values differs compared to drought monitoring. In fire mapping NDVI is used to detect vegetation loss after a wildfire, as burned areas exhibit significantly lower NDVI due to the elimination of forest cover. NDVI could ensure timely action in high risk or heavily impacted areas. Table 7 depicts the NDVI values used for the delineation of burned areas.

Table 7: NDVI values for delineation of burned areas.

NDVI values	Category
-1.0	Cloud, water
-0.1 - 0.1	Bare soil
0.2 – 0.4	Sparse vegetation
>0.5 -1	Dense vegetation

Differenced Normalized Vegetation Index (dNDVI)

The Differenced Normalised Vegetation Index (dNDVI) is the difference in NDVI between images pre and post the fire event. Pre/post-fire difference indices have the advantage of allowing for obvious discrimination between unburned sparsely vegetated areas and charred areas, which is difficult to achieve using mono-temporal images. The differenced Normalized Differenced Vegetation Index (dNDVI) improves the ability to measure absolute changes by using imagery captured in pre-fire and post-fire dates.

$$dNDVI = NDVIpre - NDVIpost$$

In dNDVI, the reflectance scale is reversed, and high raster values imply a higher shift from pre-fire values, whilst values close to zero suggest insignificant fire damage (see Table 8).

Table 8. dNDVI values for Sentinel-2 images (Source: Fire Mapping chapter, working paper ver. 1.0, 2018)

dNDVI per fire severity class	Class description	Colour
dNDVI <=0.3	Low severity	
0.3 > dNDVI <= 0.55	Moderate severity	
dNDVI > 0.55	High severity	

Normalized Burn Ratio (NBR)

The **NBR** index is considered the most appropriate choice to detect burnt areas. NBR values range between -1 to 1. Healthy green vegetation is expected to have a high NBR value, while burned vegetation a low value (Abdikan S. et al., 2022).

Below, the NBR math type using Sentinel 2 images is shown:

$$NBR = (B_{8A} - B_{12})/(B_{8A} + B_{12})$$

Furthermore, the dNBR which is the differenced Normalized Burn Ratio is calculated in a similar way to NDVI. Thus, it is estimated as the difference between pre and post fire situation and its value is more relevant than NBR value alone in determining the burnt area. dNBR can be used to define the burn severity of fire. Classification of burn severity values given in Table 9.

$$dNBR = NBRpre - NBRpost$$

Table 9: USGS proposed classification table for interpretation of burn severity values (Source: Keeley, J.E., 2009).

Severity Level	dNBR Range (scaled by 10^3)	dNBR Range (not scaled)
Enhanced Regrowth, high (post-fire)	-500 to -251	-0.500 to -0.251
Enhanced Regrowth, low (post-fire)	-250 to -101	-0.250 to -0.101
Unburned	-100 to +99	-0.100 to +0.99
Low Severity	+100 to +269	+0.100 to +0.269
Moderate-low Severity	+270 to +439	+0.270 to +0.439
Miderate-high Severity	+440 to +659	+0.440 to +0.659
High Severity	+660 to +1300	+0.660 to +1.300

Relativized Burn Ratio (RBR)

RBR Index is used to identify recently burned areas and distinguish bare ground from other non-vegetated areas. RBR is used for reliability when the biomass before burning is low or diverse. A high RBR value indicates healthy vegetation, whereas a low value indicates bare soil or recently burned areas (Da Simone et al., 2020). The index has been divided into three burn severity classes (Moderate-Low, Moderate-High and High severity). The computation for RBR is:

$$RBR = dNBR/(NBRprefire + 1001)$$

RBR displayed improved average classification accuracy across numerous fires, indicating its robustness in assessing burn severity across various ecosystems.

Relative Differenced Normalized Burn Ratio (RdNBR)

The dNBR algorithm evaluates the absolute difference between pre- and post-fire images, while the **RdNBR** algorithm determines burn severity using pre-fire reflectance and calculates the relative change caused by fire. The index is mostly sensitive in areas with low pre-fire vegetation cover (Parks S.A. et al., 2014).

$$RdNBR = {dNBR}/{|(NBRprefire)|^{0.5}}$$

RdNBR is also used to assess potential limitations of dNBR in characterizing fire severity on low biomass sites and potentially enhance inter-fire comparability of the results at a broader ecological scale (https://www.mtbs.gov/mapping-methods). This normalization accounts for variations in vegetation density and other pre-existing factors, making it more reliable for comparing burn severity across diverse landscapes.

Normalized Difference Built-up Index (NDBI)

The Normalized Difference Built-up Index NDBI is used to detect changes in built-up areas pre and post fire. It's primary application in fire events is the assessment of structural damage in urban and peri-urban areas. Additionally, it helps map accessibility issues in roads and

transportation network. NDBI can provide insights on the extent of urbanization and land cover change. NDBI values range from − 1 to + 1, with higher values indicating more impervious surfaces, while lower values represent non-built-up areas such as vegetation and water (Alademomi, A. S et al., 2022). The formula is:

$$NDBI = \frac{B11 - B8}{B11 + B8}$$

Integrating NDBI may improve assessments in urban areas where fires impact both vegetation and built environments. Furthermore, when combined with other vegetation indices, it can help in evaluation of urban areas recovery in comparison to natural.

In summary, satellite remote sensing is an important tool for detecting burned areas, assessing and monitoring the ecological impact of fire over large areas and in short time. Key spectral indices as Normalized Difference Vegetation Index (NDVI) and dNDVI are useful for monitoring vegetation loss whilst Normalized Difference Built-up (NDBI) is important in evaluating fire impact in urban and peri-urban areas. Finally, Normalized Burn Ratio (NBR) and Differenced Normalized Burn Ratio (dNBR) and Relativized Burn Ratio (RdNBR) have effective results in determining burn severity.

Integration of multiple indices offer multi-faceted analysis on natural and urban landscapes. Additionally, a comprehensive impact assessment across diverse land cover types and important ecological and urban recovery insights. Moreover, the indices' synergy enables the development of effective mitigation strategies such prioritization of areas for reforestation, urban reconstruction and erosion control based on the severity. This approach ensures a robust and reliable fire impact assessment and support effective recovery planning.

3.5 Flood Susceptibility and Exposure Mapping

3.5.1 Flood susceptibility and exposure Mapping

Climate change has been a significant driver of many hazards over the past decade. Extreme weather events, such as heatwaves, rising global temperatures, reduced precipitation, and declining soil moisture, have led to drier land and prolonged drought periods. However, extreme events are not limited to temperature alone; changes in precipitation have also played a critical role. Increased rainfall intensity in recent years has resulted in floods, which can take different forms, such as fluvial floods, pluvial floods, or flash floods. According to the International Glossary of Hydrology, a flash flood is defined as a flood of short duration with a relatively high peak discharge (WMO).

Urban areas and impervious land are particularly vulnerable to flash floods due to the limited infiltration of water through impermeable surfaces, with other factors exacerbating the risk. Deforestation, wildfires, urbanization, climate change, and human activities are key contributors to flash flood events as a consequence of extreme rainfall. Flash floods often lead to agricultural damage, landslides, loss of human and animal lives, and the destruction of infrastructure and communication networks (Diakakis et al., 2016).

The assessment and mapping of flash flood events are crucial for urban planning, emergency response, policy-making, and enhancing community resilience. The literature provides methodologies for flood mapping, with a focus on identifying areas at risk. Several studies highlight the importance of the Flash Flood Potential Index (FFPI) in mapping vulnerability. The FFPI integrates topographical, hydrological, soil, and land use data, offering a comprehensive

analysis to predict areas prone to flash floods. Developed by Smith in 2003 at the Colorado Basin River Forecast Center, the FFPI was created in response to unsatisfactory results from previous predictions of torrential floods (Durlević U, et al., 2021). The mathematical formula for calculating FFPI is:

$$FFPI = \frac{M + S + L + V}{4}$$

, where *M* is terrain slope, *S* is soil type, *L* is the land use and *V* is vegetation.

FFPI calculation requires high-resolution datasets. For slope calculation, a Digital Elevation Model (DEM) is essential. High spatial resolution open data, such as the EU-DEM from Copernicus, is commonly used for this purpose. The mathematical formula for slope is:

$M = 10^{n/30}$, where *n* is a terrain slope expressed in %.

Steeper slopes result in faster runoff, thereby increasing the risk of flooding.

For soil type data, the FAO's[25] Soil Map of the World, including continent-specific maps, serves as an open-source resource when country-level data is unavailable. Soil type coefficients are derived based on their characteristics and supported by literature. Different soil textures significantly influence runoff behaviour.

The third parameter is land cover. For this, data from the CORINE Land Cover (CLC) inventory can be used, with classes assigned values according to their susceptibility to torrential floods. A table of land use coefficients provided by Durlević U, et al. (2021) identifies the most flood-susceptible terrains as non-irrigated agricultural areas and regions with sparse vegetation. Land use is a critical factor affecting water absorption and flow dynamics.

The final parameter is the vegetation density index, calculated using the Bare Soil Index (BSI). The BSI is derived from satellite imagery, specifically Sentinel-2 satellite images. The mathematical formula for BSI is:

$$BSI = \frac{(B11 + B4) - (B8 + B2)}{(B11 + B4) + (B8 + B2)}$$

Then, the V coefficient is calculated as:

$$V = 7.68 * In(BSI + 1) + 8$$

Recent scientific literature also incorporates the NDVI index to estimate vegetation parameters, noting that NDVI can effectively quantify flood extent across various flood events (Mohd Y. et al., 2023).

For the final calculation of the Flash Flood Potential Index (FFPI), each parameter is calculated, reclassified, and combined. Each parameter is assigned a score ranging from 1 to 10, where 1 indicates low probability and 10 indicates high probability of flash floods. A table (Table 10) provided by Mohd Y. et al. (2023) outlines FFPI parameter values based on slope, land use, vegetation cover, and soil type as follows:

[25] https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/

Table 10: FFPI values depend on each dataset (Source: Mohd Y. et al., 2023).

FFPI value	Slope (%)	Land Use / Land Cover	Vegetation cover (%)	Soil type
1	3 and below	Water	90-100	Water / Alluvial
2	6	Woody Wetlands	80-89	Sand
3	9	Evergreen forest	70-79	Sandy Loam
4	12	Mixed Forest	60-69	Silty Loam, Loamy Sand
5	15	Deciduous Forest	50-59	Silt / Organic matter
6	18	Pasture Hey, Cultivated areas	40-49	Loam
7	21	Developed – open space – Barren Land	30-39	Sandy Clay Loam, Silty Clay Loam
8	24	Developed low	20-29	Clay
9	27	Developed medium	10-19	Clay
10	30 and above	Developed heavy	0-9	Bed, Rock, Impervious land

The result is a number, as FFPI is a relatively static index. After calculating all coefficients, the FFPI index is determined, and its values can be visualized to represent areas at risk of flash floods. Over the years, some modifications have been made to FFPI components and their weights (Brewster, 2009; Kruzdlo & Ceru, 2010; Ceru, 2012). The most recent modification to the FFPI index includes the addition of two parameters: profile curvature, which affects the way water flows across terrain, and flow accumulation, which indicates potential flooding hotspots (Tincu R., et al., 2018).

Finally, the FFPI index serves as an early warning system for potential flash flood events, enabling communities to prepare and implement mitigation strategies.

Flash flood mapping plays a crucial role in mitigating the impacts of floods. Scientific literature highlights the value of hazard mapping, particularly methodologies for identifying high-risk areas. FFPI has been extensively applied across different geographical regions. Additionally, many other approaches have been implemented for flash flood hazard mapping, leveraging diverse datasets to provide a comprehensive understanding of risk.

These approaches consider various factors, including topography, hydrology, meteorology, geology, land use/land cover, soil characteristics, and drainage. Remote sensing indices like NDVI, NDWI, and BSI also contribute valuable information for flood analysis. The factors are often assigned weights based on statistical analysis, expert opinions, machine learning approaches, or their relative importance through Multi-Criteria Decision Analysis (MCDA).

Further research into flood risk mapping integrates a wide range of factors influencing an area's susceptibility to flooding. According to Elkhrachy (2015), key risk factors include rainfall, soil type, surface slope and roughness, hydrographic network density, vegetation, and proximity to streams. These factors can be extracted from open datasets such as satellite imagery and geographic information systems (GIS) and transformed into raster format. The Analytical Hierarchy Process (AHP) is then used to weight these factors based on their importance and combine them to display flood risk for a given area. This method provides a

cost-effective and comprehensive approach to flood risk assessment and is feasible for various geographic regions.

Several approaches and methods are available for determining areas at risk from different hazards. Integrating high-resolution satellite data with vector data and socioeconomic data in GIS platforms provides a multidimensional view of risks. This combination enhances a region's overall disaster resilience and improves mitigation measures to prevent disasters or minimize their adverse impacts.

3.5.2 Satellite-Derived Indices for Flood Area detection and Impact Assessment – Water Spectral Indices

Flood events, as previously mentioned, are among the most destructive natural hazards, comparable to fires. The damage to ecosystems, infrastructure, and human lives is extensive. Although the immediate effects of a flood event are significant, the aftermath is often worse: affected areas are covered with mud, drinking water becomes contaminated, and power outages occur, which can lead to the spread of diseases. Therefore, accurate flood impact assessment is crucial, and satellite missions can provide near-real-time data with little or no cost. This information is vital due to its timely availability.

Satellite images from the Copernicus program, such as Sentinel-1 (radar) and Sentinel-2 (optical), are commonly used to provide timely and accurate data. Rapid and efficient flood mapping is crucial in flood hazard and risk analysis. Due to their lower cloud penetration compared to optical sensors, radar systems like Sentinel-1 are often preferred for monitoring floods (Sheng et al., 2001). Optical sensors, however, are mainly used to determine flood extent, but they are limited by cloud cover during the flood event, and their revisit cycle may be too long for real-time monitoring (Lin L. et al., 2016).

Sentinel-2, with its high-resolution multispectral imagery, provides invaluable data for flood impact analysis and enables detailed assessment of flood-affected areas. Satellite-derived indices are widely used for flood extent mapping and water distribution analysis. Vegetation indices like NDVI, dNDVI, and NDMI can be used to assess post-flood changes. Among these, the NDMI index, which will be described below, is particularly relevant.

Additionally, there are several indices that specifically highlight water bodies, such as the Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), and Normalized Difference Inundation Index (NDII).

Normalized Difference Moisture Index (NDMI)

The Normalized Difference Moisture Index (NDMI) is used to determine vegetation water content and is derived as the ratio between the NIR (B8) and SWIR (B11) bands

$$NDMI = \frac{NIR-SWIR}{NIR+SWIR}$$

The values of the NDMI range between -1 to 1. Higher values indicate more soil moisture (larger quantities of water). Moreover, dNDMI can be calculated as the difference between pre and post flood and serve as absolute change.

$$dNDMI = NDMIpre - NDMIpost$$

Normalized Difference Water Index (NDWI) - Modified Normalized Difference Water Index (MNDWI)

The Normalized Difference Water Index (NDWI) is designed to enhance the detection of water bodies and is widely used for water body analysis. It was initially developed to detect surface water and facilitate its measurement (McFeeters, S.K. et al., 1996). NDWI highlights water bodies and efficiently identifies areas inundated by floods, helping delineate flood extents (Stoyanova, 2023). However, it is sensitive to built-up areas and can sometimes overestimate water bodies. The equation for NDWI is as follows:

$$NDWI = \frac{Green\ (B3) - Nir(B8)}{Green\ (B3) + Nir\ (B8)}$$

NDWI values range from -1 to 1. According to McFeeters et al. (1996), a threshold value of 0 is commonly used: values greater than 0 indicate water, while values less than 0 represent non-water features. However, when NDWI is applied in regions with built-up land backgrounds, it often yields inaccurate results. The retrieved water information in such areas is frequently contaminated by noise from built-up land characteristics, which can contribute positively to the NDWI image (Xu, 2006). To address this limitation, the Modified Normalized Difference Water Index (MNDWI) was developed. It uses the SWIR band to improve the distinction between water and non-water features, particularly in areas with built-up land. The formula for MNDWI is:

$$MNDWI = \frac{Green - SWIR}{Green + SWIR}$$

The SWIR band effectively absorbs terrestrial surface reflections, suppressing non-water features and making MNDWI more reliable for delineating flooded areas.

Normalized Difference Inundation Index – NDII

The Normalized Difference Inundation Index (NDII) is a spectral index designed for mapping inundated areas to identify spectral differences between pre-flood and peak-flood conditions. According to Levin and Phinn (2022), the NIR reflectance decreases significantly after a flood event. The NDII leverages this change by comparing spectral differences in the NIR band before and during peak flooding. The formula for NDII is:

$$NDII = \frac{NIR_{pre-flood} - NIR_{peak-flood}}{NIR_{\ pre-flood} + NIR_{peak-flood}}$$

A threshold of NDII > 0.1 is commonly used to identify inundated areas. While NDII shows promise as an effective tool for flood mapping, its application in existing research is limited, given that it is a relatively new index. Nonetheless, NDII is particularly useful for understanding water distribution in flood-affected areas.

In conclusion, Sentinel-2 imagery provides significant capabilities for flood assessment and mitigation, particularly in the aftermath of flood events. Accurately identifying flood-inundated areas and mapping the extent of damaged regions are critical for effective post-disaster response. The integration of calculated indices such as NDWI and MNDWI with geospatial data offers stakeholders powerful tools for flood management.

Sentinel-2 data enhances the applicability of NDII due to its high spatial resolution (10 to 30 meters) and short temporal revisit time (5 days). These features make Sentinel-2 an invaluable

resource for analysing and monitoring inundated regions. Flood extents derived from Sentinel-2 imagery provide essential insights into hydraulic functioning and help delineate features of floodplain areas (Schumann et al., 2009). Additionally, the high resolution of Sentinel-2 imagery can be utilized to estimate soil moisture, a critical factor influencing flood forecasting models and predictions of flood characteristics (Grillakis et al., 2016).

For flood impact assessment, Sentinel-2 data enables the generation of Land Use/Land Cover (LULC) maps, which are crucial for understanding runoff dynamics, prioritizing rehabilitation efforts, and estimating economic losses (Zope et al., 2016; Kuntla, 2021). These datasets provide valuable information at various stages of flood management, ensuring a comprehensive and effective approach to mitigating flood impacts.

3.5.3 LISFLOOD Model by Copernicus EMS

LISFLOOD is a GIS-based distributed hydrological model created by the Joint Research Centre (JRC) of the European Commission. Designed to operate on large and transnational catchments, the model integrates various hydrological and meteorological data sources to simulate rainfall-runoff processes and river channel routing. LISFLOOD captures complex hydrological processes such as interception, infiltration, surface runoff, soil moisture dynamics, and groundwater flow, making it a robust tool for assessing water resource balance and predicting flood events. Embedded in the European Flood Awareness System (EFAS) of Copernicus EMS, LISFLOOD provides a crucial foundation for large-scale flood forecasting and flood impact studies across Europe (Van Der Knijff et al., 2010).

Purpose, Use, and Application

LISFLOOD's primary purpose is to enhance emergency management by forecasting flood events, assessing water availability, and supporting flood risk mitigation. It is used in both operational and research settings, providing critical insights into flood forecasting (Ramirez et al., 2016) and early warning. The model is applied in a variety of contexts, including:

- **Flood forecasting and real-time monitoring**: As part of the European Flood Awareness System (EFAS), LISFLOOD supports early flood warnings, which are essential for reducing disaster impacts by enabling timely evacuation and resource deployment.
- **Climate change impact assessments**: LISFLOOD is used to simulate future flood scenarios based on changing precipitation patterns and land use, helping to evaluate climate adaptation measures for flood-prone regions (Feyen et al., 2008).
- **Water resource management**: Beyond flood forecasting, LISFLOOD supports long-term planning for water resource allocation, especially in areas affected by droughts or variable precipitation patterns (De Roo et al., 2003).

Input Parameters

LISFLOOD requires a variety of spatial and temporal data inputs, allowing it to simulate complex hydrological processes. Key input parameters include:

- **Meteorological data**: Precipitation, temperature, evapotranspiration, and wind speed data are fundamental for estimating rainfall-runoff dynamics.
- **Topographical data**: Digital Elevation Models (DEMs) are used to understand surface flow paths, slope, and channel networks.

- **Soil and land cover information**: Soil texture, depth, and land use (e.g., forests, urban areas) inform the model's calculations of infiltration, evapotranspiration, and surface runoff.
- **Streamflow and discharge data**: These are used for calibration to align LISFLOOD's predictions with observed data, enhancing accuracy for operational forecasting (Van Der Knijff et al., 2010).

Output

LISFLOOD produces comprehensive hydrological outputs including river discharge, flood extent, soil moisture, and evapotranspiration, which can be visualized as time series data, spatial maps, or custom reports. Outputs are available in flexible formats (e.g., NetCDF, raster maps) to support a range of user needs from high-resolution flood depth maps for urban planning to large-scale hydrological overviews for water resource allocation. This versatility allows LISFLOOD outputs to be integrated into other systems or used directly in GIS platforms to analyse spatial flood risks (De Roo et al., 2000).

Focus on Historic, Real-Time, and Forecast Data

The model supports simulations based on historic data, real-time inputs, and forward-looking scenarios. Historical flood simulations enable model validation and retrospective flood assessments, while real-time data integration facilitates continuous flood monitoring across Europe. Additionally, the model's forecasting capabilities—particularly within EFAS—enhance flood preparedness by offering projections of flood risk up to several days in advance, allowing regional authorities to make proactive decisions during extreme weather events (Arnal et al., 2018).

Image from the Model Output

An example LISFLOOD output visualization includes flood inundation maps showing areas at risk or soil moisture maps indicating water distribution across a basin. These graphical outputs are valuable for decision-makers in assessing the spatial extent and depth of potential flood events, as well as for mapping past floods to inform urban and regional planning.

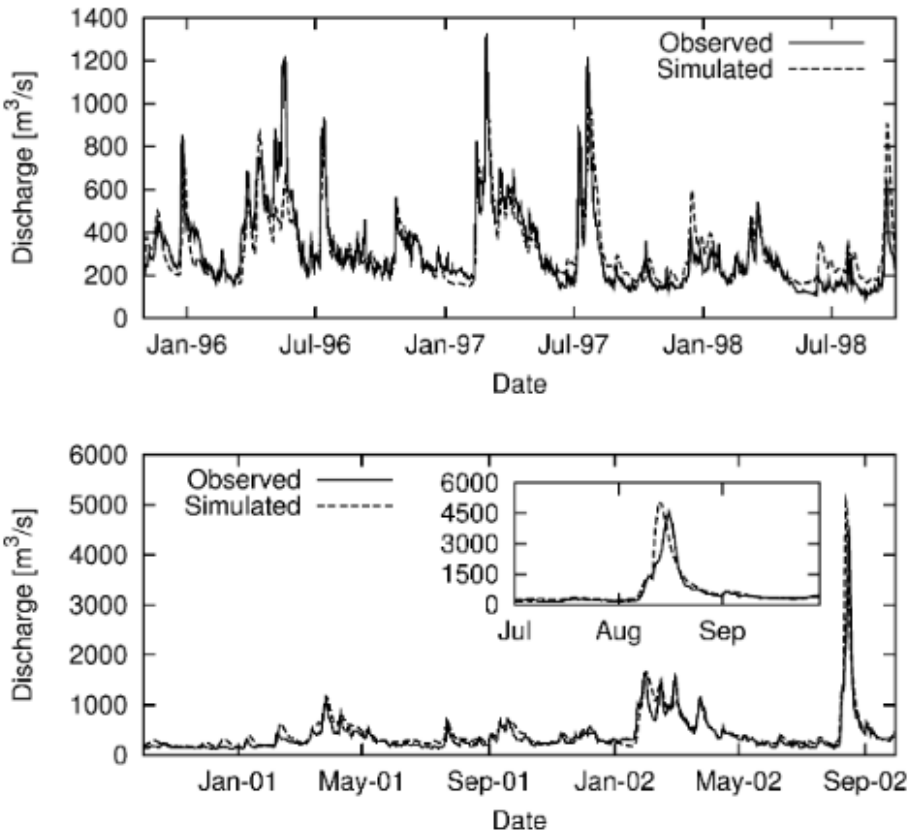


Figure 5: Observed and simulated discharge at Dresden for calibration (top) and validation (bottom) period. Note that both model runs were preceded by a one-year warm-up period, which is not shown here. Inset bottom graph zooms in on the August 2002 flood (Van Der Knijff et al., 2010).

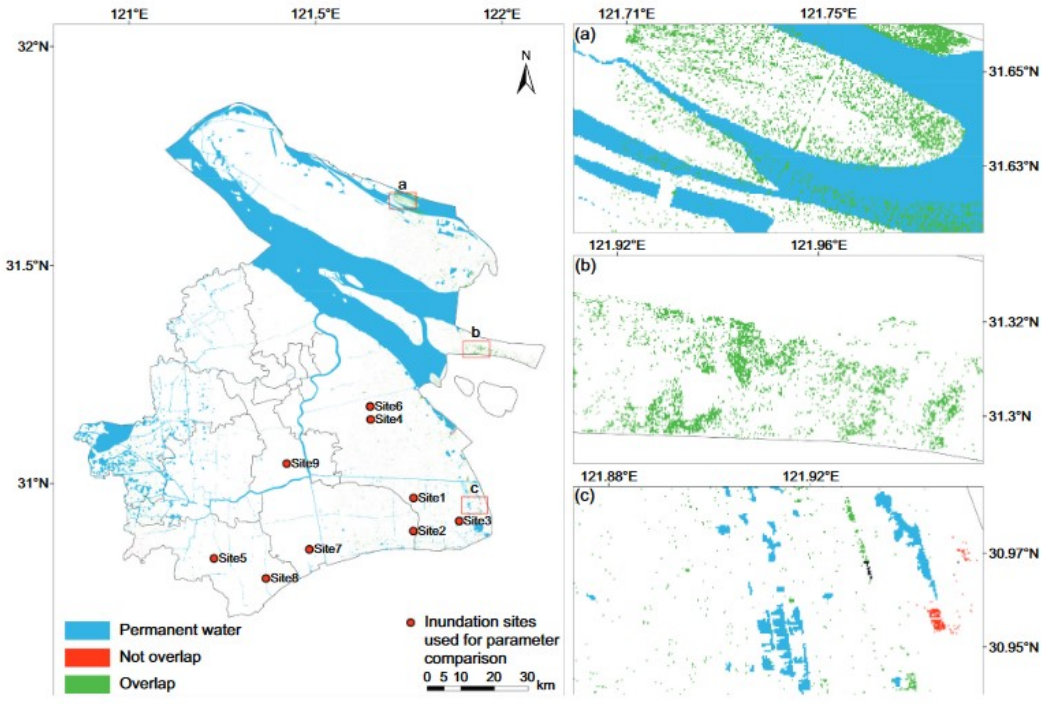


Figure 6: Flood extent validation using LISFLOOD on July 25th 2021. The distribution of inundation sites used for parameter comparison based on Li et al., 2024.

Open Data and Documentation

The Copernicus EMS provides open access to LISFLOOD, including extensive documentation, code repositories, and calibration tools. This transparency facilitates a broad user base among public agencies, research institutions, and disaster management entities. Documentation available via the JRC's GitHub repositories provides setup guides, data requirements, and technical support for model customization across various catchments, enabling diverse applications in flood and water management (https://github.com/ec-jrc/lisflood-code).

3.6 The EUROSTAT grid for exposure and socio-economic vulnerability mapping

3.6.1 Data source and description

The EUROSTAT 1km² grid is a standardized spatial framework distributed via GISCO (Geographical Information System of the Commission). It provides a standardized spatial framework to systematically map socio-economic variables across the European Union at a good spatial resolution. Its harmonized structure ensures robust cross-border comparability, thereby facilitating the identification of regional disparities and informing targeted policy interventions. Mandated by Regulation (EU) No 1089/2010 and Regulation (EU) 2018/1799 under the INSPIRE directive, this framework constitutes a critical element of the Union's spatial data infrastructure, fostering coherence in data collection, analysis, and planning activities. Furthermore, the grid directly supports the achievement of several United Nations Sustainable Development Goals (SDGs), including Goal 3 (Good Health and Well-being) by enabling assessments of population vulnerability to environmental and socio-economic stressors, Goal 10 (Reduced Inequalities) by highlighting socio-economic disparities, Goal 11 (Sustainable Cities and Communities) by facilitating data-driven urban planning, and Goal 13 (Climate Action) by providing a foundation for assessing climate risks and guiding adaptation strategies. This integration underscores the grid's strategic importance in advancing health, equity, sustainability, and resilience across Europe.

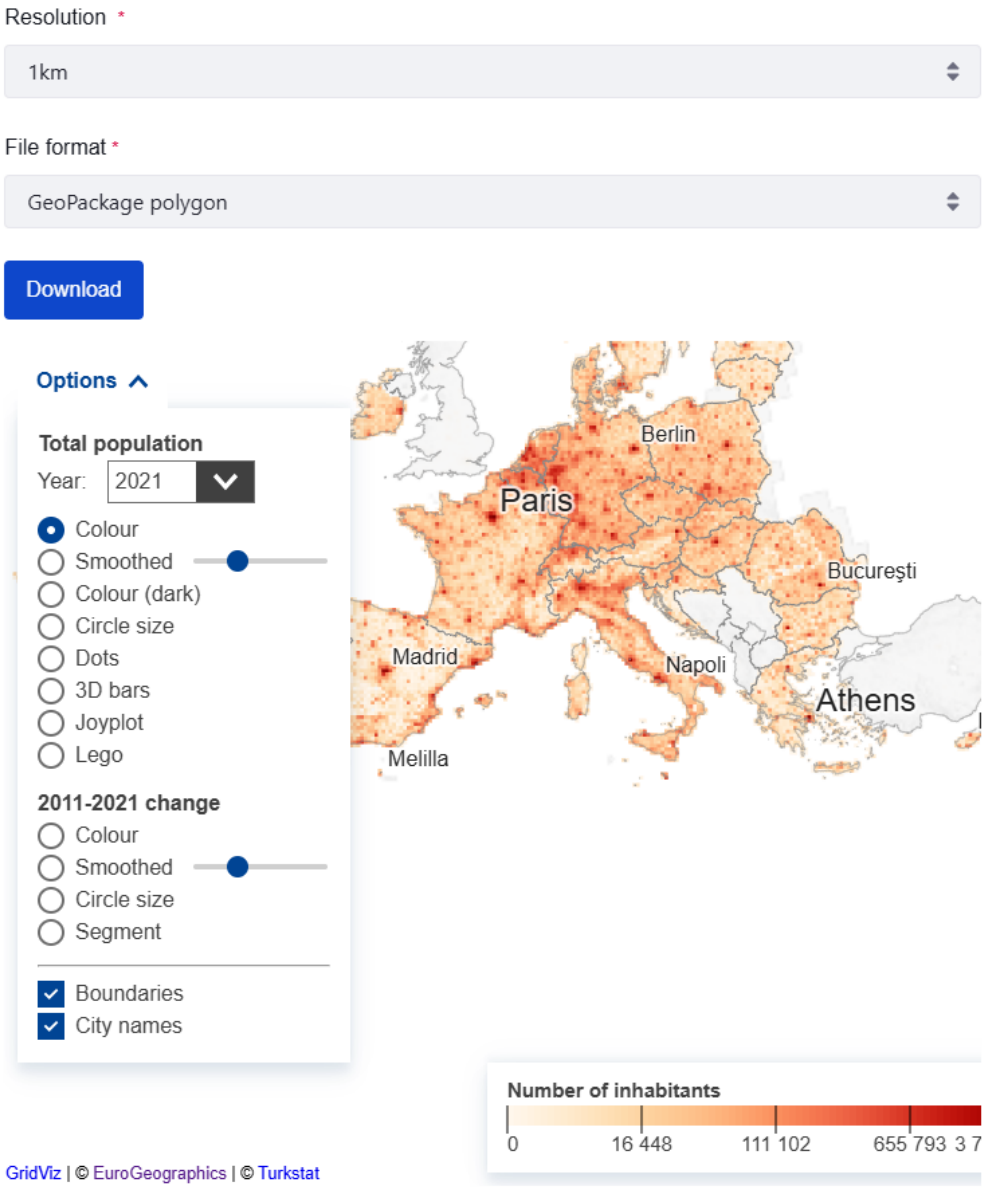


Figure 7: Screenshot of the Eurostat grid download page on the GISCO platform (europa.eu/eurostat/web/gisco/geodata/grids)

3.6.2 Basic variables and further data integration

The basic grid downloadable from GISCO in Geopackage format contains the 13 variables required by the EU Regulation, aggregated by cell and provided by census year.

Census topic categories to be broken down on the 1 km² reference grid

0.	SEX.0.: Total population
1.	SEX.1.: Male
2.	SEX.2.: Female
3.	AGE.G.1.: Under 15 years
4.	AGE.G.2.: 15 to 64 years
5.	AGE.G.3.: 65 years and over
6.	CAS.L.1.1.: Employed persons ([1])
7.	POB.L.1.: Place of birth in reporting country
8.	POB.L.2.1.: Place of birth in other EU Member State
9.	POB.L.2.2.: Place of birth elsewhere
10.	ROY.1.: Place of usual residence one year prior to the census unchanged
11.	ROY.2.1.: Place of usual residence one year prior to the census: move within reporting country
12.	ROY.2.2.: Place of usual residence one year prior to the census: move from outside of the reporting country

Figure 8: Regulation (EU) 2018/1799 - List of the 13 mandatory census variables included in the 1km resolution grid, which is available for download from the GISCO platform (ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/geostat)

The statistical institutes of the Member States validate the data of this grid and in some cases further parameterize the correlated attribute table, as in the case of INE, for Spain, where the additional variables are 144, ranging from socio-demographic data to data concerning dwellings, such as the occupied surface or the type of property.

3.6.3 Purpose and Use Cases

The EUROSTAT grid serves as a crucial tool for harmonizing spatial data collection and supporting various types of spatial analyses, including socio-economic and environmental vulnerability or hazard assessments. It enables the construction of indicators within its cells, which can then be used as standard reporting units to provide recommendations based on the data contained in each cell.

The strategic objective SO3 of the ClimEmpower project *"involves analysing climate change and impact indicators, creating a common set of indicators, and evaluating the effects of measures and strategies related to climate resilience"* which is an ambitious goal to achieve and the EUROSTAT grid provides a solution for the spatial aggregation of indicators, thus using the 1km cell as a standard reporting unit, and for the delivery of recommendations for climate resilience measures and strategies.

3.7 Conclusions

There are many and different gaps and needs between the regions. But the most common of them is the lack of detailed tools, hazard maps, access to climate related data, predictive models, risk scenarios, collaboration between agencies, awareness and education of people and detailed climate data.

A summary of hazards and the identified datasets, indices and services given in Table 11.

Table 11. Hazards connection with datasets, indices and services.

Hazards	Datasets	Indices (Satellite, Meteorological) , Models	Services
Flood	• European Flood Awareness System (EFAS) • Hydrology-related climate impact indicators from 1970 to 2100 derived from bias adjusted European climate projections • Climate indicators for Europe from 1940 to 2100 derived from reanalysis and climate projections	Flood extent: NDMI, dNDMI, NDWI, MNDWI, NDII LISFLOOD hydrological model	• GFWS (Climate data rescue management & mining, Climate Analysis Monitoring, Climate prediction, Climate projection) Climate Risk Dashboard • Flood Extent Enhancement and Water Depth Estimation Tool (FLEXTH) • European Risk Typology Map
Drought	• Climate indicators for Europe from 1940 to 2100 derived from reanalysis and climate projections • Fire danger indicators for Europe from 1970 to 2098 derived from climate projections	Drought : NDVI, EVI, VCI, NDWI, LSWI, NDDI, NMI Meteorological: SPEI, Keetch-Byram, VPD Soil Moisture	• GFWS (Climate data rescue management & mining, Climate Analysis Monitoring, Climate prediction, Climate projection) • Pan-European Fuel Map-server Climate Risk Dashboard • Flood Extent Enhancement and Water Depth Estimation Tool (FLEXTH) • European Risk Typology Map
Fire	• Fire danger indicators for Europe from 1970 to 2098 derived from climate projections • Climate indicators for Europe from 1940 to 2100 derived from reanalysis and climate projections	Fire impact: NDVI, dNDVI, NBR, RBR, RdNBR, NDBI	• Climate Risk Dashboard (Indicators for climate variability, urban heat stress) • GFWS (Climate data rescue management & mining, Climate Analysis Monitoring, Climate prediction, Climate projection) • European Risk Typology Map
Heatwave	Climate indicators for Europe from 1940 to 2100 derived from reanalysis and climate projections	EHF (EHI)	• Climate Risk Dashboard (Indicators for climate variability, urban heat stress) • European Risk Typology Map • GFWS (Climate data rescue management & mining, Climate Analysis Monitoring, Climate prediction, Climate projection) • Climate Risk Dashboard

The proposed methodologies including hazard mapping (fire, flood) and drought monitoring, address the identified gaps. Datasets and indices can be integrated to provide actionable insights at regional level. Their combination with operational tools enables climate monitoring and when combined with exposed elements provide a risk overview and impact assessment tool.

4. Development of spatial downscaling Copernicus Satellite images

4.1 Introduction to spatial downscaling and overview on downscaled methods

One of the identified gaps is the lack of High Resolution (HR) multispectral satellite data for environmental applications with a relatively high temporal resolution. HR data with a relatively high observation frequency are provided by Copernicus, namely Sentinel-2 satellite twins, however not all of their spectral bands offer high enough spatial resolution, which can be a potential drawback in various environmental applications that utilize these spectral bands. For this reason, in this task the focus is on exploring the effect of various spatial downscaling methodologies applied on the low spatial resolution spectral bands of Sentinel-2.

This provides a theoretical overview of the purpose of this deliverable, focusing on filling data gaps in high-resolution multispectral satellite imagery using spatial downscaling techniques on Sentinel-2 data. Section 4.1.1 outlines the problem definition and challenges. These include the limitations of Sentinel-2's multispectral bands and the need for advanced pansharpening techniques. Section 4.1.2 describes the Sentinel-2 satellite specifications and data processing levels. Section named “Related Work” highlights the challenge posed by the lack of a panchromatic band and reviews existing pansharpening techniques and their applications. Section 4.3.1 describes the methodology in detail, including the pre-processing of the data and the generation of synthetic panchromatic band. Section 4.3.2 discusses the different pansharpening methods, including Brovey, IHS, Wavelet, and CNN-based approaches. Section 4.4 describes the results, highlighting the effectiveness of these methods in improving spatial and spectral resolution, with Brovey, IHS, and PNN performing best. Section 4.5 concludes with a summary of the results, highlighting the importance of the selection of appropriate pansharpening techniques and pointing out the limitations of some methods, such as the Wavelet approach.

4.1.1 Problem definition and challenges

In the field of remote sensing, the Copernicus program's Sentinel-2 imagery stands out for its diverse applications in land monitoring, agriculture, and climate change studies. Despite its utility, the relatively coarse spatial resolution of the Sentinel-2A multispectral (MS) bands presents challenges for precise feature localization. This requires advanced techniques, such as pansharpening, to improve image quality by combining high-resolution panchromatic (PAN) imagery with MS imagery to produce enhanced spatial and spectral resolution.

The coarse spatial resolution of the Sentinel-2A MS bands can limit the accuracy of mapping features, such as morphological features in urban environments. This limitation arises from the mechanical and technological limitations in satellite sensor design, which prevent a single sensor from acquiring images with both high spectral and spatial resolution (Dong et al., 2021; Hashim et al., 2022). For example, sensors such as Quickbird, Worldview, and GeoEye have a high-resolution PAN band that is not available on Sentinel-2 and Rapid-Eye sensors, which only acquire images in multispectral mode (Mahmoud, 2021). According to Kumar et al. (2014), having images with both high spectral and spatial resolution is critical for mapping complex morphological structures.

There are various ways which can be used for tackling the lack of higher spatial resolution satellite data. Some of them are data fusion, super-resolution algorithms and interpolation methods. This task focuses on pansharpening the 20m Sentinel-2 bands.

4.1.2 Copernicus Satellite Data (Description of data sources)

The data source we used was the Copernicus Sentinel-2 series. Sentinel-2 is an Earth observation mission of the Copernicus program for land monitoring, agricultural applications, and climate change studies. The mission consists of two identical satellites; Sentinel-2A, launched in 2015, and Sentinel-2B, launched in 2017, 180 degrees apart in the same orbit where each satellite has a 10-day revisit time at the equator. The twin satellite configuration allows the typical 10-day revisit cycle to be enhanced into just 5 days. The Sentinel-2 satellite platform includes a Multispectral Instrument (MSI) that measures the Earth's reflected radiation in 13 spectral bands ranging from the visible to the shortwave infrared regions of the electromagnetic spectrum, producing a 290 km swath with the VNIR and SWIR sensors, each consisting of 12 detectors arranged in two offset rows.

Specifically, it includes four 10-meter spatial resolution bands; the standard RGB bands (blue at about 493 nm, green at about 560 nm, and red at about 665 nm) and a near-infrared band at about 833 nm. These are bands 2, 3, 4, and 8 (visible and near infrared). In addition, there are six 20-meter spatial resolution bands; four narrow bands in the VNIR vegetation red-edge spectral region (approximately 704 nm, 740 nm, 783 nm, and 865 nm) and two broader SWIR bands (approximately 1610nm and 2190nm) corresponding to bands 5, 6, 7, 8A (red edge), and 11, and 12 (shortwave infrared), respectively. These bands are useful for tasks such as detecting snow, ice or clouds and assessing vegetation moisture stress. There are also three 60-meter spatial resolution bands primarily for cloud screening and atmospheric correction, including a band at about 443 nm for aerosol detection, one at 945 nm for water vapor, and another at 1374 nm for cirrus cloud detection, which are bands 1, 9, and 10, respectively. All bands along with their wavelength and spatial resolution are shown in *Figure 1*.

Copernicus Sentinel-2 products are available in two processing levels: Level-1C (L1C) and Level-2A (L2A). L1C products contain top-of-the-atmosphere reflectance data, i.e. they are not corrected for atmospheric effects. In contrast, L2A products provide surface reflectance data

that has been atmospherically corrected using the sen2cor algorithm [26] which is a generic algorithm[27].

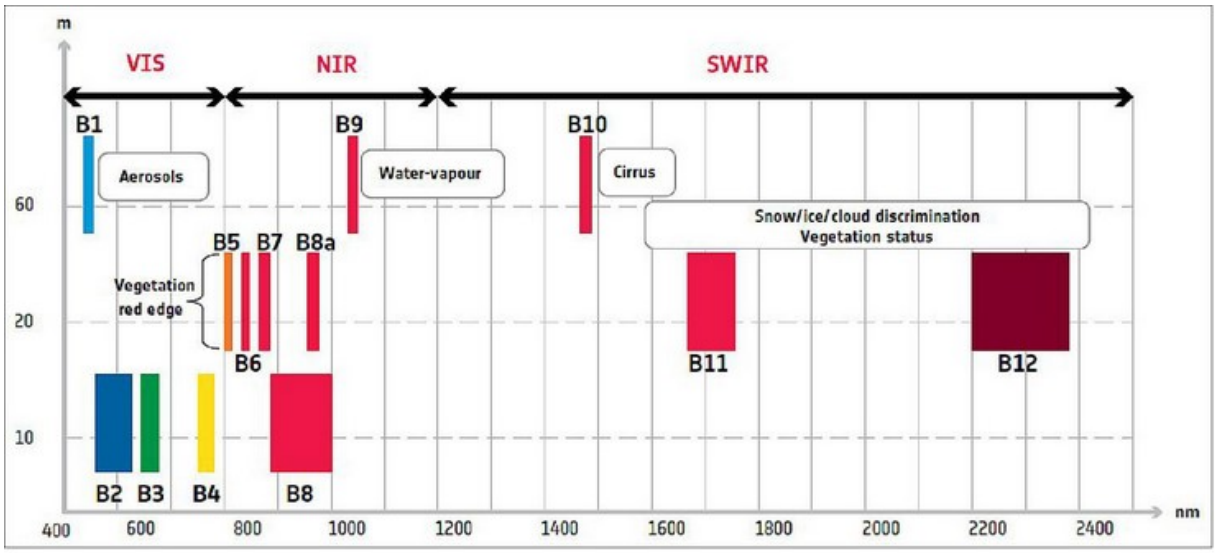


Figure 9: Spectral bands and band names and aliases for Sentinel-2 MSI (source: https://www.eoportal.org/satellite-missions/copernicus-sentinel-2).

4.2 Related Work

Pansharpening is a promising tool to simultaneously achieve high spatial and spectral resolution by merging MS and PAN bands (Park et al., 2017); Dong et al., 2021; Mallick et al., 2021; Hashim et al., 2022). The MS bands are enhanced using high-resolution PAN images, potentially improving classification accuracy (Gašparović and Jogun, 2017; Zheng et al., 2017; Hashim et al., 2022). The enhanced images take advantage of the strengths of both image types, which is particularly beneficial for mapping complex environments (Kumar et al., 2014).

Pansharpening approaches are typically classified into three categories: component substitution (CS), statistics-based methods (ST), and multi-resolution analysis (MRA) (Zhen et al., 2017). While many researchers have highlighted the superiority of CS methods over MRA approaches in achieving high-resolution image fusion (Mhangara et al., 2020; Mahmoud, 2021; Mallick et al., 2021), others argue that MRA-based algorithms can achieve better spectral and spatial quality (Zheng et al., 2017). Ultimately, the effectiveness of a fusion method depends on its ability to preserve the spectral properties and spatial information of the multispectral and panchromatic data, respectively (Park et al., 2017; Zheng et al., 2017; Kaplan and Avdan, 2018).

One challenge with Sentinel-2 is the lack of a panchromatic band required for pansharpening. Several methods have been developed to create a synthetic panchromatic band:

[26] https://step.esa.int/main/snap-supported-plugins/sen2cor/

[27] https://sentiwiki.copernicus.eu/web/s2-mission

1) *Pan1*: This method involves averaging all the fine multispectral bands (Bands 2-4 and Band 8) to create a panchromatic band.

2) *Pan2*: This method uses the average of band 4 (red) and band 8 (NIR) as a panchromatic band for downscaling bands 5-7, and band 8 as a panchromatic band for downscaling bands 8a, 11, and 12.

3) *Pan3*: Wang's method includes a linear combination of four fine bands and 20m bands, using NIR for fusion with Bands 6, 7, and 8a, and Bands 3 and 4 for fusion with Bands 5, 11, and 12 (Gordana Kaplan - Ugur Avdan, 2018).

Pansharpening of Sentinel-2A imagery has received considerable attention due to the availability of 10 m bands that have the potential to improve the resolution of 20 m bands (Zhang, 2004; Park et al., 2017; Phiri et al., 2020). This approach has been used in various remote sensing applications, including water body mapping (Feng et al., 2012; Che et al., 2015; Wu and Liu, 2015; Du et al., 2016), mineral mapping (Ge et al., 2020), leaf area index and chlorophyll content retrieval (Zhang et al., 2019), and land use/land cover (LULC) classification (Gašparović and Jogun, 2017; Zheng et al., 2017).

Despite its potential to improve mapping applications by increasing spatial resolution, pansharpening is associated with spectral and spatial distortions depending on the approach used (Jawak and Luis, 2013; Pandit and Bhiwani, 2015; Grochala and Kedzierski, 2017; Park et al., 2017). To mitigate these distortions, some researchers recommend incorporating image texture and contextual information to ensure a comprehensive evaluation of the pansharpened images (Palsson et al., 2011; Du et al., 2016; Zheng et al., 2017).

4.3 Methodology of spatial downscaling

4.3.1 Data and Preprocessing

Our methodology focuses on the application of pansharpening techniques for the downscaling of satellite imagery. The selected area of interest is Chennai, India, with geographical coordinates (E80.076061, N13.385915) and (E80.260207, N13.567355) and the image was acquired on 29/04/2024. This area was selected for its concentration of features that would facilitate the observation of downscaling effects. Our aim was to include a range of features, such as buildings, water bodies and agricultural crops, so that we could analyse the effect of downscaling in different environments. We also sought a region with enough detail to be able to see the variations after downscaling. Upon review, we found that this region met these criteria, making it an ideal choice. We used Sentinel-2 data for this study. Sentinel-2 lacks a true panchromatic band, so in order to apply the pansharpening to downscale the images from 20 m to 10 m resolution, we had to create a synthetic panchromatic band using the red, green, blue, and NIR bands. We accessed Sentinel-2 data, which includes bands at 10m, 20m, and 60m resolution, but we did not use the 60m bands as they are dedicated mostly for atmospheric correction and cloud detection, among others.

A number of preprocessing steps were performed using Python libraries (e.g., Rasterio), including:

4) **Definition of bounding box and windowing:** A bounding box was defined to extract the area of interest for Chennai, India.

5) **Read and Subset of the data:** Data within the defined bounding box was read and cropped from both the 10m and 20m resolution bands of the Sentinel-2 imagery.

6) **Generation of synthetic Panchromatic band:** Since Sentinel-2 lacks a true panchromatic band, a synthetic panchromatic band was created by pixel-wise averaging the red, green, blue, and NIR bands. This step is critical for improving spatial resolution.

Pansharpening was performed using the synthetic panchromatic band and the multispectral bands at 10 m resolution. This technique combines the high spatial resolution of the panchromatic band with the spectral information of the multispectral bands, resulting in downscaled images with enhanced spatial detail suitable for various applications.

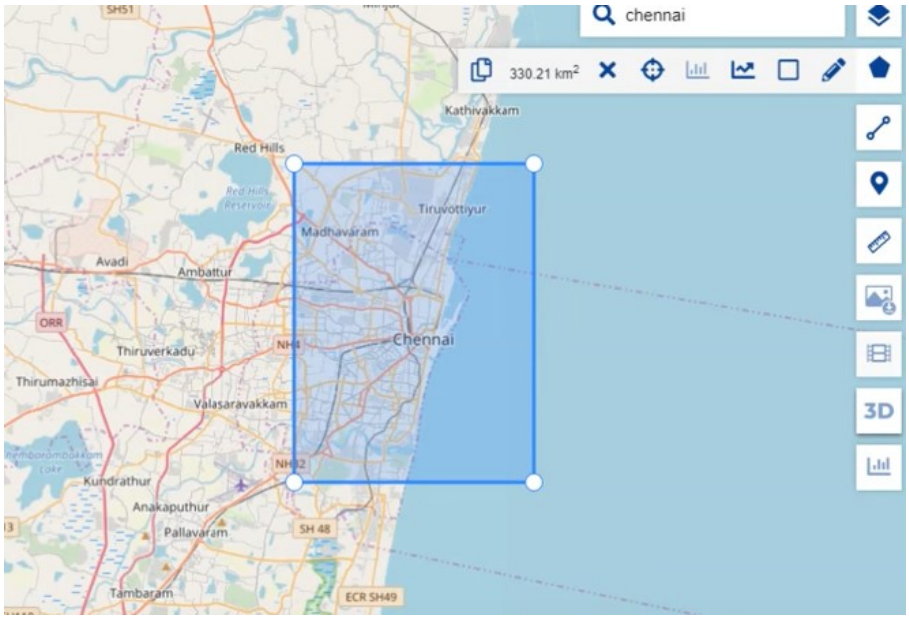


Figure 10: Figure depicts the location selected for the spatial downscaling in Copernicus Data Space Ecosystem Browser.

4.3.2 Pansharpening Methods

For pansharpening, we used three relatively traditional algorithms and a convolutional neural network (CNN) that have been previously used in published articles. These were the Brovey Transform (BT), the Intensity, Hue, and Saturation Transform (IHS), the Wavelet Transform (WPC), and the Pansharpening Neural Network (PNN). Each image fusion algorithm belongs to a different category of pansharpening approaches as described in the Related Work section.

Brovey Transform (BT)

BT is described as a simple method (Palsson et al., 2011; Pandit and Bhiwani, 2015) that uses a mathematical combination of the MS bands and the PAN band for pansharpening (Pandit and Bhiwani, 2015). This method is defined by Eq. (1). It uses an algebraic expression to integrate the overall brightness of the PAN image into each pixel of the MS image (Pushparaj and Hegde, 2016). However, according to Pandit and Bhiwani (2015), this approach can sometimes distort the radiometric properties of the bands.

$$DN_{fusedMSi} = \frac{(DNbi)}{DNb1 + DNb2 + \cdots + DNbnDNbi} DN_{PAN}$$

where DN is the digital number of the specific band and bi refers to a specific band of the MS image.

Intensity, Hue, and Saturation (IHS) Transform

IHS is considered to be the most widely used pansharpening technique (Sarp, 2014; Niazi et al., 2015) due to its fast computational capabilities (Pandit and Bhiwani, 2015). This method is defined by equation (3). This method relies on colour space transformation (Grochala and Kedzierski, 2017), which involves converting a composite of red, green, and blue (RGB) bands into an IHS colour space (Zhang, 2004; Sanli et al., 2017). The transformation separates the intensity of the two colour components (Du et al., 2007). In the fusion process, a high-resolution PAN image replaces the intensity (I) band (Zhang, 2004; Borana et al., 2019), while an interpolation technique resamples the hue (H) and saturation (S) bands to the higher resolution pixel size (Nikolakopoulos, 2008). To obtain a pansharpened image, a reverse IHS transformation is performed on the PAN, H, and S bands (Zhang, 2004). Although IHS can cause significant spectral distortion (Choi, 2006), it is still considered a simple and efficient approach (Niazi et al., 2015).

$$I_L \sum_{i=1}^{N} w_i MS_i \quad (2)$$

$$MS_k = MS_k + (i = 1 \sum Nwi) - 1(P - IL) \qquad k = 1, \dots, N \quad (3)$$

in which IL follows from (2) and MSK is the original MS image. In general, the coefficients {wi}k=1,...,N are all equal to 1/N (Tu et al., 2001). Alternatively, they can be optimized to the responses of the spectral channels, i.e., MS and PAN (Tu et al., 2004).

Wavelet Transform

Wavelet transform converts high-resolution images into intensity values by decomposing the image into multiple channels based on local frequency content. This approach provides a framework for decomposing images into multiple new images, each with a different resolution. Unlike the Fourier transform, which represents the frequency content of an image, the wavelet transform provides an intermediate representation between the Fourier and spatial domains, allowing for good localization in both the frequency and spatial domains.

However, not all wavelet bases are ideal for image fusion. The orthonormal wavelet basis transform lacks shift invariance, which can be problematic in signal analysis, pattern recognition, or data fusion. To achieve shift-invariant discrete wavelet decomposition for images, biorthogonal spline wavelets are used. These wavelets construct scaling functions and wavelets from splines without relying on orthogonalization tricks (King et al., 2001).

Pansharpening Neural Network (PNN)

The proposed CNN-based pansharpening model takes advantage of the successful super-resolution method previously described in (Dong et al., 2015) and adapts it for pansharpening in remote sensing (Masi et al., 2016). This adaptation aims to improve the spatial resolution of MS images using PAN data, while integrating domain-specific knowledge from remote sensing image processing. The following description outlines the basic architecture, dataset handling, and specific implementation details of this pansharpening approach.

Basic Architecture

The basic architecture of the proposed model closely mirrors the Super-Resolution Convolutional Neural Network (SRCNN) architecture specifically tailored for Ikonos and GeoEye-1 satellite data, which include 4-band multispectral components and a higher resolution panchromatic band. In our case, the six low-resolution spectral bands are first upsampled and interpolated to match the resolution of the panchromatic band. These upsampled bands are then combined with the high-resolution PAN band to form a 7-component input to the CNN. As a result, the network operates at the target resolution from the start, eliminating the need for further up/down-sampling. The output consists of the original six multispectral bands, but with the enhanced panchromatic resolution.

The network architecture retains a three-layer structure as used in the SRCNN, but with one modification: the 1×1 kernels in the central layer are replaced by 5×5 kernels, which preliminary experiments have shown to improve performance. This relatively flat architecture is advantageous for remote sensing, where large training datasets are often not available, making it easy to train the model with limited data. The output of the CNN is a high-resolution multispectral image that approximates what would be acquired by a multispectral sensor operating at the same spatial resolution as the PAN component. This model is trained and validated using the Wald protocol, which downsamples a multi-resolution image to create a valid reference for the original multispectral component.

Implementation Details

The CNN is trained to produce a high-resolution multispectral image that closely matches a hypothetical multispectral image with the same spatial resolution as the panchromatic image. This process is complicated by the fact that such high-resolution multispectral images do not exist, creating challenges for performance evaluation and network training. To address these challenges, the Wald protocol is used, which involves the creation of downsampled multi-resolution images using the original multispectral component as a reference (Wald et al., 1997).

During training, the original image is downsampled and its multispectral component is interpolated. The resulting image, now with an additional band, is tiled and fed into the CNN. The original multispectral component is also tiled and used as a reference. The CNN parameters are adjusted to minimize the mean square error between the pansharpened tiles and the reference tiles, following a backpropagation approach with stochastic gradient descent. The training parameters are updated using a momentum parameter to reduce randomness, and the learning rate is varied to speed up convergence. The trained parameters are then used to pansharpen real multi-resolution images at their original resolution.

Training Process and Components

The PNN model combines low resolution multispectral and high-resolution panchromatic images into a single input to the CNN. The model architecture consists of three convolutional layers:

7) **Input Layer**: The model takes two inputs - the low-resolution MS image of size (height, width, channels) = (32, 32, 3) and the high resolution PAN image of size (height, width, channels) = (32, 32, 1).

8) **First Convolutional Layer:** A convolutional layer with 64 filters of size (9, 9), using ReLU activation.

9) **Second Convolutional Layer:** A convolution layer with 32 filters of size (5, 5), using ReLU activation.

10) **Output layer:** A convolutional layer with a number of filters equal to the number of bands in the low-resolution MS input (3 in this case), using a filter size of (5, 5) and ReLU activation.

The model is built using the Adam optimizer with a learning rate of 5×10^{-4} and a loss function based on mean square error (MSE). The performance metric used is the peak signal-to-noise ratio (PSNR).

Training Data Preparation

By generating patches from the input images, the training data is prepared. The preprocessing pipeline consists of several transformations applied to the input images. First, high resolution PAN and low-resolution MS images are augmented by cropping and downgrading. Next, images are divided into smaller 32x32 patches with a step size of 8. This approach ensures a large number of training samples. Then, Bicubic interpolation is used to upsample the low-resolution MS patches to match the resolution of the High resolution PAN patches. Finally, the extracted patches are shuffled to create the training set, ensuring a mix of different image regions.

Training Process

The training process is as follows. First, the PNN model is compiled using the Adam optimizer and the MSE loss function with PSNR as the performance metric. To adjust the learning rate during training, a learning rate scheduler is implemented that reduces the learning rate at certain epochs to ensure better convergence. Based on the validation PSNR, a model checkpoint is used to store the best model weights. The model is trained for 50 epochs with a batch size of 32. During each epoch, the model parameters are updated to minimize the MSE between the sharpened output and the high-definition reference patches.

Remote sensing specific architecture

While the basic architecture provides satisfactory results, further analysis specific to remote sensing imagery led to additional enhancements. The focus has been on first-order filters, which are easier to interpret. For example, analysis of WorldView-2 images with nine input bands showed that some filters, particularly low-energy filters, were not capturing valuable information and could be removed to streamline the network. High-energy filters, on the other hand, showed well-defined patterns, with spatial structure appearing primarily in the panchromatic band and spectral selectivity in the other bands. This behaviour is consistent with known spectral indices such as Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI), which were effectively captured by specific filters in the network.

The model is better adapted to the specific characteristics of the imagery by refining the filters and using prior knowledge from remote sensing to improve performance. These refinements ensure that the network is robust and effectively trained, even with a limited amount of high-resolution training data. The overall approach shows significant potential for improving the spatial resolution of multi-spectral imagery while preserving the spectral fidelity of the imagery, making it a valuable tool for remote sensing applications (Masi et al., 2016).

4.4 Results and discussion

In our research, we applied several pansharpening techniques described above to increase the spatial and spectral resolution of Sentinel-2 imagery. In order to evaluate the effectiveness of these methods, we performed a series of quality checks and compared the results to the reference image.

The original reference image for the red edge band 5, as shown in Figure *12a*, presents difficulties in clearly distinguishing field boundaries and the roads between them. The area circled for analysis shows this lack of clarity. The fields appear noisy and lack the definition needed for accurate interpretation. In the appendix, you can find similar visual comparisons for the other downscaled bands.

After applying the pan-sharpening techniques, we observed a significant improvement in the clarity of the images, with some of the techniques performing better than others. The Brovey transformation (*Fig. 12b*) produced a much clearer image. It significantly improved the contrast between fields and roads and reduced the noise present in the reference image, allowing for more precise identification of field boundaries. Similarly, the IHS transform (*Fig. 12c*) provided a sharp and clear image with improved contrast and spatial detail, making the fields and roads easily distinguishable. The PNN transform (*Fig. 12e*) also delivered high quality results, providing a clear image with well-defined boundaries and minimal noise, comparable to the Brovey and IHS techniques. The Wavelet transform (*Fig. 12d*), on the other hand, did not perform as well. Although it offered some improvement over the reference image, the resulting image was less sharp, and the presence of horizontal lines, a byproduct of the Wavelet method, further compromised its overall quality.

Besides the visual inspection, we also examined the pixel values of the pansharpened images (*Fig. 11*). Specifically, we found that the pixel values in the pansharpened images were an order of magnitude larger than those in the reference images. This increase in pixel intensity is a reflection of the improved spatial resolution achieved by pansharpening, and an indication of successful transformation.

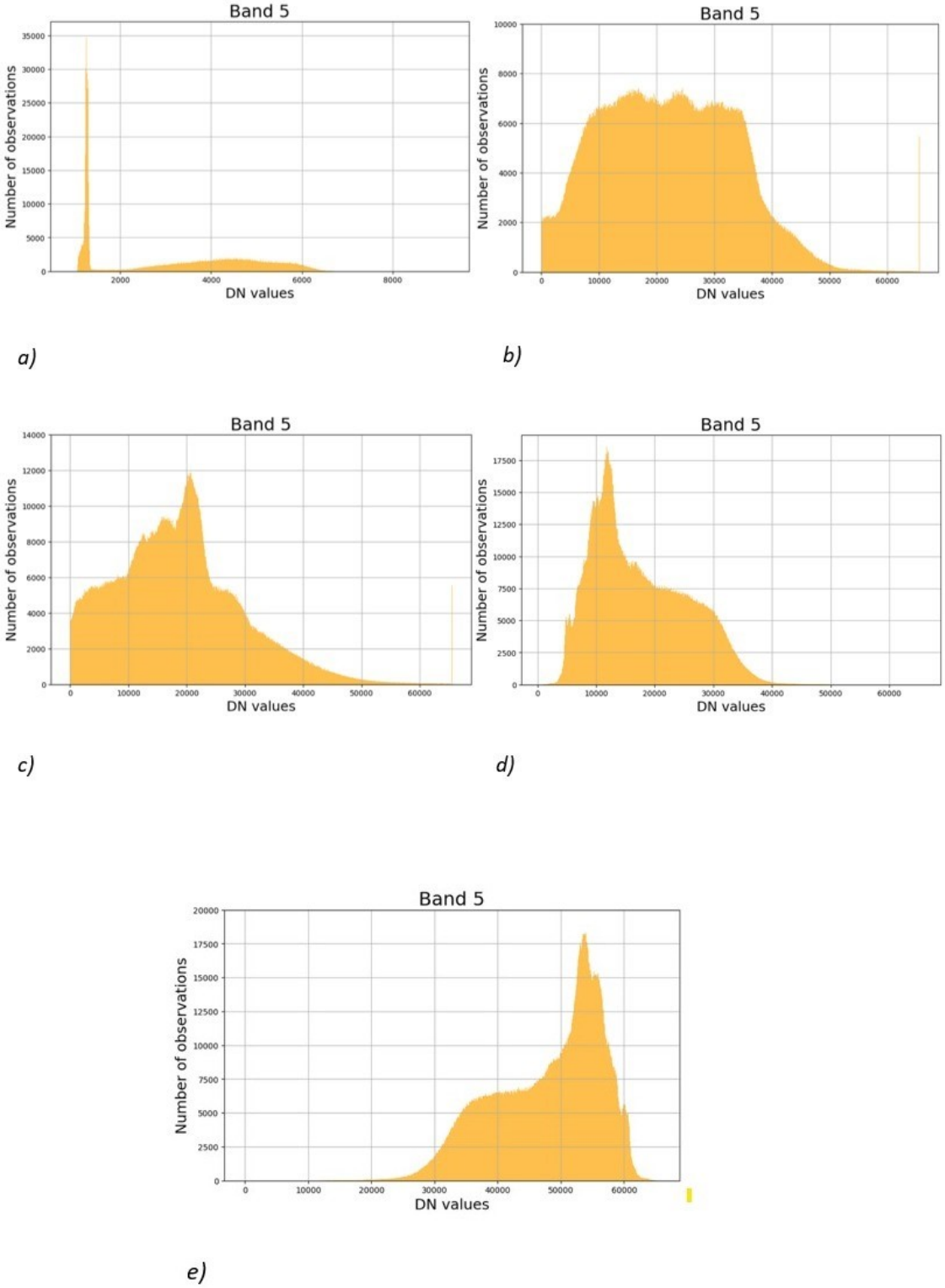


Figure 11: Histograms representing the distribution of DN values for Band 5 of Sentinel-2 across different pansharpening techniques where a) Reference image b) Brovey Transform c) IHS Transform d) Wavelet Transform e) PNN

a) b)

c) d)

e)

Figure 12: Figures depict the various downscaling techniques applied on band 5 of Sentinel-2 where a) Reference image b) Brovey Transform c) IHS Transform d) Wavelet Transform e) PNN

4.5 Conclusions

To address the challenges posed by the lack of a high-resolution panchromatic band, the research focused on improving the spatial resolution of Sentinel-2 imagery through various pansharpening techniques. The application of these methods, including the Brovey Transform, IHS Transform, Wavelet Transform, and PNN, demonstrated their potential to significantly improve the spatial detail and clarity of Sentinel-2 imagery, particularly for environmental monitoring and land use classification.

The results showed that pansharpening can effectively bridge the gap between the high spatial resolution required for detailed analysis and the coarser resolution available in certain Sentinel-2 spectral bands. Brovey and IHS methods, along with PNN, produced the most visually pleasing, high-quality imagery with clear field boundaries and streets, reduced noise, and improved contrast. These methods showed great promise for applications where accurate feature delineation is critical, such as agriculture and urban planning.

However, while the Wavelet transform provided some improvements, it did not perform as well as the other techniques. This suggests that not all pansharpening methods are equally suitable for enhancing Sentinel-2 imagery. The presence of artifacts such as horizontal lines is an indication of the limitations of the wavelet approach for this specific application.

Furthermore, the increase in pixel intensity observed in the pansharpened images relative to the reference images confirms the effectiveness of these techniques in improving spatial resolution. This improvement highlights the importance of selecting the appropriate pansharpening method depending on the specific requirements of the analysis and the characteristics of the images.

In conclusion, pansharpening is a valuable tool for improving the usability of Sentinel-2 data in various environmental applications, although it presents some challenges, particularly with respect to spectral distortions and the creation of synthetic panchromatic bands.

5.Social Media Crawler & Analysis

Social media crawling is the process of automatically collecting and extracting data from social media platforms like X (formerly Twitter). This technique enables the gathering of unstructured data shared by users, which can be processed and analysed for various applications such as disaster management (Karimiziarani et al. 2022) and monitoring (Sufi et al. 2022). It involves deploying web crawlers or automated tools to scan social media platforms for relevant posts, hashtags, and mentions, systematically harvesting the content for further analysis. In the context of natural disasters, social media crawling becomes an indispensable tool for providing timely information on events like extreme weather, floods, droughts, and wildfires.

Social media crawling is a crucial tool for real-time monitoring of weather conditions, floods, and wildfires by analysing user-generated content on platforms like X. By tracking weather-related keywords, hashtags, authorities can gather early indicators of extreme weather events such as storms, heatwaves, or heavy rainfall, often providing updates that traditional monitoring systems might miss. For floods, social media data, including posts about waterlogged areas or rising water levels, can help map and track the progression of flooding in real-time, enhancing situational awareness and guiding evacuation efforts. Similarly, during wildfires, social media posts about fire locations, size, and movement provide valuable

information for responders to prioritize resources and optimize firefighting strategies, especially in remote or hard-to-reach regions. Integrating this data with geographic information systems (GIS) further improves disaster management, aiding in accurate forecasting, timely warnings, and effective decision-making.

While extracting social media data provides valuable insights for disaster monitoring, it also presents challenges. The sheer volume of information requires advanced algorithms and machine learning techniques to filter out irrelevant content and accurately identify relevant information. Moreover, determining the location referenced in a post can be difficult, as users may not include explicit geographic details or use non-standardized language. To address this, natural language processing (NLP) methods are employed to analyse the text and infer the location based on contextual clues such as place names, landmarks, or event-specific hashtags.

In summary, social media crawling offers a powerful approach to monitoring weather events, floods, and wildfires, providing emergency responders with timely, situational awareness to improve disaster management. By utilizing the vast amount of information shared on social media, authorities can make informed decisions to protect lives, property, and the environment during natural disasters.

5.1 Social Media Framework

The Social Media Sensing framework (**Fehler! Verweisquelle konnte nicht gefunden werden.**) is a system developed to gather, analyse, and process social media posts concerning natural disasters. It incorporates various components to enhance the data and identify relevant events, which are subsequently stored in a MongoDB database and made available to end-users.

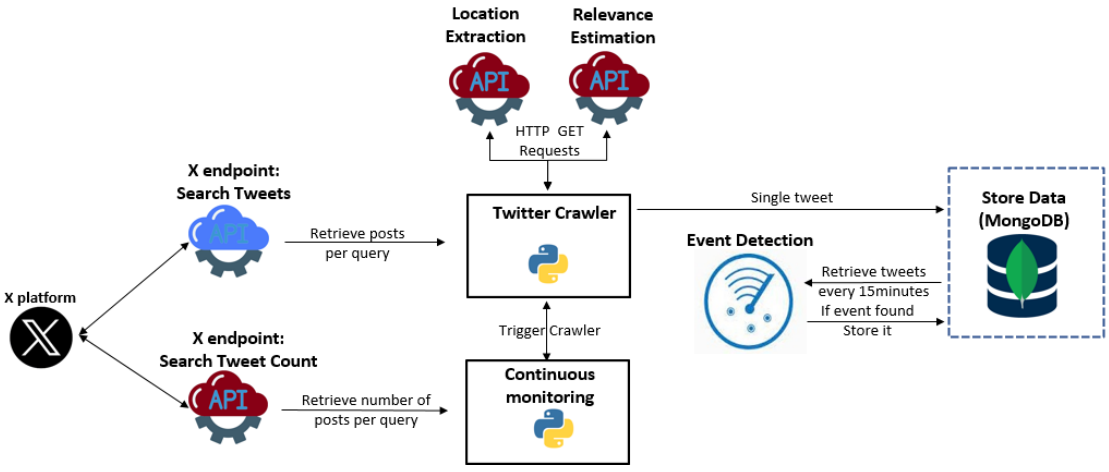


Figure 13: Social Media Framework

The process begins with the continuous monitoring of disaster-related activity on *X* through the Continuous Monitoring Module. When this module detects an unusually high level of activity, it triggers the crawler. The crawler then formulates complex queries based on predefined search criteria to collect posts that may be related to extreme weather events, such as floods, droughts, and wildfires.

Once the posts are collected, they are passed through two key processing components: Location Extraction and Relevance Estimation. These modules operate through independent

API calls that enrich each post with geographical data and relevance score. The Location Extraction component identifies geographical information within posts, while the Relevance Estimation component assesses the likelihood of each tweet being related to the specific crisis event under investigation. These enriched posts, now containing valuable metadata, are stored in the MongoDB database.

The Event Detection mechanism is a core component of the framework, designed to scan incoming posts for specific keywords, patterns, or signals that indicate the occurrence of an event, such as a fire. If an event is detected, the system groups the relevant posts as an event and stores them in database. This step is crucial for identifying actionable information within the vast amount of social media data, focusing on posts that are most likely to represent real-world incidents.

5.2 Search Criteria

Before the development of the social media crawler, it was crucial to establish the search criteria, including keywords or phrases, X accounts used for filtering content when querying X. The selection of these criteria was done in collaboration with pilot leaders, who could leverage their expertise in the relevant scientific fields and knowledge of their native language. This joint effort ensured a more precise identification of search parameters, leading to the retrieval of valuable social media data.

5.2.1 Definition of search criteria

The first step was to gather information about the area of interest and establish communication with the end users of ClimEmpower. This involved understanding their experience with monitoring social media, determining how social data would contribute to their pilot case, and defining search criteria related to weather, floods, droughts, and fire events for crawling on platform X.

Following the establishment of communication with the end users and definition of the initial criteria, a comprehensive list of weather-related keywords, flood and drought-related terms, and fire-related keywords was compiled for use in monitoring and data crawling on X. These keywords were translated into the respective languages of the regions involved in the ClimEmpower project, including English, Croatian, Italian, Spanish, Greek.

For each weather-related phenomenon, such as tempests, heatwaves, torrential rains, and dust storms, the corresponding keywords were localized to capture relevant social media conversations within each participating country. This also extended to flood and drought scenarios, where terms related to flash floods, urban floods, desertification, and rising sea levels were identified to enable precise tracking of relevant discussions across platforms. Furthermore, keywords related to fire events, such as forest fires and wildfires, were included for regions prone to these hazards.

In addition to the keywords, specific regions of interest within each country were identified to ensure that data collection is focused and relevant. These regions include Osijek-Baranja County in Croatia, Sicily in Italy, Andalusia in Spain, Troodos in Cyprus, and Central Greece. These regions were chosen for their significance in the management of protected areas, environmental sustainability, and their vulnerability to climate-related events. By combining the localized keywords and targeted regions, ClimEmpower aims to enhance social data

monitoring and analysis, contributing to more effective climate adaptation strategies for each case.

Fehler! Verweisquelle konnte nicht gefunden werden.2 displays all the weather-related search criteria, including terms for various weather phenomena such as tempests, precipitation, heatwaves, extreme heat, and weather alerts, translated into English, Croatian, Italian, Spanish, Greek, and Cypriot. This ensures comprehensive monitoring of social media conversations related to weather events across all regions.

Table 12: Weather related criteria.

WEATHER RELATED KEYWORDS					
ENGLISH	**CROTIA**	**ITALY**	**SPAIN**	**CYPRUS**	**GREECE**
tempest	oluja	tempesta	tempestad	καταιγίδα	
precipitation	oborine / padaline	precipitazione	precipitación	υετός	
Heatwave	toplinski val	ondate di calore	ola de calor	καύσωνας	
extreme heat	vrućina	calore estremo	calor extremo		
heat stroke	toplinski udar	colpo di calore	golpe de calor	θερμοπληξία	
high temperature	visoka temperatura	caldo killer	altas temperaturas	υψηλή θερμοκρασία	
weather alert	vremenska uzbuna	allerta meteo	alerta meteorológica	ειδοποίηση καιρού	
torrential rains	obilne kiše	piogge torrenziali	lluvias torrenciales	καταρρακτωδης βροχή καταρρακτωδης	
hail	tuča / grad / krupa	salve	granizo	χαλάζι	
pouring rain	pljusak	pioggia battente	lluvia torrencial	καταρρακτώδης βροχή	
Saharan dust				σκόνη Σαχάρας	
African dust				Αφρικανική σκόνη	

Table 13 presents the floods and drought-related search criteria, including key terms for different types of floods (e.g., flash floods, river floods, urban floods) and drought conditions. These terms are essential for identifying discussions about flood risks, water scarcity, and desertification, particularly in areas vulnerable to such events.

Table 13: Floods and drought-related search criteria.

FLOODS / DROUGHTS RELATED KEYWORDS					
ENGLISH	**CROATIA**	**ITALY**	**SPAIN**	**CYPRUS**	**GREECE**
flash flood	bujična poplava	bombe d'acqua	inundación repentina	αστραπιαία / ξαφνική πλημμύρα	

FLOODS / DROUGHTS RELATED KEYWORDS				
river flood	riječna poplava			
urban flood	urbana poplava	allagamenti urbani	inundación urbana	αστική πλημμύρα
drought	suša		sequía	ξηρασία
desertification		desertificazione	desertificación	ερημοποίηση
pluvial flood			inundación pluvial	
coastal flood			inundación costera	

Table 14 contains the search criteria related to fires, specifically forest fires and wildfires. These keywords will help track social media activity related to fire outbreaks, which are particularly common in Mediterranean and forested regions during hot, dry conditions.

Table 14: Fire related keywords.

FIRES					
ENGLISH	**CROATIA**	**ITALY**	**SPAIN**	**CYPRUS**	**GREECE**
forest fire			incendio forestal	Δασική πυρκαγιά	
wildfire				πυρκαγιά	

Table 15 outlines the areas of interest for each participating region in the ClimEmpower project. It includes specific locations in Croatia (Osijek-Baranja County), Italy (Sicily), Spain (Andalusia), Cyprus (Troodos region), and Greece (Central Greece), ensuring that the monitoring efforts are geographically targeted to regions most susceptible to climate-related events such as floods, droughts, and wildfires.

Table 15: Area of Interest (AOI) of ClimEmpower regions.

REGIONS					
	CROATIA	**ITALY**	**SPAIN**	**CYPRUS**	**GREECE**
English	Osjek - Baranja (Public Institution for the Management of Protected Natural Values in Osijek-Baranja County)	Sicily	Regional Ministry for Sustainability, Environment and Blue Economy of Andalusia	Troodos Development Company LTD (ANET)	Region of Central Greece
Mother Tongue	Osječko-baranjske (Javna ustanova za upravljanje zaštićenim prirodnim vrijednostima na području Osječko-baranjske županije)	Regione Siciliana	Consejería de Sostenibilidad, Medio Ambiente y Economía Azul de la Junta de Andalucía (Junta de Andalucía)	Troodos Development Company LTD (ANET) ανετ Τροόδους Αναπτυξιακή εταιρεία κοινοτήτων	Περιφέρεια Στερεάς Ελλάδας

REGIONS	
	περιοχής Τροόδους

5.3 Social Media Crawling

X is a popular platform with millions of unique users who participate in a range of activities, including reporting disasters such as fire incidents. CERTH has developed a crawler that gathers posts in near real-time from the X API, offering a valuable tool for monitoring and analysing live data. This section describes the social media crawler.

The primary objective of the crawler is to collect social media posts related to weather, floods, droughts and wildfires, ensuring a reliable and continuous flow of data. Key features include real-time monitoring of X's public stream, filtering options that focus on specific keywords or user accounts, and the ability to handle large data volumes efficiently. Moreover, the crawler adheres to ethical standards and complies with X's terms of service, ensuring that all data is collected lawfully.

5.3.1 X API

To access X's endpoints, you need an X account, a developer account on the X portal[28], and at least Free tier access (the lowest level of access to X's API) and a bearer token[29]. In particular the X API offers Free tier, Basic tier and Pro tier. The free tier allows reading a limited number of posts (i.e., 100 posts per month) and the Basic tier, which costs $200 per month, only permits reading up to 15,000 posts per month and does not include access to real-time data or short-term historical data. Access to these functionalities gives the Pro tier, which comes with a price tag of $5000 per month, allowing users to read up to 1,000,000 posts per month providing also available real-time data and historical data features. This information is accurate as of November 15, 2024 (Table 16).

Table 16: X API tier costs.

Access Type	Current API Costs		
Cost	**Free**	**Basic**	**Pro**
Write Post	0$	200$ / month	5000$ / month
Read Post	✓	✓	✓
Read Post rate limit	**X**	15,000 posts/ month	1,000,000 posts/ month

From the above-mentioned tiers, CERTH chooses the Basic version as it offers a balance between cost and functionality. At $200 per month, it provides a feasible solution for maintaining a stream of relevant data without the prohibitive costs associated with the Pro version. While the Pro version offers more extensive capabilities, its high cost makes it less viable.

[28] https://developer.twitter.com/en/portal/dashboard
[29] https://developer.twitter.com/en/docs/authentication/oauth-2-0/bearer-tokens

In response to the rate limits in X API, CERTH has created a monitoring-based crawling strategy. This approach is designed to improve the efficiency of post retrieval while mitigating the impact of rate limits. By periodically monitoring the number of posts and concentrating crawling efforts during periods of significant activity spikes, CERTH can more effectively capture data related to specific search queries. This approach optimizes the collection of relevant posts.

5.3.2 Continuous Monitoring

Continuous monitoring involves the real-time tracking of social media activity to detect emerging events. This approach is essential for quickly identifying significant increases in post volumes, which may indicate unfolding incidents and thus a good opportunity to use the efficiently the X API rate limit.

During periods of high activity, the crawler queries the Tweets Counts[30] endpoint of the X API at 30-minute intervals, adhering to the platform's rate limits, which restrict calls to this endpoint to once every 30 minutes. Using the search criteria outlined in section 5.2, the crawler utilizes the Recent Tweet Counts endpoint to monitor social media activity by detecting significant increases in X post counts since the previous query. A noticeable spike signals a potential event, prompting the retrieval of posts from the specific time window, while the absence of such an increase leads the crawler to refrain from collecting posts during that period. This strategy ensures the efficient collection of relevant posts while staying within the platform's rate limits.

To achieve continuous monitoring, we employed the Isolation Forest algorithm (Liu, F., 2008) a powerful method for detecting anomalies in time-series data. In our system, we aggregate post data into 30-minute intervals, summing the number of posts in each period. This time series of post counts is then analysed using Isolation Forest to identify periods of unusually high activity, which could signal significant incidents. An analysis was conducted on a dataset derived from the social media platform X, focusing on historical data collected between June 2019 and August 2022. This dataset, consisting of 7,990 posts, focuses on posts related to wildfires within Greece.

The aim was to develop and evaluate the ongoing monitoring of social media activity, especially during times of increased posting activity. The research had two key objectives: firstly, to pinpoint and visually highlight periods of significant activity on the platform; and secondly, to identify if the isolation forest algorithm is effective for real-time monitoring and detecting anomalies in anticipation of future events.

In general, this algorithm is highly effective at identifying outliers without requiring pre-labelled anomaly data. It was configured with a contamination parameter that controlled the proportion of expected anomalies in the dataset. Anomalies were then identified based on the timestamps provided by the algorithm's output.

One of the key strengths of this approach is its robustness, especially in identifying outliers across varying post patterns. Whether the dataset was dense or sparse, the algorithm

[30] https://developer.x.com/en/docs/x-api/tweets/counts/introduction

successfully detected a wide range of anomalies. However, it is worth noting that this method can be computationally demanding and requires careful adjustment of the contamination parameter to achieve optimal performance.

In the time series of aggregated post counts, as illustrated in Figure 14, anomalies are highlighted with red markers. This visual representation makes it easy to identify periods of unusually high posting activity. What sets this method apart is its ability to distinguish true anomalies from regular cyclical patterns, such as daily or weekly fluctuations. This capability significantly reduces the likelihood of false positives, ensuring that the identified anomalies genuinely reflect significant deviations from the norm. By distinguishing true anomalies, the system minimizes false positives and focuses on deviations that are truly significant.

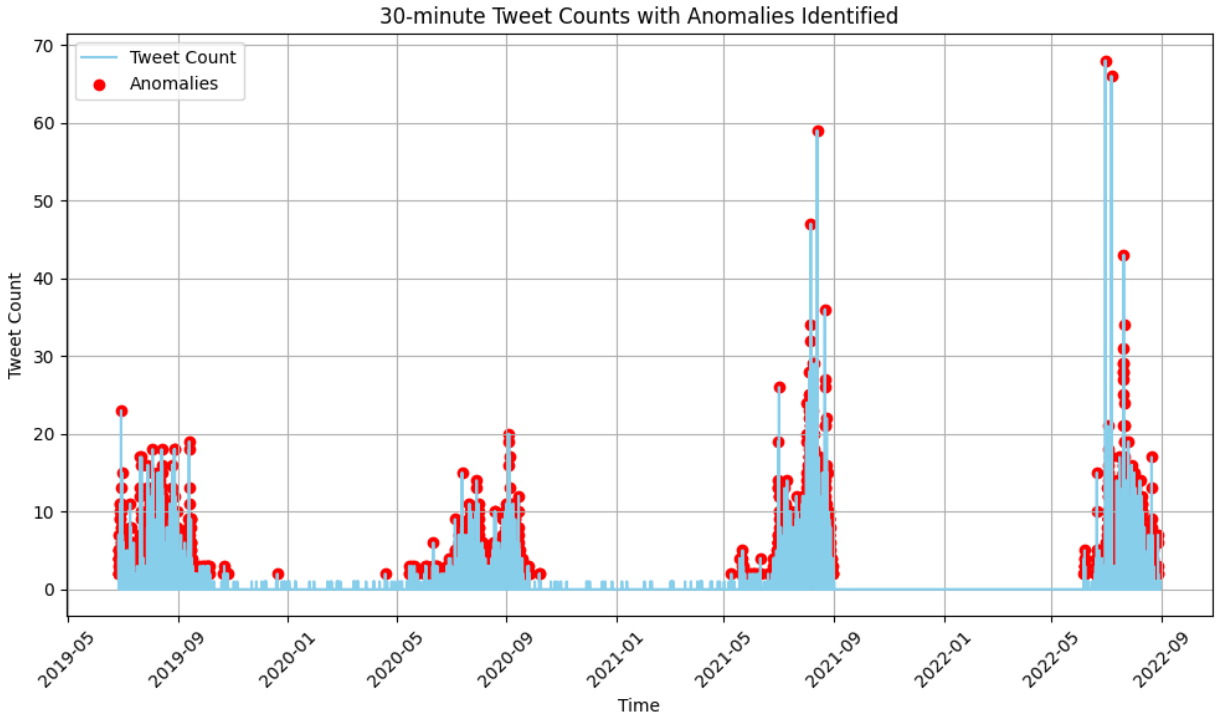


Figure 14: Results for Isolation Forest method

5.3.3 X Crawler

CERTH has developed an almost real-time X crawler using Python and the Tweepy[31] library, which simplifies interaction with X endpoints. Specifically, the crawler connects to X 's Search Tweets endpoint[32], giving access to public data. By leveraging Tweepy, the crawler establishes an open connection and retrieves posts in near real-time through complex queries based on various streaming rule sets[33].

[31] https://www.tweepy.org/
[32] https://developer.x.com/en/docs/x-api/tweets/search/introduction
[33] https://developer.twitter.com/en/docs/twitter-api/tweets/filtered-stream/api-reference/post-tweets-search-stream-rules

To access the Search Tweets endpoint, users need an X account, a developer account on the X portal[34], and at least Free tier access, which is the basic level of API access. A bearer token[35], created through the X developer portal, provides secure API access in X API v2. CERTH has also acquired Basic tier access to benefit from higher usage limits on the platform.

For data collection, the crawler defines specific stream rule sets that include keywords, phrases, accounts, and language filters, along with additional query parameters. These parameters enable the retrieval of new posts that match the defined rules and contain relevant information. The crawler uses four main query parameters (see Table 17):

- **Tweet.fields**: Grants access to details about the post, such as the post ID, timestamp, language, text, media attachments, referenced posts, entities (like hashtags or mentions), and whether the post is potentially sensitive.
- **Media.fields**: Provides metadata about any attached media, including media type, URL, and media key.
- **User.fields**: Supplies information about the user who posted the X post, such as their username.
- **Expansions**: Requests related data objects, like media identifiers and referenced posts IDs, for a more comprehensive dataset.

Table 17: X's JSON attributes coming from API

Query Parameter	Attribute Name	Description	Type
tweet.fields	id	A unique identifier for the post	String
	created_at	The time and date when the post was created	Date (ISO8601)
	lang	The language of the post	String
	text	The text of the post	String
	attachments	Specifies the type of attachments (if any) present in this post (image, video)	JSON Object
	referenced_tweets	A list of posts that this post refers to. For example, if the parent post is a Repost , it will include the related post that is referenced to by its parent.	Array
	entities	Contains tags that are extracted from the text and have special meaning	JSON Object
	possible_sensitive	Indicates if this post contains sensitive information	Boolean
media.fields	type	The type of media contained in the post (animated_gif, photo, video)	String

[34] https://developer.twitter.com/en/portal/dashboard
[35] https://developer.twitter.com/en/docs/authentication/oauth-2-0/bearer-tokens

	url	A direct URL to the media file on X.	String
	media_key	A unique identifier for the media contained in the post	String
user_fields	name	The username of that the user defined in X profile	String
expansions	attachments.media_keys	List of unique identifiers of the media attached to this post	String
	referenced_tweets.id	The unique identifier of the referenced posts	String

The crawler applies a set of rules within the Recent Search endpoint, which may involve one or more types of operators[36]. The four rule operators used by the crawler are:

- Keyword: Fetches posts containing a specific word in the post text (e.g., posts with the keyword "wildfire").
- Exact phrase match: Retrieves posts that include an exact phrase (e.g., posts with the phrase "forest fire").
- From: Retrieves posts from a specific user account (e.g., posts from a designated X user1).
- lang: Fetches posts classified by X as being in a particular language (e.g., posts in English).

These operators can be combined using logical tools:

- AND logic: A space between operators implies Boolean "AND".
- OR logic: Using "OR" creates Boolean "OR" logic.
- Not logic, negation: Adding a dash (-) before an operator applies Boolean "NOT" logic.
- Grouping: Parentheses can be used to group operators together.

For example, the rule (wildfire ("windy day" OR "hot day")) (from:User1) (lang:en) retrieves posts from X user @User1 in English that contain the keyword "wildfire" and either of the exact phrases “windy day” or “hot day” in the post text.

Once the crawler retrieves posts in JSON format, it extracts the relevant information from the JSON attributes. Afterward, the content undergo analysis which includes services such as Location Extraction (to identify locations mentioned in the post) Relevance Estimation (to assess the use case). Finally, after analysis, the posts and their extracted data are stored in a MongoDB database in JSON format.

5.3.4 X Policy Risks

It is crucial to consider the potential legal, financial, and reputational impacts that could result from the policies of X[37]. Like other social media platforms, X must address complex legal and ethical challenges associated with user-generated content, such as hate speech, harassment, and misinformation. If these issues are not adequately managed, they can lead to various consequences, including lawsuits and regulatory penalties.

[36] https://developer.x.com/en/docs/x-api/v1/rules-and-filtering/search-operators
[37] https://developer.twitter.com/en/developer-terms/agreement-and-policy

For a crawler accessing X, policy risks mainly involve potential violations of X's terms of service, including the risk of account suspension or legal repercussions. Specifically, X's terms of service prohibit automated data scraping or collection without proper authorization through its official API. In our case, all data collection is conducted through the official API endpoint provided by X.

X's data storage policy no longer imposes restrictions on how long posts can be stored and analysed. Although previous guidelines limited data retention to six months (Article I.F.2.b.ii), these restrictions have since been lifted.

X's data usage policy authorizes the use of pre-processed data derived from tweet content, including text and images. At CERTH, all analyses conducted through the Social Media Analysis modules focus solely on textual content, ensuring that no potentially harmful information is generated. As a result, we fully comply with X's policies.

To adhere to European GDPR regulations[38] and safeguard privacy, CERTH does not store, display, or share any personal user information, such as usernames, addresses, or other sensitive details. This measure ensures that private information about X's users is not compromised.

Overall, a crawler can operate within X's policy guidelines as long as the developer adheres to the API rules and terms of service. Additionally, any data collected must be used ethically and responsibly, in accordance with relevant laws and regulations.

5.4 Analysis Modules

The analysis modules developed by CERTH aim to enhance the utility of social media data for various applications, particularly in the context of event detection. Given the limitations of social media platforms, such as the sparse availability of geolocation data and the challenge of filtering relevant content, CERTH has created specialized modules to address these gaps. These modules include a location extraction component that enhances geospatial accuracy by identifying and mapping location references within social media text, and a relevance estimation component that uses AI and NLP techniques to determine the significance of posts related to specific events, such as wildfires.

5.4.1 Location Extraction

Social media platforms often lack essential geolocation information for posts, which limits their utility for tasks such as linking satellite data to social media content or pinpointing events on a map. The X API, for instance, includes geolocation data for only a small fraction of posts. To address this limitation, CERTH has developed a localization module aimed at enhancing geolocation capabilities. This module searches for location references within the text of social media posts and maps them to precise coordinates based on the World Geodetic System (WGS 84). Initially created as part of the European EOPEN project, the module has since been further refined by CERTH.

[38] https://gdpr.eu/

Specifically, the localization module processes text by pre-processing it and feeding it into a Long Short-Term Memory (LSTM) network. The network assigns Named Entity Recognition (NER) labels to any identified locations, which are then used to query the OpenStreetMap API. The API returns the geolocation data, providing the exact coordinates for each identified place.

The module is available as a standalone web service, enabling its seamless integration into various systems. The service accepts as input social media post text and outputs a JSON Object Array containing the full name of the location, the type of coordinate reference system used, and its precise coordinates. This geolocation data is then appended as an attribute to the JSON representation of the social media posts gathered by the crawler, enriching as many posts as possible with location information. An example of the JSON structure can be found in Figure 15.


```
"extracted_locations": [
    {
     "placename": "Jurien Bay, Shire Of Dandaragan, Western Australia,
6516, Australia",
      "crs": "WGS84"
      "geometry": {
        "type": "Point",
        "coordinates": [
             115.0406027,
             -30.3040478
        }]
      }
    }
]
```


Figure 15: CERTH's location extraction JSON output.

6.4.2 Relevance Estimation

Effectively filtering textual data from platforms like X can significantly enhance the efficiency of response efforts by authorities. CERTH has developed a relevance estimation module that utilizes Artificial Intelligence (AI) and Natural Language Processing (NLP) to assess the relevance of posts to wildfire events.

AI and NLP are expansive research areas, encompassing numerous models designed for a wide range of tasks. The intersection of these fields often involves leveraging Deep Learning (DL) and Machine Learning (ML) techniques, particularly for text-related applications. Within the SILVANUS framework, relevance estimation is addressed as a supervised classification problem. Here, "supervised" refers to the training approach, where the model is trained on labelled datasets. These datasets are annotated with binary labels: 0 for irrelevant posts and 1 for relevant ones. The developed model uses this training data to predict the likelihood that an unseen post is related to wildfire or not.

To demonstrate the module's functionality, consider two hypothetical examples containing keywords similar to those encountered during data collection. A relevant post might state, "The fire is burning the whole forest right now!", while an irrelevant one could be, "Arson's rap is on fire!". The model outputs a probability distribution for both outcomes, indicating the likelihood of a post being irrelevant and, conversely, the likelihood of it being relevant. It is worth noting that these examples are fictional and may differ from actual posts, especially in terms of length and typical social media features like URLs, hashtags, or emoticons.

Ultimately, the aim of this approach is to filter content from social platforms effectively, ensuring that only wildfire-relevant information is presented to users. The module will be designed to process posts in English.

Relevance estimation module can be used as a standalone web service, making it easy to integrate into existing systems. All that's required is a text of a social media post and the service will return a JSON string that shows if the post text is related with fire incidents and a score. An example of the returned JSON string is shown below (Figure 16):

```
output: {
          fire_related: true,
          score: 0.92
        }
```

Figure 16: An example of relevance estimation service output.

5.4.3 Event Detection Module

The Event Detection module plays a crucial role in identifying potential incidents by harnessing social media data. This system employs a combination of temporal and spatial clustering techniques to group posts, enabling near real-time event detection across various geographic areas. By analysing both the timestamps and locations of posts, the module ensures that fire events are identified based on the most current information available.

Clustering is a fundamental aspect of the event detection process, allowing the system to group social media posts according to their temporal and spatial proximity. This approach is vital for identifying and analysing potential fire incidents using social media data. The workflow begins with retrieving relevant data from MongoDB. A key component of this workflow is dynamically determining the date range for analysis, focusing on data from the previous month. This strategy ensures that the system operates with the most up-to-date information, enhancing its ability to respond quickly to real-time events. Utilizing current data increases the system's sensitivity to detecting and analysing potential incidents.

The clustering phase is critical, as it groups posts based on their temporal and spatial proximity to identify potential incidents. This is achieved using the DBSCAN algorithm (Deng et al. 2020), configured with parameters like the minimum number of posts per event and the time window size. For example, setting the "minimum posts per event" parameter to 10 ensures that only clusters with at least 10 posts are considered significant enough to indicate a possible fire, reducing false positives and improving detection reliability. The "time window size" parameter dictates the duration for which posts are aggregated for clustering, impacting the granularity of event detection.

Additionally, the clustering process organizes posts into weekly segments based on their chronological order, facilitating easier analysis. Within each segment, the system evaluates both the geographic and temporal proximity of posts. The spatial proximity check groups posts within a 10-kilometer radius, while the temporal check considers posts within the same batch period. Posts meeting both criteria are clustered together, representing potential events. These clusters, containing the specified minimum number of posts, are then formatted into JSON for further analysis.

This methodology combines temporal batching, spatial proximity checks, and clustering, ensuring efficient processing of large volumes of data while accurately identifying relevant events based on their spatial and temporal context. Subsequently, the system extracts

keywords from each identified cluster to discern the dominant themes and topics associated with potential fire incidents. This keyword extraction, carried out by a specialized module, identifies and ranks relevant keywords, providing valuable context and insights into the nature and scope of detected incidents.

Throughout the detection process, the system continuously monitors its performance, logging execution times for clustering to maintain efficiency and deliver timely insights into emerging fire events. Implemented in Python, the system operates through a Flask-based RESTful API, enabling it to receive HTTP GET requests and initiate the detection process every 15 minutes. This configuration allows for remote triggering and ensures a prompt response to new data inputs. Once fire events are identified, they are stored along with their associated metadata in MongoDB.

5.4.4 LLMs and context extraction

Integrating social media data with large language models (LLMs) offers an exciting and potentially effective way to analyse disasters in real time. By using Retrieval-Augmented Generation (RAG) (Lewis et al., 2021) and prompt engineering techniques (Chen et al., 2024), these models can keep up-to-date and tackle key questions like identifying the hardest-hit areas from social media posts, finding user-shared media of the damage (e.g., images in the posts), tracking trending hashtags, and figuring out the main worries of those affected.

There are many tools to make the most of open-source LLMs for this work. LangChain[39], for instance, simplifies setting up LLMs quickly in real-world situations, enabling efficient integration with various data sources and applications. OpenWebUI[40] provides a user-friendly platform that supports various LLMs, making it easier to manage data from social media. Platforms like Ollama[41] are great for getting open-source LLMs running on local systems with ease. Using open-source models like LLaMA 3.0 (Touvron et al., 2023) and Mistral 7B[42] not only means more transparency but also allows for tailoring disaster response systems without relying on proprietary or closed models (e.g., chatgpt).

In the sequel, we provide a proof-of-concept (PoC) example using OpenWebUI's RAG system and Ollama with Meta's Llama 3 model, offering insights into the disaster's impact from social media data alone. Specifically, the model is tasked with answering the following questions using input data retrieved from social media.

- Identify the most affected regions mentioned in tweets concerning the disaster.
- List any user-generated videos or photos showcasing the disaster's damage.
- Summarize the trending hashtags and main topics related to the disaster over the past X hours.
- Highlight the predominant concerns or needs frequently mentioned by users affected by the disaster.

The answers to these questions are displayed in Figure 17**Fehler! Verweisquelle konnte nicht gefunden werden.**, beneath each question. Regarding the dataset used for this PoC, it

[39] https://www.langchain.com/
[40] https://openwebui.com/
[41] https://ollama.com/
[42] https://mistral.ai/news/announcing-mistral-7b/

includes a curated subset of 20,000 social media posts collected from X (a small sample of the dataset is shown in Figure 18). These posts gathered based on keywords related to fire danger and wildfires (Figure 19), cover the period from May 1 to August 31 across the years 2019 to 2022. The primary geographical focus is Greece, including Athens and surrounding areas (Figure 20).

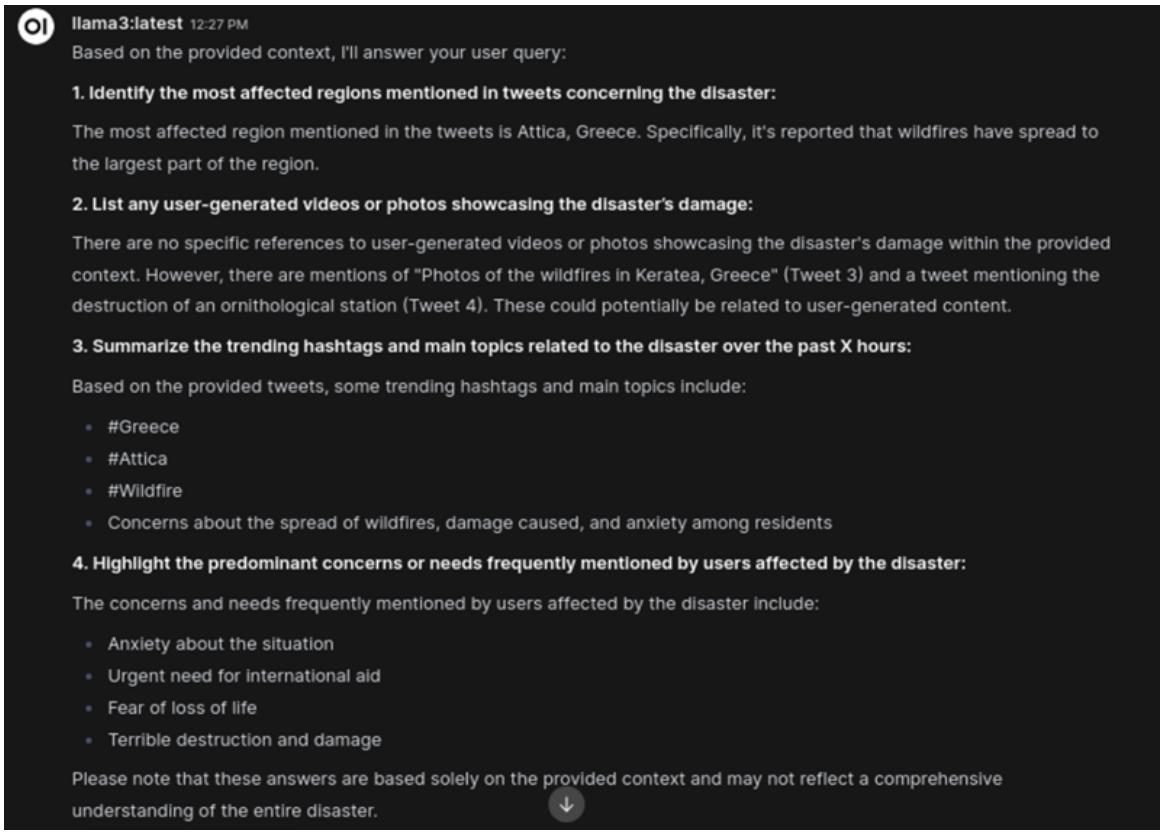


Figure 17: PoC of using the Llama3 model with social media data to answer questions strictly based on the information available, without going beyond the given social media context. The data consists of aggregated tweets concerning wildlife events around Attica, near Athens, Greece.

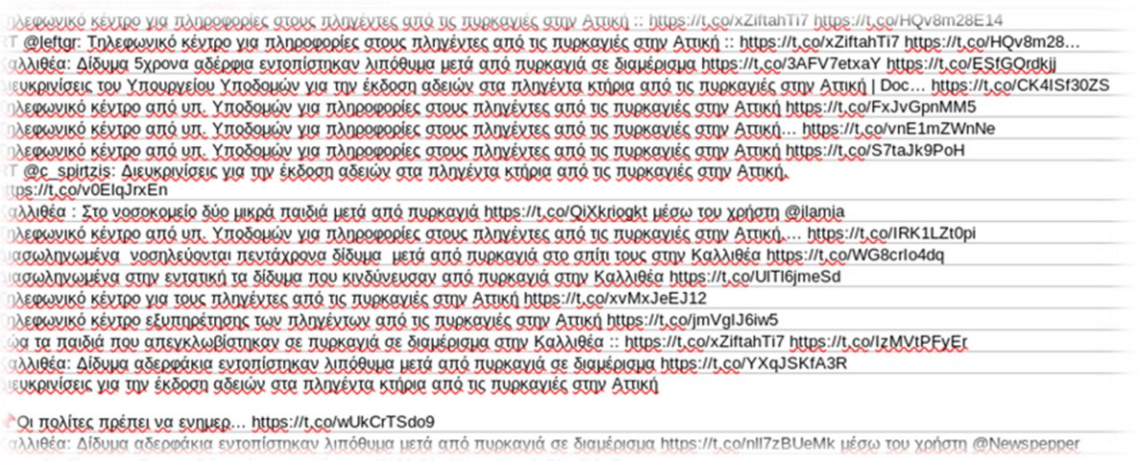


Figure 18: Sample of social media data used for the PoC.

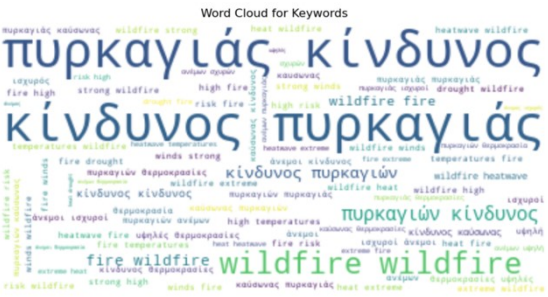


Figure 19: Keywords found in the social media dataset.

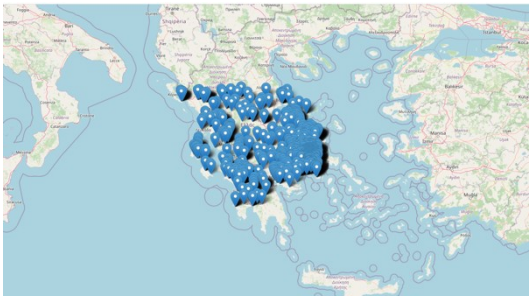


Figure 20: Location of the social media posts found in the social media dataset.

6.Conclusions

This deliverable presents a comprehensive framework for enhancing climate resilience by integrating diverse data sources, advanced methodologies, and innovative tools. Key components include the collection, analysis, and integration of open datasets and climate services, the development of indices based on satellite and meteorological data, and the application of state-of-the-art techniques for hazard and vulnerability mapping.

The exploration of drought, heatwave, wildfire, and flood susceptibility underscores the importance of high-resolution data and tailored approaches to assessing regional climate hazards. Satellite-derived indices and meteorological indicators have been leveraged to address data gaps, enabling more accurate and actionable insights for decision-making. Furthermore, the development of **spatial downscaling methods for Copernicus satellite imagery has resolved challenges in spatial resolution, offering improved datasets suitable for localized climate risk assessments**.

The inclusion of **social media crawling** as a real-time data collection strategy enriches this deliverable by providing **dynamic insights into the impacts of climatic hazards**. This innovative approach leverages user-generated content to support hazard monitoring and response planning, particularly in rapidly evolving situations.

The integration of these approaches and datasets aligns with the project's overarching objectives of addressing data gaps, improving stakeholder usability, and fostering long-term climate adaptation solutions. The methodologies and tools outlined here contribute directly to ClimEmpower's mission by equipping regions with the resources and knowledge required to effectively monitor, assess, and respond to climate hazards.

Looking ahead, the approaches and technologies outlined in this deliverable pave the way for scalable and transferable solutions that can support climate resilience projects across Europe and beyond. By refining these methodologies and fostering stakeholder collaboration, this work makes a meaningful contribution to building a more sustainable and climate-resilient future.

This deliverable in combination with deliverable 2.1 and respect to existing climate services allow a better understanding of climate impacts which is valuable for Deliverable 2.4 to evaluate the effectiveness of adaptation and mitigation measures. Additionally, as for the WP3, methodology of spatial data downscaling offers a foundational service which provides high resolution spatial data and analysis of social media data integrates into on demand service in disaster monitoring and event detection. All of these outputs (hazard indices, real time monitoring data) could be components in HTML5 application and ensuring that end users can take input from these processed data without technical expertise. Finally, as for WP4, high resolution spatial data, hazard mapping and social sensing can be shared with CoPs to enhance the regional climate challenges. All these will be assessed by trials to ensure user satisfaction.

References

Abdikan, S., Bayik, C., Sekertekin, A., Bektas Balcik, F., Karimzadeh, S., Matsuoka, M., & Balik Sanli, F. (2022). Burnt area detection using multi-sensor SAR, optical, and thermal data in Mediterranean pine forest. Forests, 13(2), 347.

Abdo HG, Almohamad H, Al Dughairi AA, Al-Mutiry M. (2022). GIS-Based Frequency Ratio and Analytic Hierarchy Process for Forest Fire Susceptibility Mapping in the Western Region of Syria. *Sustainability*. 2022; 14(8):4668. https://doi.org/10.3390/su14084668

Abedi Gheshlaghi, H., Feizizadeh, B., Blaschke, T. (2020). GIS-based forest fire risk mapping using the analytical network process and fuzzy logic. *J. Environ. Plan. Manag.* **2020**, *63*, 481–499.

Adams, H. D., Guardiola-Claramonte, M., Barron-Gafford, G. A., Villegas, J. C., Breshears, D. D., Zou, C. B., ... & Huxman, T. E. (2009). Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proceedings of the national academy of sciences*, *106*(17), 7063-7066.

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. Reviews of Geophysics, 53(2), 452-480.

Alademomi, A. S., Okolie, C. J., Daramola, O. E., Akinnusi, S. A., Adediran, E., Olanrewaju, H. O., et al. (2022). The interrelationship between LST, NDVI, NDBI, and land cover change in a section of Lagos metropolis, Nigeria. *Applied Geomatics*, *14*(2), 299-314.

Almouctar, M. A. S., Wu, Y., Zhao, F., & Qin, C. (2024). Drought analysis using normalized difference vegetation index and land surface temperature over Niamey region, the southwestern of the Niger between 2013 and 2019. Journal of Hydrology: Regional Studies, 52, 101689.

Aznar-Siguan, G., & Bresch, D. N. (2019). CLIMADA v1: a global weather and climate risk assessment platform. Geoscientific Model Development, 12(7), 3085-3097.

Bentekhici, N., Bellal, S. A., & Zegrar, A. (2020). Contribution of remote sensing and GIS to mapping the fire risk of Mediterranean forest case of the forest massif of Tlemcen (North-West Algeria). *Natural Hazards*, *104*(1), 811-831.

Betterle, A., & Salamon, P. (2024). Water depth estimate and flood extent enhancement for satellite-based inundation maps. *Natural Hazards and Earth System Sciences*, *24*(8), 2817-2836.

Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, *58*(3), 257-266.

Borana, S. L., Yadav, S. K., Parihar, S. K., & Kaplan, G. (2019). PanSharpening of Landsat-8 Using Synthetic Sentinel-2 PAN Data. In Esri India User Conference (pp. 1-11).

Breshears, D. D., Cobb, N. S., Rich, P. M., Price, K. P., Allen, C. D., Balice, R. G., ... & Meyer, C. W. (2005). Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Sciences*, *102*(42), 15144-15148.

Brewster, J. (2009). Development of the Flash Flood Potential Index. http://www.erh.noaa.gov/bgm/research/ERFFW/presentations/june_02_2010/Brewster_Jim_Development_of_FFPI.ppt. 2009

Busico, G., Giuditta, E., Kazakis, N., & Colombani, N. (2019). A hybrid GIS and AHP approach for modelling actual and future forest fire risk under climate change accounting water resources attenuation role. *Sustainability*, *11*(24), 7166.

Ceru, J. (2012). The Flash Flood Potential Index for Pennsylvania. Retrieved from https://proceedings.esri.com/library/userconf/feduc12/papers/user/joeceru.pdf

Che, X., Feng, M., Jiang, H., Song, J., & Jia, B. (2015). Downscaling MODIS surface reflectance to improve water body extraction. Advances in Meteorology, 2015(1), 424291.

Chen, B., Zhang, Z., Langrené, N., & Zhu, S. (2024). *Unleashing the potential of prompt engineering in Large Language Models: A comprehensive review* (arXiv:2310.14735). arXiv. http://arxiv.org/abs/2310.14735

Cheng, Y., Zhang, K., Chao, L., Shi, W., Feng, J., & Li, Y. (2023). A comprehensive drought index based on remote sensing data and nested copulas for monitoring meteorological and agroecological droughts: A case study on the Qinghai-Tibet Plateau. *Environmental Modelling & Software*, *161*, 105629. https://doi.org/10.1016/j.envsoft.2023.105629

Choi, M. (2006). A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter. IEEE Transactions on Geoscience and Remote Sensing, 44(6), 1672–1682. doi:10.1109/TGRS.2006.869923.

Chuvieco, E., Martin, M. P., & Palacios, A. (2002). Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, *23*(23), 5103-5110.

Cocke, A. E., Fulé, P. Z., & Crouse, J. E. (2005). Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data. *International Journal of Wildland Fire*, *14*(2), 189-198.

De Simone, W., Di Musciano, M., Di Cecco, V., Ferella, G., & Frattaroli, A. R. (2020). The potentiality of Sentinel-2 to assess the effect of fire events on Mediterranean mountain vegetation. *Plant Sociology*, *57*, 11-22.

Deng, D. (2020, September). DBSCAN clustering algorithm based on density. In *2020 7th international forum on electrical engineering and automation (IFEEA)* (pp. 949-953). IEEE.

Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307.

Dong, W., Yang, Y., Qu, J., Xiao, S., & Du, Q. (2021). Hyperspectral pansharpening via local intensity component and local injection gain estimation. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.

Du, Q., Younan, N. H., King, R., & Shah, V. P. (2007). On the performance evaluation of pan-sharpening techniques. IEEE Geoscience and Remote Sensing Letters, 4(4), 518-522.

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sensing, 8(4), 354.

Durlević, U., Novković, I., Lukić, T., Valjarević, A., Samardžić, I., Krstić, F., Batoćanin, N., Mijatov, M. & Ćurić, V. (2021). Multihazard susceptibility assessment: A case study – Municipality of Štrpce (Southern Serbia). *Open Geosciences*, *13*(1), 1414-1431. https://doi.org/10.1515/geo-2020-0314

Elkhrachy, I. (2015). Flash flood hazard mapping using satellite images and GIS tools: a case study of Najran City, Kingdom of Saudi Arabia (KSA). *The Egyptian Journal of Remote Sensing and Space Science*, *18*(2), 261-278.

Kogan, F., & Sullivan, J. (1993). Development of global drought-watch system using NOAA/AVHRR data. *Advances in Space Research*, *13*(5), 219-222.

Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. Advances in space research, 15(11), 91-100.

Feng, L., Hu, C., Chen, X., Cai, X., Tian, L., & Gan, W. (2012). Assessment of inundation changes of Poyang Lake using MODIS observations between 2000 and 2010. Remote Sensing of Environment, 121, 80-92.

Fraisse, C. W., Breuer, N. E., & Zierden, D. (2011). Drought Decision-Support Tools: Introducing the Keetch Byram Drought Index—KBDI1. *University of Florida*.

Gannon, C. S., & Steinberg, N. C. (2021). A global assessment of wildfire potential under climate change utilizing Keetch-Byram drought index and land cover classifications. *Environmental Research Communications*, *3*(3), 035002.

Gašparović, M., & Jogun, T. (2017). The effect of fusing Sentinel-2 bands on land-cover classification. International journal of remote sensing, 39(3), 822-841.

Ge, W., Cheng, Q., Jing, L., Wang, F., Zhao, M., & Ding, H. (2020). Assessment of the capability of sentinel-2 imagery for iron-bearing minerals mapping: A case study in the cuprite area, Nevada. Remote Sensing, 12(18), 3028.

Ghosh, A., & Kar, S. K. (2018). Application of analytical hierarchy process (AHP) for flood risk assessment: a case study in Malda district of West Bengal, India. *Natural Hazards*, *94*, 349-368.

Grillakis, M. G., Koutroulis, A. G., Komma, J., Tsanis, I. K., Wagner, W., & Blöschl, G. (2016). Initial soil moisture effects on flash flood generation–A comparison between basins of contrasting hydro-climatic conditions. *Journal of Hydrology*, *541*, 206-217. https://doi.org/10.1016/j.jhydrol.2016.03.007

Grochala, A., & Kedzierski, M. (2017). A method of panchromatic image modification for satellite imagery data fusion. Remote sensing, 9(6), 639.

Hashim, F., Dibs, H., & Jaber, H. S. (2022). Adopting gram-schmidt and brovey methods for estimating land use and land cover using remote sensing and satellite images. Nature Environment and Pollution Technology, 21(2), 867-881.

Hayes, M. J., Svoboda, M. D., Wardlow, B. D., Anderson, M. C., & Kogan, F. (2012). Drought monitoring: Historical and current perspectives.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, *83*(1-2), 195-213. https://doi.org/10.1016/S0034-4257(02)00096-2

Matthews, J. R. (2018). Annex I: glossary. *Global warming of*, *1*, 541-562.

Balch, J. K., Abatzoglou, J. T., Joseph, M. B., Koontz, M. J., Mahood, A. L., McGlinchy, J., ... & Williams, A. P. (2022). Warming weakens the night-time barrier to global fire. *Nature*, *602*(7897), 442-448.

Jang, J. D., Viau, A. A., & Anctil, F. (2006). Thermal-water stress index from satellite images. *International Journal of Remote Sensing*, *27*(8), 1619-1639.

Jawak, S. D., & Luis, A. J. (2013). A comprehensive evaluation of PAN-sharpening algorithms coupled with resampling methods for image synthesis of very high resolution remotely sensed satellite data. Advances in Remote Sensing, 2013.

Brown, J. F., Wardlow, B. D., Tadesse, T., Hayes, M. J., & Reed, B. C. (2008). The Vegetation Drought Response Index (VegDRI): A new integrated approach for monitoring drought stress in vegetation. *GIScience & Remote Sensing*, *45*(1), 16-46. https://doi.org/10.2747/1548-1603.45.1.16

Ji, L., & Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote sensing of Environment*, *87*(1), 85-98. https://doi.org/10.1016/S0034-4257(03)00174-3

Christian, J. I., Basara, J. B., Lowman, L. E., Xiao, X., Mesheske, D., & Zhou, Y. (2022). Flash drought identification from satellite-based land surface water index. *Remote Sensing Applications: Society and Environment*, *26*, 100770.

Kaplan, G., & Avdan, U. (2018). Sentinel-2 pan sharpening—comparative analysis. In Proceedings (Vol. 2, No. 7, p. 345). MDPI.

Karimiziarani, M., Jafarzadegan, K., Abbaszadeh, P., Shao, W., & Moradkhani, H. (2022). Hazard risk awareness and disaster management: Extracting the information content of twitter data. *Sustainable Cities and Society*, *77*, 103577.

Keeley, J. E. (2009). Fire intensity, fire severity and burn severity: a brief review and suggested usage. *International journal of wildland fire*, *18*(1), 116-126. https://doi.org/10.1071/WF07049

Keetch, J. J., & Byram, G. M. (1968). *A drought index for forest fire control* (Vol. 38). US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station.

King, R. L., & Wang, J. (2001, July). A wavelet based algorithm for pan sharpening Landsat 7 imagery. In *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217)* (Vol. 2, pp. 849-851). IEEE.

Kostopoulou, E., & Giannakopoulos, C. (2023). Projected Changes in Extreme Wet and Dry Conditions in Greece. *Climate*, *11*(3), 49. https://doi.org/10.3390/cli11030049

Kruzdlo, R., & Ceru, J. (2010). *National Oceanic and Atmospheric Administration. Retrieved from Flash Flood Potential Index for WFO Mount Holly/Philadelphia*.

Kumar, L., Sinha, P., & Taylor, S. (2014). Improving image classification in a complex wetland ecosystem through image fusion techniques. Journal of Applied Remote Sensing, 8(1), 083616-083616.

Kuntla, S. K. (2021). An era of Sentinels in flood management: Potential of Sentinel-1,-2, and-3 satellites for effective flood management. *Open Geosciences*, *13*(1), 1616-1642.

Kutchartt, E., González-Olabarria, J. R., Aquilué, N., Garcia-Gonzalo, J., Trasobares, A., Botequim, B., et al. (2024). Pan-European fuel map server: an open-geodata portal for supporting fire risk assessment. *Geomatica*, 100036.

Levin, N., & Phinn, S. (2022, January). Assessing the 2022 Flood Impacts in Queensland Combining Daytime and Nighttime Optical and Imaging Radar Data. Remote Sensing, 14(19), 5009.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.. http://arxiv.org/abs/2005.11401

Lin, L., Di, L., Yu, E. G., Kang, L., Shrestha, R., Rahman, M. S., ... & Hu, L. (2016, July). A review of remote sensing in flood assessment. In *2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* (pp. 1-4). IEEE.

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth ieee international conference on data mining* (pp. 413-422). IEEE.

Liu, X., Zhu, X., Pan, Y., Li, S., Liu, Y., & Ma, Y. (2016). Agricultural drought monitoring: Progress, challenges, and prospects. *Journal of Geographical Sciences*, *26*, 750-767.

Diakakis, M., Deligiannakis, G., Katsetsiadou, K., Lekkas, E., Melaki, M., & Antoniadis, Z. (2016). Mapping and classification of direct effects of the flood of October 2014 in Athens. *Bulletin of the Geological Society of Greece*, *50*(2), 681-690., 10.12681/bgsg.11774

Mahmoud, A. G. (2021). Machine learning and pan-sharpening of Sentinel-2 data for land use mapping in arid regions: a case study in Fayoum, Egypt. Journal of Soil Sciences and Agricultural Engineering, 12(11), 717-723.

Mallick, J., Talukdar, S., Pal, S., & Rahman, A. (2021). A novel classifier for improving wetland mapping by integrating image fusion techniques and ensemble machine learning classifiers. Ecological Informatics, 65, 101426.

Masi, G., Cozzolino, D., Verdoliva, L., & Scarpa, G. (2016). Pansharpening by convolutional neural networks. Remote Sensing, 8(7), 594.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, *17*(7), 1425-1432.

McKee, T. B., Doesken, N. J., & Kleist, J. (1993, January). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, No. 22, pp. 179-183).

Mhangara, P., Mapurisa, W., & Mudau, N. (2020). Comparison of image fusion techniques using satellite pour l'Observation de la Terre (SPOT) 6 satellite imagery. Applied Sciences, 10(5), 1881.

Mishra, A. K., Singh, V. P., & Desai, V. R. (2009). Drought characterization: a probabilistic approach. *Stochastic Environmental Research and Risk Assessment*, *23*, 41-55.

Mohd Yassin, N. A. M., Adnan, N. A., & Sadek, E. S. M. (2023). Analysis of Flash Flood Potential Index (FFPI) and scenarios assessment in Shah Alam using GIS approach. *Planning Malaysia*, *21*. https://doi.org/10.21837/pm.v21i26.1255

Naumann, G., Spinoni, J., Vogt, J. V., & Barbosa, P. (2015). Assessment of drought damages and their uncertainties in Europe. *Environmental Research Letters*, *10*(12), 124013. **DOI** 10.1088/1748-9326/10/12/124013

Niazi, S., Mokhtarzade, M., & Saeedzadeh, F. (2015). A Novel Ihs-Ga Fusion Method Based on Enhancement Vegetated Area. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 40, 543-548.

Nikolakopoulos, K. G. (2008). Comparison of nine fusion techniques for very high resolution data. Photogrammetric Engineering & Remote Sensing, 74(5), 647-659.

Palmer, W. C. (1965). *Meteorological drought* (Vol. 30). US Department of Commerce, Weather Bureau.

Palsson, F., Sveinsson, J. R., Benediktsson, J. A., & Aanaes, H. (2011). Classification of pansharpened urban satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(1), 281-297.

Pandit, V. R., & Bhiwani, R. J. (2015). Image fusion in remote sensing applications: A review. International journal of computer applications, 120(10).

Park, H., Choi, J., Park, N., & Choi, S. (2017). Sharpening the VNIR and SWIR bands of Sentinel-2A imagery through modified selected and synthesized band schemes. Remote Sensing, 9(10), 1080.

Park Williams, A., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., ... & McDowell, N. G. (2013). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature climate change*, *3*(3), 292-297.

Parks, S. A., Dillon, G. K., & Miller, C. (2014). A new metric for quantifying burn severity: the relativized burn ratio. *Remote Sensing*, *6*(3), 1827-1844.

Perkins, S. E., & Alexander, L. V. (2013). On the measurement of heat waves. *Journal of climate*, *26*(13), 4500-4517.

Perkins-Kirkpatrick, S. E., & Lewis, S. C. (2020). Increasing trends in regional heatwaves. *Nature communications*, *11*(1), 3357.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing*, *12*(14), 2291.

Politi, N., Vlachogiannis, D., Sfetsos, A., Nastos, P. T., & Dalezios, N. R. (2022). High resolution future projections of drought characteristics in Greece based on SPI and SPEI indices. *Atmosphere*, *13*(9), 1468.

Patil, P. P., Jagtap, M. P., Khatri, N., Madan, H., Vadduri, A. A., & Patodia, T. (2024). Exploration and advancement of NDDI leveraging NDVI and NDWI in Indian semi-arid regions: A remote sensing-based study. *Case Studies in Chemical and Environmental Engineering*, *9*, 100573. https://doi.org/10.1016/j.cscee.2023.100573

Pushparaj, J., & Hegde, A. V. (2017). Evaluation of pan-sharpening methods for spatial and spectral quality. *Applied Geomatics*, *9*, 1-12.

Bajgain, R., Xiao, X., Wagle, P., Basara, J., & Zhou, Y. (2015). Sensitivity analysis of vegetation indices to drought over two tallgrass prairie sites. *ISPRS Journal of Photogrammetry and Remote Sensing*, *108*, 151-160.

Restaino, C. M., Peterson, D. L., & Littell, J. (2016). Increased water deficit decreases Douglas fir growth throughout western US forests. *Proceedings of the National academy of Sciences*, *113*(34), 9557-9562.

Richaud Bertrand (2019). Vegetation index, NDVI and VCI indices. UNCCD-led Drought toolbox, Antalya, 2019. Reporting Application (https://www.unccd.int/sites/default/files/inline-files/04_Vegetation_index.pdf)

Sanli, F. B., Abdikan, S., Esetlili, M. T., & Sunar, F. (2017). Evaluation of image fusion methods using PALSAR, RADARSAT-1 and SPOT images for land use/land cover classification. *Journal of the Indian Society of Remote Sensing*, *45*, 591-601.

Sarp, G. (2014). Spectral and spatial quality analysis of pan-sharpening algorithms: A case study in Istanbul. European Journal of Remote Sensing, 47(1), 19-28.

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009). Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. *Reviews of Geophysics*, *47*(4). 10.1029/2008RG000274.

Seager, R., Hooks, A., Williams, A. P., Cook, B., Nakamura, J., & Henderson, N. (2015). Climatology, variability, and trends in the US vapor pressure deficit, an important fire-related meteorological quantity. *Journal of Applied Meteorology and Climatology*, *54*(6), 1121-1141. doi:10.1175/JAMC-D-14-0321.1, 2015.

Sepulcre-Canto, G., Horion, S., Singleton, A., Carrao, H., & Vogt, J. (2012). The European Drought Observatory (EDO): Current state and future developments. **Natural Hazards and Earth System Sciences**, 12(7), 2141–2151. doi:10.5194/nhess-12-2141-2012

Singh, C., Jain, G., Sukhwani, V., & Shaw, R. (2021). Losses and damages associated with slow-onset events: Urban drought and water insecurity in Asia. *Current Opinion in Environmental Sustainability*, *50*, 72-86. https://doi.org/10.1016/j.cosust.2021.02.006

Sheng, Y., Gong, P., & Xiao, Q. (2001). Quantitative dynamic flood monitoring with NOAA AVHRR. *International Journal of Remote Sensing*, *22*(9), 1709-1724.

Smith, A. M. S., Drake, N. A., Wooster, M. J., Hudak, A. T., Holden, Z. A., & Gibbons, C. J. (2007). Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: comparison of methods and application to MODIS. *International Journal of Remote Sensing*, *28*(12), 2753-2775.

Stefanidou, A., Gitas, I. Z., Stavrakoudis, D., & Eftychidis, G. (2019). Midterm fire danger prediction using satellite imagery and auxiliary thematic layers. *Remote Sensing*, *11*(23), 2786. https://doi.org/10.3390/rs11232786

Stoyanova, E. (2023). Remote sensing for flood inundation mapping using various processing methods with Sentinel-1 and Sentinel-2. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 339-346.

Sufi, F. K., & Khalil, I. (2022). Automated disaster monitoring from social media posts using AI-based location intelligence and sentiment analysis. *IEEE Transactions on Computational Social Systems*.

Sun, G. F., Qin, X. L., Liu, S. C., Li, X., Chen, X., & Zhong, X. (2019). Potential analysis of typical vegetation index for identifying burned area. *Remote Sensing for Land and Resources*, *31*(1), 204-211.

Tincu, R., Lazar, G. & Lazar, I. (2018). Modified flash flood potential index in order to estimate areas with predisposition to water accumulation. *Open Geosciences*, *10*(1), 593-606. https://doi.org/10.1515/geo-2018-0047

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models* (arXiv:2302.13971). arXiv. http://arxiv.org/abs/2302.13971

Tsakiris, G., Pangalou, D., & Vangelis, H. (2007). Regional drought assessment based on the Reconnaissance Drought Index (RDI). *Water resources management*, *21*, 821-833.

Tu, T. M., Huang, P. S., Hung, C. L., & Chang, C. P. (2004). A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery. IEEE Geoscience and Remote sensing letters, 1(4), 309-312.

Tu, T. M., Su, S. C., Shyu, H. C., & Huang, P. S. (2001). A new look at IHS-like image fusion methods. Information fusion, 2(3), 177-186.

Shashikant, V., Mohamed Shariff, A. R., Wayayok, A., Kamal, M. R., Lee, Y. P., & Takeuchi, W. (2021). Utilizing TVDI and NDWI to classify severity of agricultural drought in Chuping, Malaysia. *Agronomy*, *11*(6), 1243. 10.3390/agronomy11061243

Varghese, D., Radulović, M., Stojković, S., & Crnojević, V. (2021). Reviewing the potential of Sentinel-2 in assessing the drought. *Remote sensing*, *13*(17), 3355.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, *23*(7), 1696-1718.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, *23*(7), 1696-1718.

Vicente-Serrano, Sergio M. & National Center for Atmospheric Research Staff (Eds). Last modified 2024-04-22 "The Climate Data Guide: Standardized Precipitation Evapotranspiration Index (SPEI)." Retrieved from https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-evapotranspiration-index-spei on 2024-10-17.

Jayawardhana W.& Chathurange M., 2020. Investigate the Sensitivity of the Satellite-Based Agricultural Drought Indices to Monitor the Drought Condition of Paddy and Introduction to Enhanced Multi-Temporal Drought Indices.Journal of Remote Sensing & GIS, Vol.9 Iss.2. DOI: 10.35248/2469-4134.20.9.272

Wald, L., Ranchin, T., & Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. Photogrammetric engineering and remote sensing, 63(6), 691-699.

Takeuchi, W., Darmawan, S., Shofiyati, R., Khiem, M. V., Oo, K. S., Pimple, U., & Heng, S. (2015, October). Near-real time meteorological drought monitoring and early warning system for croplands in asia. In *Asian Conference on Remote Sensing 2015: Fostering Resilient Growth in Asia* (Vol. 1, pp. 171-178).

Wilhite, D. A. (2000). Drought as a natural hazards: concept and definition in: Wilhite. *Drought: A Global Assessment Routledge*.

Wu, G., & Liu, Y. (2015). Downscaling surface water inundation from coarse data to fine-scale resolution: Methodology and accuracy assessment. Remote Sensing, 7(12), 15989-16003.

Xiao, C., Wu, Y., & Zhu, X. (2023). Evaluation of the Monitoring Capability of 20 Vegetation Indices and 5 Mainstream Satellite Band Settings for Drought in Spring Wheat Using a Simulation Method. *Remote Sensing*, *15*(19), 4838. https://doi.org/10.3390/rs15194838

Xiang, K., Yuan, W., Wang, L., & Deng, Y. (2020). An LSWI-based method for mapping irrigated areas in China using moderate-resolution satellite data. *Remote Sensing*, *12*(24), 4181.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

Zhang, S., Bai, M., Wang, X., Peng, X., Chen, A., & Peng, P. (2023). Remote sensing technology for rapid extraction of burned areas and ecosystem environmental assessment. *PeerJ*, *11*, e14557. https://doi.org/10.7717/peerj.14557

Zhang, M., Su, W., Fu, Y., Zhu, D., Xue, J. H., Huang, J., ... & Yao, C. (2019). Super-resolution enhancement of Sentinel-2 image for retrieving LAI and chlorophyll content of summer corn. European Journal of Agronomy, 111, 125938.

Zhang, Y. (2004). Understanding image fusion. Photogramm. Eng. Remote Sens, 70(6), 657-661.

Zheng, H., Du, P., Chen, J., Xia, J., Li, E., Xu, Z., ... & Yokoya, N. (2017). Performance evaluation of downscaling Sentinel-2 imagery for land use and land cover classification by spectral-spatial features. Remote Sensing, 9(12), 1274.

Zope, P. E., Eldho, T. I., & Jothiprakash, V. (2016). Impacts of land use–land cover change and urbanization on flooding: A case study of Oshiwara River Basin in Mumbai, India. *Catena*, *145*, 142-154.. 10.1016/j.catena.2016.06.009.

Appendix

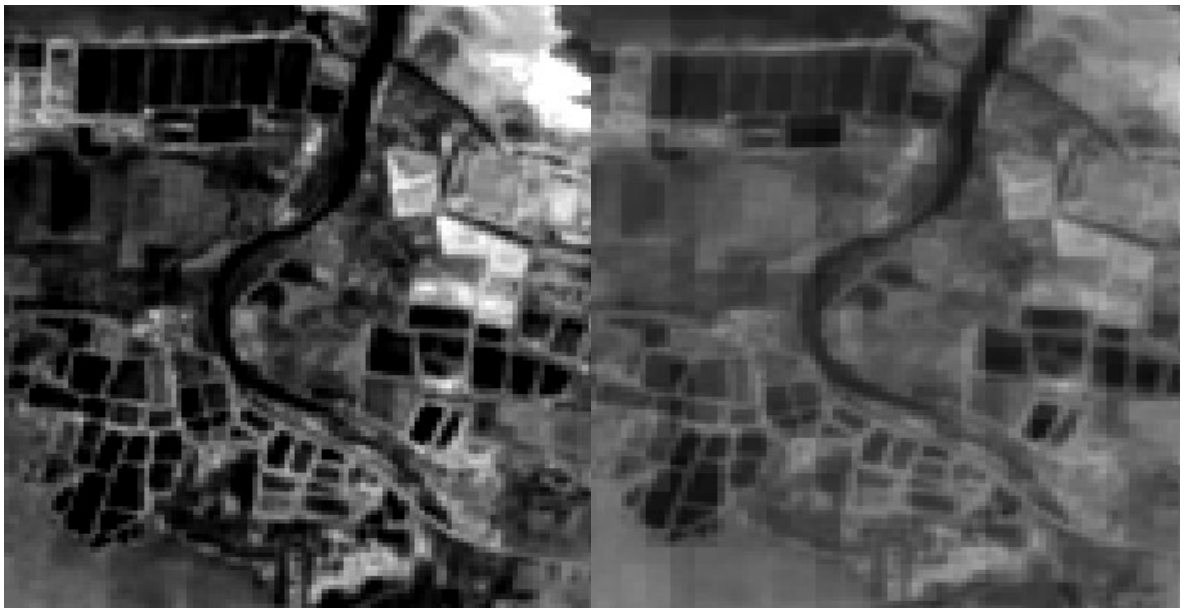
This section contains visual comparisons of the reference image, and the downscaled images processed after various pansharpening techniques. In particular, it includes images after application of the Brovey Transform, the IHS Transform, the Wavelet Transform and the PNN for the red edge bands (5, 6, 7), the NIR band (8A) and the SWIR bands (11, 12). These techniques have been used for the downscaling of Sentinel-2 imagery from 20 m to 10 m spatial resolution.

Red edge bands:

Band 5:

Reference image

Brovey Transform



IHS Transform

Wavelet Transform

PNN

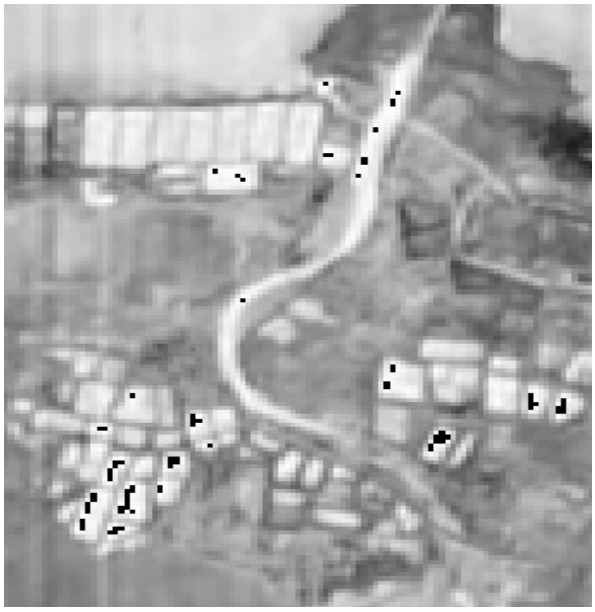
Band 6:

Reference image

Brovey Transform

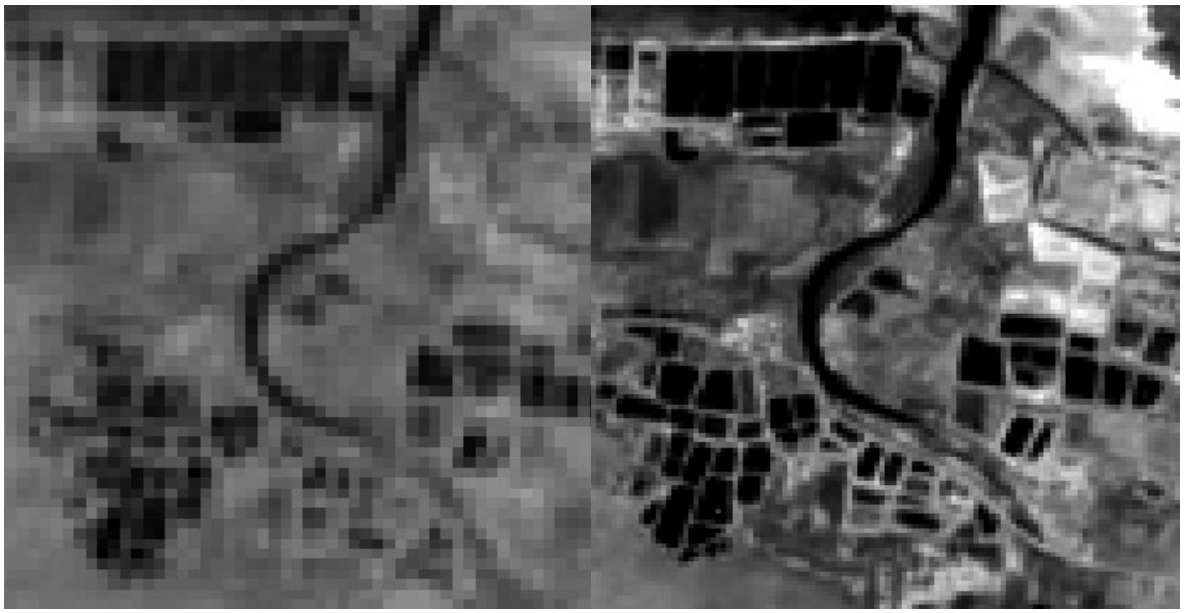
IHS Transform

Wavelet Transform



PNN

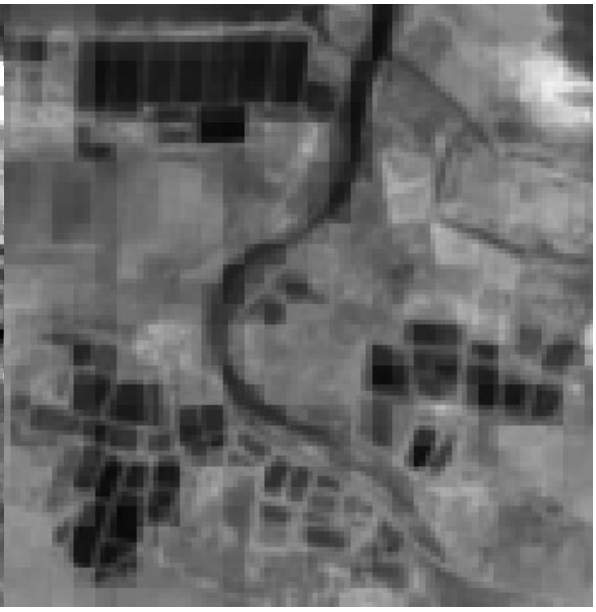
Band 7:



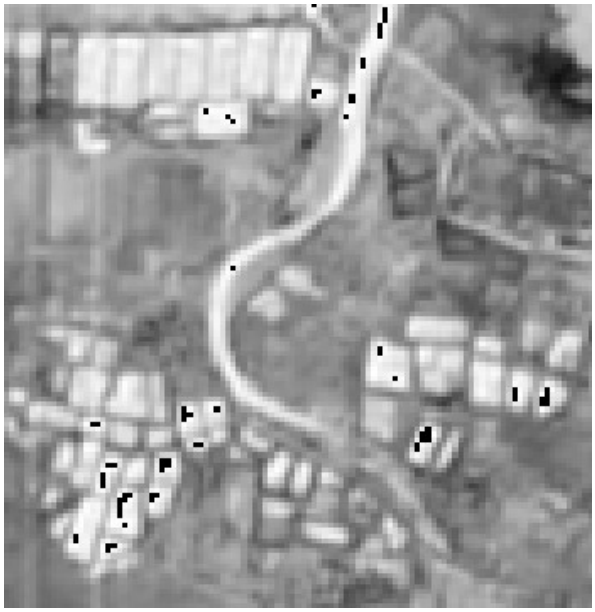
Reference image

Brovey Transform

IHS Transform

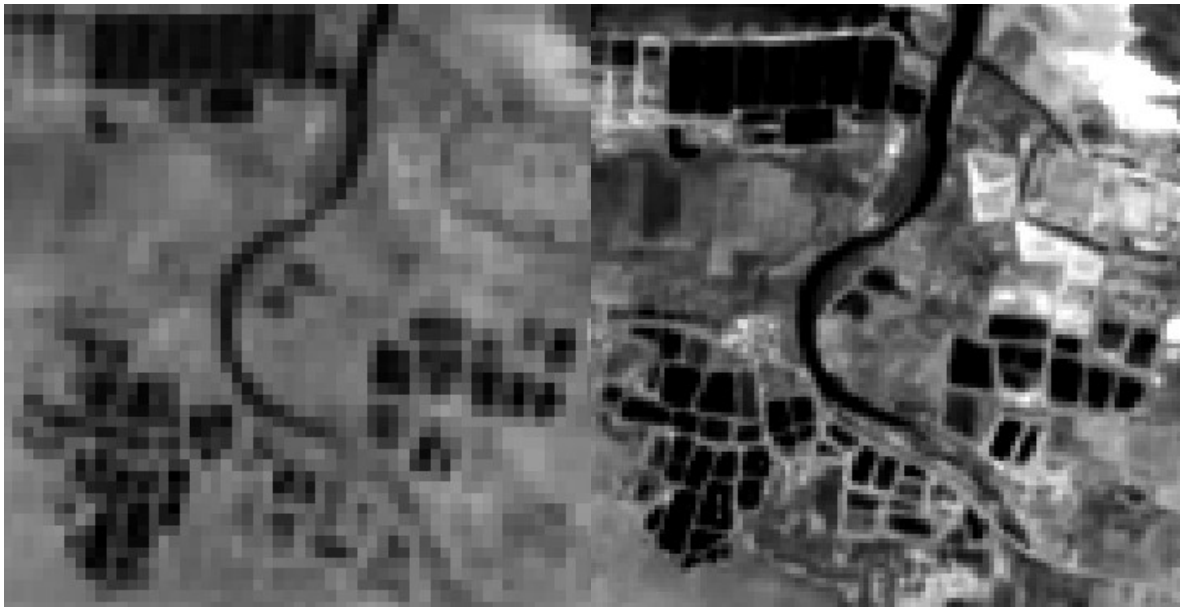


Wavelet Transform



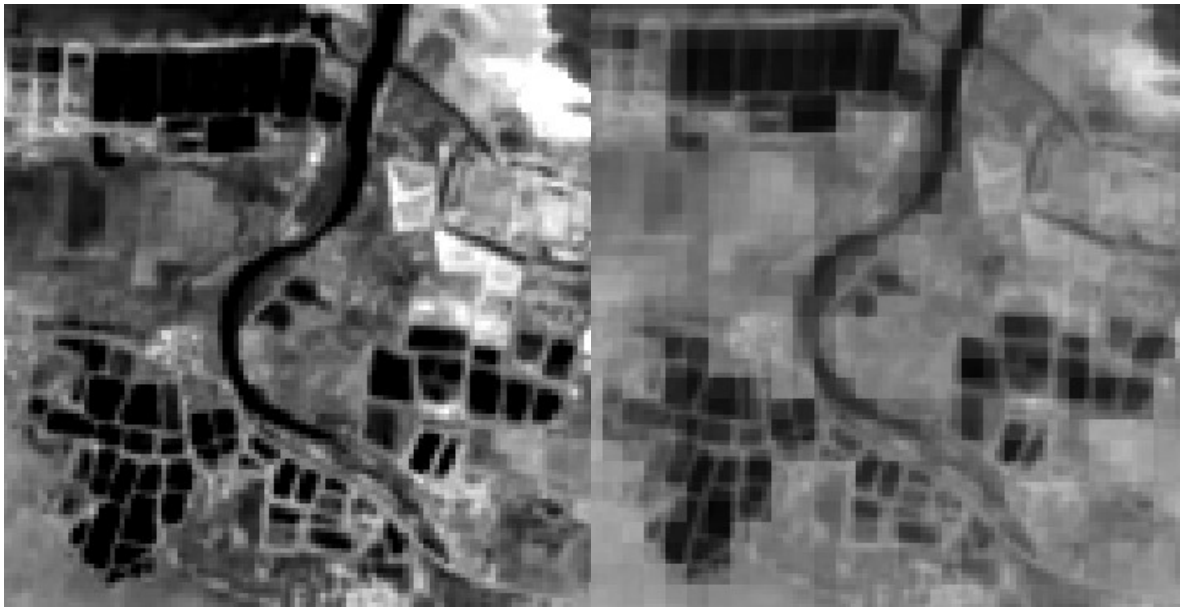
PNN

Band 8A (NIR):



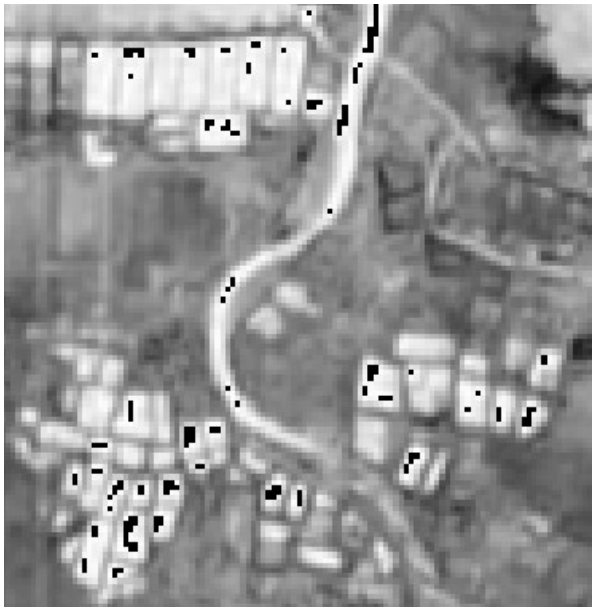
Reference image

Brovey Transform



IHS Transform

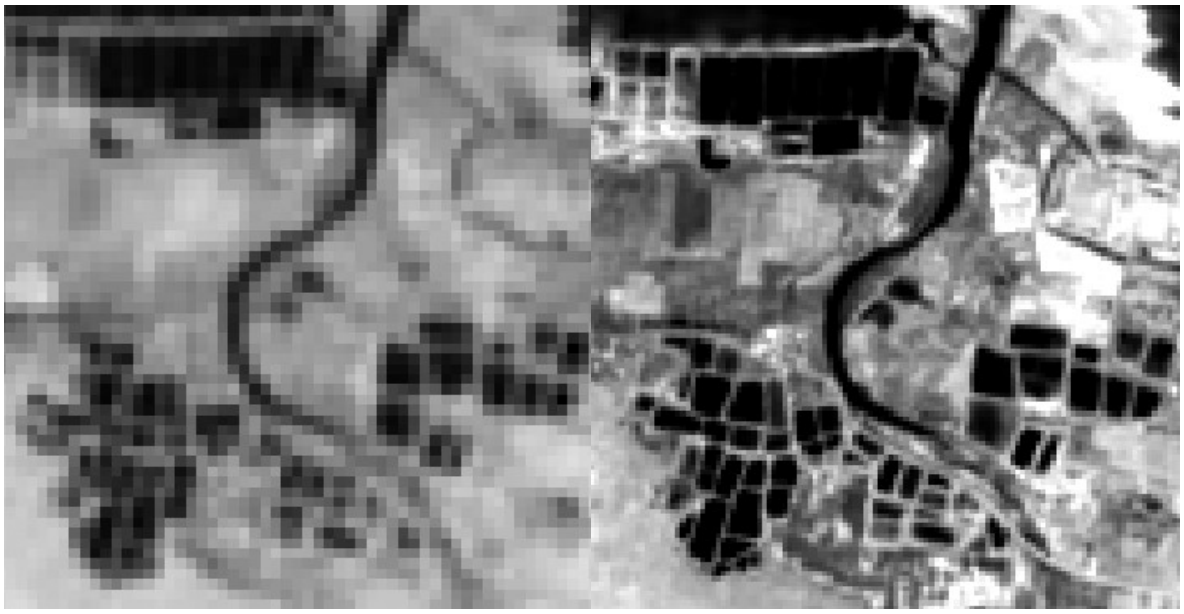
Wavelet Transform



PNN

SWIR bands:

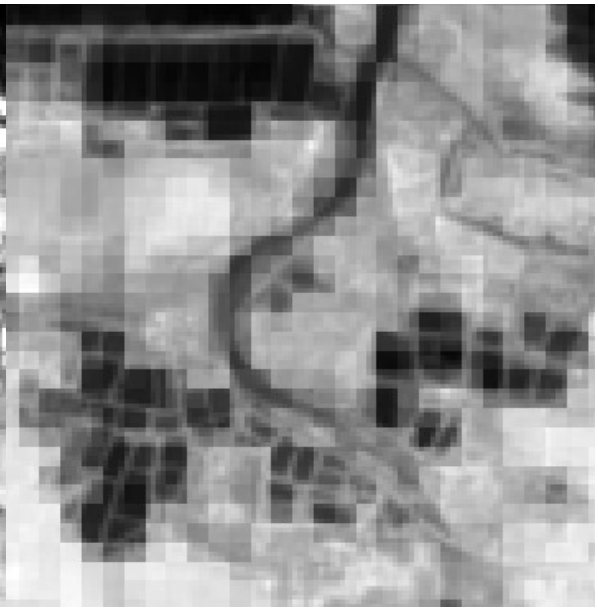
Band 11:



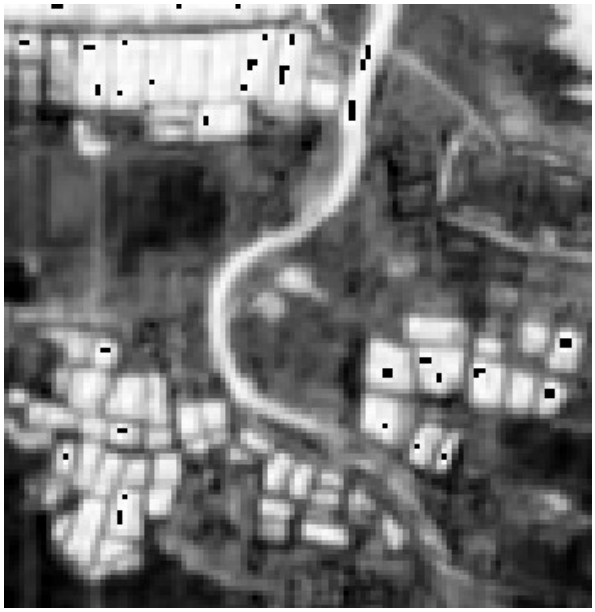
Reference image

Brovey Transform

IHS Transform



Wavelet Transform



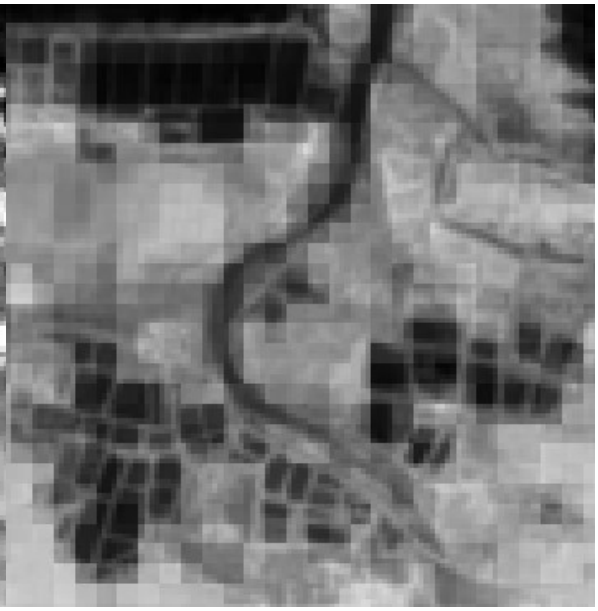
PNN

Band 12:

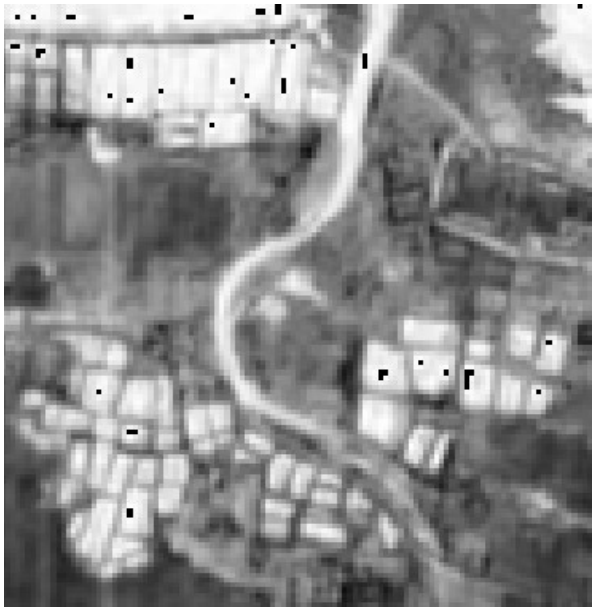
Reference image

Brovey Transform

IHS Transform



Wavelet Transform



PNN