Università degli Studi di Catania

Dipartimento di Ingegneria Elettrica Elettronica e Informatica

PhD in Systems, Energy, Computer and Telecommunications Engineering - XXXVIII cycle

---

Towards Quantifying Volcanic and Climate Interactions Through AI-Enhanced Modeling and Satellite Remote Sensing

PhD Thesis

Author
Simona Cariello

Supervisor
Prof. Maide Bucolo
Dr. Ciro Del Negro

Cosupervisor
Dr. Claudia Corradino

ACADEMIC YEAR 2024/2025

Simona Cariello

Towards Quantifying Volcanic and Climate Interactions Through AI-Enhanced Modeling and Satellite Remote Sensing

PhD Thesis

UNIVERSITY OF CATANIA

"*As Jim Morrison once said, "be like the sea, always finding the strength to rise again after breaking against the rocks." And so it was for me: before every obstacle and every fear, I found the courage to face them, to try once more, and now I can finally say.. I did it!*"

Simona Cariello

UNIVERSITÀ DEGLI STUDI DI CATANIA

Dipartimento di Ingegneria Elettrica Elettronica e Informatica

Automation Engineering and Control of Complex Systems

Towards Quantifying Volcanic and Climate Interactions Through AI-Enhanced Modeling and Satellite Remote Sensing

by Simona Cariello

ABSTRACT

Volcanic eruptions represent one of the most dynamic and complex natural phenomena on Earth, capable of profoundly affecting both local populations and global environmental systems. The importance of continuous and precise volcanic monitoring extends beyond the mitigation of immediate hazards, such as pyroclastic flows, lava emplacement, tephra dispersal, and volcanic gas emissions, to encompass the broader scientific imperative of understanding the intricate interactions between volcanic processes and the Earth's atmospheric and climatic systems. In this context, recent advancements in satellite technologies and Earth Observation (EO) systems have significantly improved our ability to monitor volcanic activity, even in remote and otherwise inaccessible regions. The integration of multi-sensor satellite data, characterized by different spatial and temporal resolutions, provides a more comprehensive understanding of volcanic phenomena. However, the ever-growing volume and complexity of these datasets necessitate the adoption of advanced computational tools and artificial intelligence (AI) techniques, which offer new perspectives for processing, analyzing, and interpreting large-scale geophysical data.

This thesis presents a set of computationally innovative AI methods for volcanic monitoring, developed by the author through the integration of observations from the visible, infrared, and ultraviolet spectral ranges, with the aim of building a prototype global monitoring platform. Several Machine Learning (ML) algorithms and feature engineering strategies have been designed to extract key information and patterns related to volcanic activity, primarily from thermal infrared (TIR) and visible imagery. Furthermore, advanced Deep Learning (DL) models have been implemented for scene classification and anomaly detection, enabling the identification of isolated (intra-crater) and extended (lava flow) thermal anomalies. The use of foundation models has also been explored to enhance the segmentation and quantification of lava flows, even for low-intensity eruptions, across multiple case studies.

The developed methodologies have been further extended to investigate volcano–climate interactions, aiming to bridge the traditional monitoring chain, comprising forecasting, detecting, tracking, quantifying, and nowcasting, with the assessment of volcanic impacts on the atmosphere and climate. This includes adapting existing models to the ultraviolet (UV) spectral range to analyze and quantify SO_2 dispersion following volcanic eruptions. Preliminary studies have also been conducted using reanalysis products such as ERA5 to explore the potential influence of recent eruptions on climate variability, particularly temperature changes, across both local and global scales and over varying temporal horizons.

In conclusion, this thesis presents newly developed AI-based algorithms for volcanic monitoring, emphasizing how technological innovation,both in methodology and instrumentation, constitutes a cornerstone in advancing our understanding of complex natural processes such as volcanic activity and its environmental implications.

Contents

List of Figures

List of Tables

To myself, who learned to keep walking even when the path was unclear...

Chapter 1

Introduction

1.1 Volcanic Phenomena and Climate interactions

Volcanic eruptions represent one of the most dynamic and complex natural phenomena on Earth, capable of profoundly impacting both local populations and global environmental systems. The importance of continuous and precise volcanic monitoring extends beyond the mitigation of immediate hazards, such as pyroclastic flows, lava flow emplacement, tephra dispersal, and volcanic gas emissions, to encompass the broader scientific imperative of understanding the complex interactions between volcanic processes, the Earth's atmospheric and climatic systems [1]. As recent events have demonstrated, even moderate eruptions can disrupt air traffic, affect agriculture, and influence weather patterns, while larger-scale eruptions may alter the Earth's radiative balance for years [2]. For instance, the 1991 eruption of Mount Pinatubo led to a global cooling of about 0.5°C due to the injection of sulfur dioxide into the stratosphere [3]. In addition to explosive eruptions, persistent low-intensity degassing from open-conduit volcanoes can also contribute cumulatively to the atmospheric budget of greenhouse and aerosol-forming gases, with potential long-term implications for regional and global climate systems [4], [5]. Real-time monitoring of volcanic activity significantly enhances our understanding of volcano dynamics, enabling timely detection of unrest, more accurate eruption forecasts, and the implementation of preventative measures to minimize risk and societal impact. However, active volcanoes pose unique challenges for direct and continuous observation, particularly during eruptive phases, when hazardous conditions often prevent field access and instrumental deployment. Paradoxically, these are also the moments when the volcano releases the most valuable information about magma dynamics and surface processes. In such contexts, satellite remote sensing becomes an indispensable complement to ground-based networks, enabling continuous, safe, and multi-spectral monitoring of volcanic activity even during high-energy eruptive events [6]. In this framework, the reinforcement of Earth Observation networks and the development of advanced data analysis

methodologies are key to effectively exploiting the increasing volume and diversity of available observations. The present thesis contributes to this goal by proposing an integrated AI-based framework designed to combine and analyze multi-sensor satellite data, thereby improving the detection, interpretation, and characterization of volcanic activity across different spatial and temporal scales. Modern technological advances offer unprecedented opportunities to monitor volcanoes with increasing resolution, continuity, and precision, not only enhancing our capacity to detect and forecast eruptive activity but also enabling more accurate quantification of gas emissions and aerosol outputs that influence atmospheric composition and climate dynamics. Satellite observations range from low spatial resolution but high temporal frequency sensors to high spatial resolution but low temporal frequency instruments, including wide-swath thermal sensors and instruments operating across various regions of the electromagnetic spectrum, provide essential datasets for detecting subtle signs of unrest, tracking plume evolution, mapping lava flow, and quantifying gas emissions [6]. Additionally, sensors dedicated to atmospheric monitoring play a key role in characterizing aerosol dispersion and gas transport in the troposphere and stratosphere, which are critical parameters for evaluating the climatic impact of volcanic eruptions [7]–[9]. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have increasingly been adopted for geophysical and remote-sensing applications, including volcanic monitoring, environmental hazard detection, and thermal anomaly recognition [10]–[12]. AI has already demonstrated strong potential for managing large, heterogeneous datasets and identifying non-linear patterns that traditional thresholding or semi-empirical methods may overlook. The integration of feature-engineering strategies, ensemble models, and data-fusion architectures is designed to bridge this gap, enabling a quantitative assessment of how and where AI provides genuine improvements in volcanic monitoring tasks. This includes the use of computer vision techniques for automated feature extraction from remote sensing images, as well as ML an DL algorithms for the identification of patterns and outliers in multivariate time series data [13]–[15]. These approaches enable the detection of potential precursors and the recognition of anomalous spatiotemporal dynamics that may precede eruptive activity, offering new avenues for early warning and system-level understanding of volcanic processes. In this evolving landscape, the advent of Foundation Models, large, pre-trained neural networks capable of adapting to a wide range of downstream tasks, represents a significant leap forward [16], [17]. These models, trained on diverse and massive datasets, offer enhanced generalization capabilities and can be fine-tuned for specific volcanic monitoring applications, improving the accuracy, adaptability, and scalability of the traditional AI-based systems. Looking ahead, emerging technologies such as Quantum Computing (QC) promise to significantly enhance our capacity to process and model large-scale volcanic and atmospheric data, opening new avenues for early warning systems and predictive modeling [18]–[20]. This thesis focuses on volcanic monitoring across all its key phases, forecasting, detecting, tracking, quantifying, and nowcasting, to construct a comprehensive picture of how a volcanic eruption unfolds over time. Through the integration of these stages, the study aims to deepen our understanding of how, when, and to what extent volcanic eruptions can impact the climate, linking geophysical processes to atmospheric responses in a coherent and multidisciplinary framework. Such a framework requires the integration of methodologies and

concepts from geophysics, atmospheric science, remote sensing, and computational modeling, reflecting the inherently interdisciplinary nature of Earth Observation (EO). The thesis is structured to explore these interconnected dimensions, beginning with a review of volcanic activity and its climate interactions, followed by an analysis of observational technologies and data analysis methods, and culminating in the proposal of a methodological framework for integrated volcanic monitoring and climate impact assessment, supported by selected case studies that illustrate the practical application of the proposed approaches. Ultimately, this work aims to contribute to the development of operational tools for volcanic hazard assessment, reinforcing the link between fundamental research and decision-making processes.

1.1.1 Volcanic Phenomena: Thermal and gas emission

A volcano is a geological structure through which a mixture of magma and gases rises to the surface from a reservoir beneath the Earth's crust. These eruptions may lead to the accumulation of volcanic material around the vent, often resulting in the formation of hills or mountains of varying morphology and size [21]–[23]. The surface opening is known as the crater, connected to deeper magma chambers by a central conduit. Beneath the surface, molten rock (magma) accumulates once it erupts, it is referred to as lava. Volcanoes represent one of the most striking expressions of Earth's internal dynamics. Magmatic eruptions result from complex geological processes involving the generation, migration, and accumulation of magma within the lithosphere. Among these processes, the exsolution and buildup of volcanic gases, such as water vapor, carbon dioxide, Sulphur dioxide and other volatiles, in order of abundance, play a fundamental role [24], [25]. The release and expansion of these gases within the magma chamber is a key mechanism that can trigger an eruption, particularly in explosive systems. The spatial distribution of volcanoes is closely linked to plate tectonics and is far from random. Most volcanoes are located along tectonic boundaries, especially at convergent margins where subduction occurs, but also along mid-ocean ridges and continental rifts. These geodynamic settings give rise to extensive volcanic chains and mountain ranges, both on land and under the sea (Figure 1.1). According to the United States Geological Survey (USGS), there are more than 1,500 potentially active volcanoes worldwide, over 500 of which have erupted in historical times. While a majority are concentrated along the Pacific Ring of Fire, notable clusters are also found in other tectonically active zones, including the Mediterranean basin. In these areas, the interaction between subduction processes and crustal dynamics produces high volcanic risk, especially due to the proximity of densely populated regions. The geodynamic setting not only controls the spatial distribution of volcanoes but also its chemical composition strongly influencing the style and behavior of volcanic activity. Depending on the tectonic environment and magma characteristics, volcanoes can exhibit a wide range of eruptive dynamics, which are generally categorized into effusive and explosive types.

Thermal and gas emissions represent two of the most direct and physically meaningful expressions of volcanic activity. They originate from complex endogenous processes involving magma dynamics, volatile degassing, magma-crust interaction, and the coupling between magmatic and hydrothermal systems. Understanding the origin, evolution, and spatio-temporal distribution

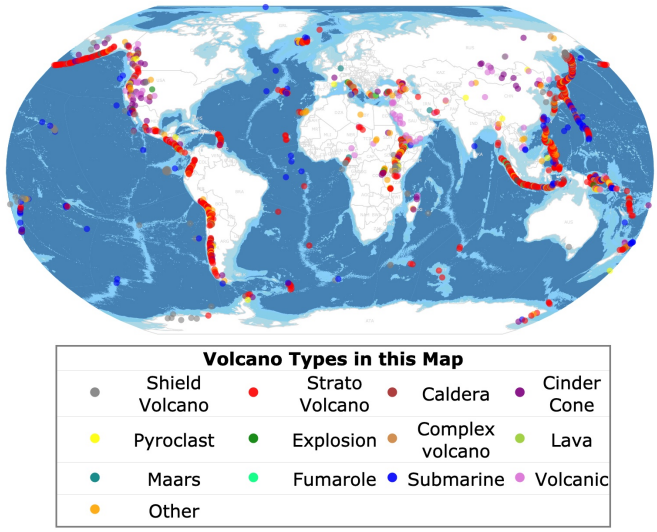

FIGURE 1.1: Map of the world showing the locations of volcanoes, with different colors indicating the various types of volcanoes. Image courtesy of https://databayou.com/.

of these signals is essential for interpreting the state of the volcano, identifying potential precursory signals, and improving eruption forecasting models.

In the volcanic context, the term thermal anomaly/emission refers to a local and persistent increase in surface temperature that significantly deviates from the typical background conditions of the surrounding environment (Figure 1.2). These anomalies are generally interpreted as evidence of abnormal heat flow from the subsurface, due to the presence or movement of magma, hot gases, or hydrothermal fluids. Thermal anomalies are dynamic and multilevel processes triggered when a significant amount of heat is transferred from the magma chamber or intermediate magmatic bodies to the surface. Heat transfer can occur through: Conduction: through rock, in the presence of a strong thermal gradient; Convection: carried by hot fluids, such as gases or vapors migrating toward the surface; Direct radiation: in cases where molten material is exposed to the atmosphere (e.g., lava flows or incandescent domes). These mechanisms lead to increased surface temperatures in specific areas, such as active craters, volcano flanks, fumarolic zones, eruptive vents, or fractures. One of the most interesting features of thermal anomalies is their ability to appear well before visible volcanic activity begins. In many documented cases, variations in thermal flux have been observed weeks or even months before an eruption, making them potential precursory signals [26]–[28]. The ascent of magma from deeper levels increases pressure and temperature in the intermediate parts of the volcanic system. In some cases, magma may rise to very shallow levels without erupting, producing a strong localized thermal anomaly, especially when accompanied by high-pressure hot gases. In this sense, a thermal anomaly can be one of the first physical indicators of the onset of a new eruptive cycle. However, not all anomalies lead to an eruption: their duration, intensity, and location must be interpreted within the geodynamic context of the volcano and in synergy with other geophysical parameters (e.g., seismicity, ground deformation, gas amount and composition variations). Time-series analysis of surface temperature, combined with

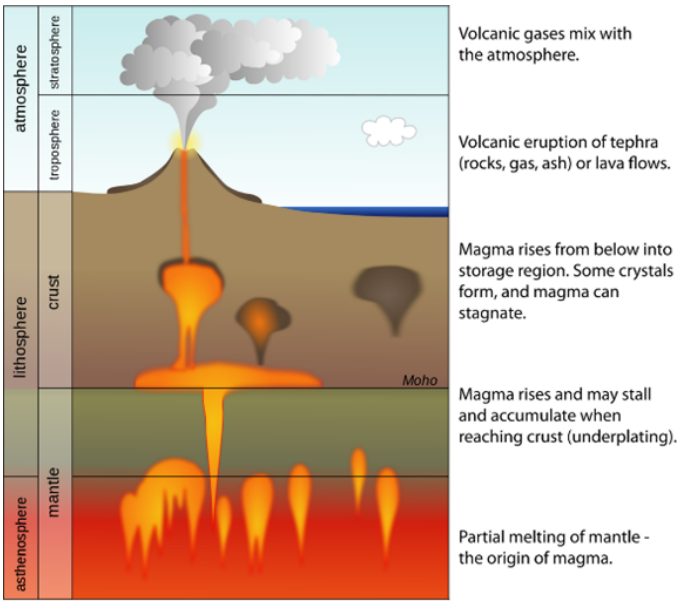

FIGURE 1.2: Schematic representation of the main volcanic processes: partial melting of the mantle generates magma that ascends through the lithosphere, may accumulate at the base of the crust (underplating) or in shallow reservoirs, and eventually reaches the surface through effusive or explosive eruptions; volcanic gases released are subsequently dispersed into the troposphere and stratosphere. Image courtesy of https://it.wikibooks.org/.

morphological and geochemical evolution of the hot area, allows a detailed reconstruction of volcanic system behavior and helps assess associated risks. Not all thermal anomalies are easily detectable. Their observability depends on various factors, including: The depth of the magma and its ability to transfer heat to the surface; The presence of vegetation, ash, or snow cover, which can obscure thermal signals; The geometry of the hydrothermal system, which may dissipate or redirect heat; Atmospheric and environmental conditions, which influence the infrared radiation detected by remote sensing. Despite these limitations, when correctly interpreted, thermal anomalies provide critical information on volcanic system evolution. Their appearance may precede an eruption, but it is often their temporal variation that contains the most informative signal. Thermal anomalies are typically detected using remote sensing systems, including satellite-based infrared sensors (e.g., MODIS, VIIRS, Sentinel-2), as well as ground-based thermal cameras.

Alongside thermal anomalies, gas emissions represent one of the main surface manifestations of the internal dynamics of a volcanic system. Both originate from deep processes related to the presence and movement of magma, and their combined analysis allows for the integration of thermal and chemical-physical signals, providing a more detailed understanding of volcanic evolution. While thermal anomalies reflect the presence of abnormal heat flows rising from underground, linked to magma, hot gases, or hydrothermal fluids, gas emissions offer a true "direct window" into the magmatic reservoir. Volcanic gases are initially dissolved in magma as volatiles (H_2O, CO_2, SO_2, etc.) and are released progressively as pressure decreases during magma ascent. Changes in chemical composition, gas ratios, and emission fluxes can serve as significant precursory signals, often observable weeks or even months before

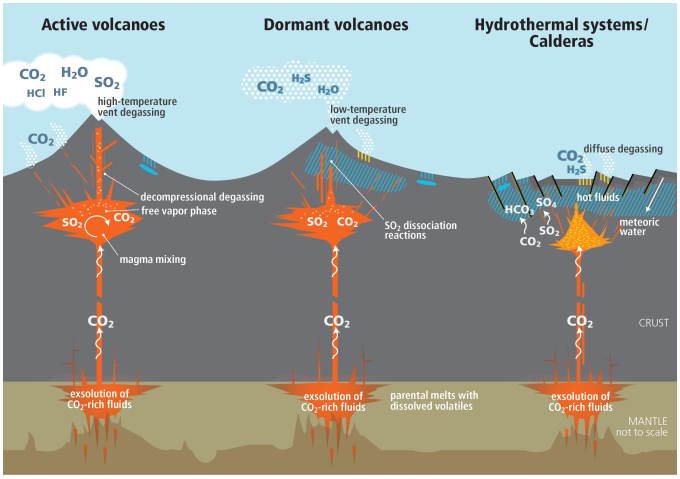

FIGURE 1.3: Schematic of the main volcanic and hydrothermal degassing processes. Image courtesy of "DCO Decadal Report, 2019".

an eruption. For instance, an increase in SO_2 flux, a gas with moderate solubility in magma and typically associated with ascending fresh magma, may indicate the opening of new degassing conduits, system depressurization, or the arrival of a new magmatic batch. These variations are often accompanied by other geophysical signals, such as seismic activity or ground deformation, which strengthen their interpretive value. It is important to note that gas emission can also occur in the absence of eruptive activity, during quiescent phases or periods of passive degassing. Though less visually striking, these emissions provide valuable insights into the state of the volcanic system and its potential evolution. In particular, the persistence or variation of gas fluxes may indicate magmatic recharge processes, decompression, or the progressive exsolution of volatiles. Gas emissions do not only affect the local volcanic system, but they can also have climatic effects on regional or even global scales. Explosive eruptions, for example, can inject large amounts of SO_2 into the stratosphere, where it converts into sulfate aerosols that reflect incoming solar radiation, leading to short-term global cooling. Water vapor and CO_2 also play a role, contributing to the radiative balance and increasing stratospheric humidity. In addition, submarine eruptions can release gases into the atmosphere or significantly alter marine chemistry. Gas emissions are monitored through spectroscopic techniques, such as DOAS (Differential Optical Absorption Spectroscopy), FTIR (Fourier-Transform Infrared Spectroscopy), and multi-GAS sensors, deployed both on the ground and via drones or aircraft [29], [30]. Anyway, both thermal and gas emission data are increasingly used to inform and constrain numerical models of volcanic behavior [31], [32]. These inputs contribute to hazard assessment tools and eruption forecasting frameworks by providing real-time constraints on magma dynamics. Taken together, thermal anomalies and gas emissions offer complementary insights into the dynamics of a volcanic system. Their early detection and interpretation, especially when combined with seismic and deformation data, are key to improving early warning systems and reducing volcanic risk for nearby populations.

1.1.2 Volcanic Activity: Passive degassing, explosive and effusive activity

Passive Degassing

Passive degassing refers to the continuous or intermittent release of volcanic gases from a volcano in the absence of significant eruptive activity. This process can occur through open-conduit systems, where magma resides at shallow depth and volatiles escape freely through the vent, as well as through diffuse emissions across the flanks, fumarolic fields, and hydrothermal systems. The principal gases released include water vapor (H_2O), carbon dioxide (CO_2), and sulfur dioxide (SO_2), along with minor species such as hydrogen sulfide (H_2S), halogen compounds, and trace metals in aerosol form. Passive degassing plays a critical role in volcanic monitoring because variations in gas fluxes or compositions can indicate changes in the underlying magmatic system, such as magma recharge, depressurization, or conduit sealing. From a climatic perspective, persistent emissions of CO_2 contribute to the long-term volcanic input to the global carbon cycle, while SO_2 emissions can affect atmospheric chemistry and radiative balance by forming sulfate aerosols that scatter solar radiation and potentially induce short-term cooling. Although the climatic effect of a single passively degassing volcano is generally small compared to major explosive eruptions, the cumulative contribution from persistently active systems, such as Masaya (Nicaragua), Etna (Italy), and Kīlauea (Hawai‘i), is significant on regional to local scales. Furthermore, sustained SO_2 output can degrade air quality locally and regionally, producing vog (volcanic smog) that impacts human health, agriculture, and ecosystems.

Explosive Activity

The intensity and magnitude of explosive eruptions are commonly evaluated using the Volcanic Explosivity Index (VEI), discussed in more detail in paragraph 1.2.2, a logarithmic scale ranging from 0 (non-explosive) to 8 (mega-colossal). The VEI considers parameters such as eruption column height, the volume of ejected tephra, and qualitative descriptors of explosiveness. Eruptions with $VEI \geq 4$ are typically capable of producing regional to global atmospheric disturbances, while events such as Pinatubo (1991, VEI 6) and Tambora (1815, VEI 7) have been linked to significant short-long term cooling. Although both eruptive styles are crucial for hazard assessment, high VEI explosive eruptions are especially relevant to climate studies due to their capacity to temporarily alter the Earth's radiative balance. Explosive eruptions are driven by high-viscosity magma and rapid gas exsolution, resulting in violent fragmentation and the release of gases, and pyroclastic materials. These phenomena arise when ascending magma contains a high concentration of volatiles that exsolve rapidly as pressure decreases near the surface. The release of gas propels molten fragments tens to hundreds of meters into the air, forming sustained or pulsating jets of incandescent lava. Lava fountains can occur in continuous episodes lasting hours to days, or as short-lived bursts during paroxysmal phases, and they are frequently associated with the emplacing of extensive short-lived lava flows as well as the generation of tephra deposits around the vent. Although lava fountains typically produce lower eruption columns than high VEI Plinian events, they can still inject significant amounts of volcanic gases, particularly SO_2, into the lower

atmosphere, contributing to local to regional air quality degradation and, under specific atmospheric conditions, potentially influencing climate-relevant aerosol formation. The high thermal output of sustained lava fountaining also drives strong convective plumes that may entrain fine ash and gas into higher atmospheric layers, increasing their dispersion range. Recent observations provide clear examples of the atmospheric implications of lava fountain. During the 2021 paroxysmal activity of Mount Etna (Italy), repeated lava fountain episodes, some exceeding 1 km in height, released substantial SO_2 fluxes measurable by TROPOMI, alongside ash plumes transported over hundreds of kilometers. These events, while moderate in VEI (typically 2–3), demonstrated how high-frequency fountaining can produce cumulative gas emissions comparable to those of a single, larger explosive event. Similarly, the 2018 lower East Rift Zone eruption of Kīlauea (Hawai'i) featured sustained lava fountains feeding extensive lava flows, with significant SO_2 output affecting air quality across the Hawaiian Islands and producing regional-scale vog (volcanic smog) that impacted human health and ecosystems. In a comparable degassing-dominated setting, the persistent activity at Masaya volcano (Nicaragua) has long been characterized by continuous, open-vent SO_2 emissions, with fluxes among the highest recorded globally for a basaltic volcano, contributing to chronic regional air quality issues and providing a natural laboratory for studying the impacts of sustained volcanic gas release on climate, environment, and public health.

As discussed before, such eruptions can generate eruption columns that inject significant amounts of volcanic gas (e.g. sulfur dioxide SO_2) and volcanic aerosols into the stratosphere, with potential climatic consequences. One of the most climatically relevant processes during an explosive volcanic eruption is the conversion of SO_2 into sulfate aerosols. Once injected into the stratosphere, SO_2 reacts with water vapor and ozone, forming sulfuric acid (H_2SO_4) in the form of tiny, suspended droplets. These aerosols reflect a portion of incoming solar radiation, reducing the amount of energy that reaches the Earth's surface and leading to a temporary cooling of the climate, which can last from several months to 2–3 years, depending on the persistence and distribution of the aerosols in the stratosphere [33], [34]. In the case of submarine volcanoes, interactions between magma and seawater can exert a significant influence on both eruptive dynamics and climate impacts. When magma meets water, it can trigger violent phreatomagmatic explosions, increasing the fragmentation of magma and the dispersal of fine ash and gases into the ocean and potentially into the atmosphere [35]. Moreover, submarine eruptions can release substantial quantities of water vapor into the atmosphere if the eruption breaches the ocean surface or produces steam-rich plumes. Water vapor, although less potent as a long-lived greenhouse gas compared to carbon dioxide, is a critical climate agent due to its strong shortwave and longwave radiative effects. In specific scenarios, especially for shallow submarine eruptions or those occurring near the surface, the additional injection of water vapor can locally enhance atmospheric humidity, influence cloud formation processes, and potentially amplify or modulate the climatic signals associated with volcanic sulfur emissions. A remarkable example is the eruption of the Hunga Tonga, Hunga Ha'apai volcano in January 2022. This powerful submarine eruption generated an exceptionally tall eruption plume that reached the mesosphere, injecting an estimated 146 ± 5 teragrams of water vapor into the stratosphere [36]. The extraordinary water

vapor injection may contribute to stratospheric heating and influence ozone chemistry, with possible implications for regional and global climate patterns over subsequent years. The specific mechanisms through which these effects occur are discussed in greater detail in paragraph 1.2.

Effusive Activity

Effusive eruptions involve the emission of low-viscosity magma, typically basaltic in composition, which allows gases to escape relatively easily from the melt without generating significant explosive pressure buildup. Effusive activity can manifest in various morphological features: Lava flows, which may travel several kilometers from vent, shaping broad basaltic fields or channelized flows depending on the local topography and lava rheology; Lava tubes, which form as the outer crust of a flow cools and solidifies, allowing molten lava to continue moving through insulated channel beneath the surface, enabling flows to extend considerable distances; Lava lakes, sustained by continuous magma supply and overturn, often found in summit craters or calderas (e.g., Kīlauea in Hawaii); Lava domes, more typical of higher viscosity magmas, although basaltic domes also occur, grow slowly through effusion. The effusive eruptions, although often persistent and capable of producing significant volumes of lava over time, generally have limited direct impact on the global atmosphere. Effusive eruptions typically involve partially degassed magma releasing modest amounts of volcanic gases, primarily H_2O with lower SO_2 emissions compared to explosive events. Nevertheless, effusive eruptions, when accompanied by explosive activity can have considerable local and regional environmental impacts. Notable examples include the Holuhraun eruption in Iceland (2014–2015), which produced one of the largest lava flows in Europe in modern times, emitting substantial SO_2 levels that significantly degraded air quality in Iceland and parts of Scandinavia [37], [38]. Similarly, prolonged volcanic activity at Hawaii's Kīlauea and Mauna Loa volcanoes has caused the persistent formation of volcanic smog (vog), a haze of SO_2 and sulfate aerosols, resulting in respiratory health concerns, crop damage, and visibility reduction [39]. Effusive eruptions also contribute to incremental topographic changes, creating new land surfaces, modifying drainage systems, and altering ecosystems [40]. The radiative properties of fresh lava surfaces, which are initially dark and highly heat-absorbing, evolve over time as weather increases their reflectance (albedo), further influencing local energy balances. Although effusive eruptions do not produce significant effects on the global radiative balance, their monitoring remains essential for hazard assessment, urban planning in volcanic regions, and evaluating localized environmental and health impacts. Thermal remote sensing, combined with gas measurements, is crucial for quantifying effusion rates and tracking changes in eruptive behavior. In particular, the estimation of parameters such as the Volcanic Radiative Power (VRP) and the Time-Averaged Discharge Rate (TADR) provides valuable insights into the volume of lava emitted and the dynamics of effusive activity.

1.1.3 Historical Context and Societal Relevance

Volcanic activity has played a significant role in shaping human history, not only by altering physical landscapes but also by influencing the development, collapse, and adaptation of societies. Major eruptions have left indelible

marks on cultural memory and historical records. One of the earliest and most iconic examples is the eruption of Mount Vesuvius in 79 AD, which buried the Roman cities of Pompeii and Herculaneum under a thick layer of pumice and ash, preserving a vivid snapshot of ancient life but also highlighting the devastating potential of sudden volcanic events [41], [42]. Other catastrophic eruptions, such as the 1815 eruption of Mount Tambora, led to widespread agricultural collapse and famine during the so-called "Year Without a Summer", while the 1883 eruption of Krakatoa generated massive tsunamis and produced atmospheric effects observable across the globe [16]. Similarly, the 1783–1784 eruption of the Laki fissure in Iceland released vast quantities of sulphur dioxide and other volcanic gases, creating a persistent haze across much of Europe. This event caused significant climatic anomalies, crop failures, and famine, effects that some historians link to the social unrest and economic hardship contributing to the French Revolution of 1789 [43]. While these historical eruptions demonstrate the dramatic potential of volcanic events, it is important to recognize that volcanism remains a daily reality for millions of people worldwide who live close to active volcanoes. Living in volcanic regions entails both opportunities and risks. From one hand, volcanic soils are often highly fertile, supporting intensive agriculture; geothermal energy offers a renewable energy source; and tourism linked to volcanic landscapes can be economically beneficial. From the other hand, the proximity to active or potentially active volcanoes exposes populations to a wide range of hazards. These include not only the obvious threats of lava flows, pyroclastic density currents, ash fall, and gas emissions, but also secondary effects such as lahars, landslides, and even tsunamis in coastal or lake-adjacent volcanic systems. Some example include Kīlauea (Hawai'i), where effusive lava flows have repeatedly inundated communities; Mount Merapi (Indonesia), renowned for its deadly pyroclastic density currents; Sakurajima (Japan), which produces frequent ash falls affecting nearby urban areas; Masaya (Nicaragua), a persistent source of volcanic gas emissions impacting air quality; Anak Krakatau (Indonesia), whose flank collapse in 2018 generated a deadly tsunami; and Nyos (Cameroon), where a limnic eruption released a lethal cloud of CO_2 from a volcanic lake. Mount Etna stands out as a paradigmatic example of the complex interplay between volcanic hazards, human settlement, and scientific observation. Located in eastern Sicily, Etna exhibits frequent effusive and explosive eruptions, often close to populated areas. The volcano poses risks to infrastructure, including highways, airports, and urban settlements, but also provides economic opportunities through agriculture, viticulture, and tourism. Its continuous activity has made it a key site for volcanological research, geophysical monitoring, and the testing of advanced technologies, such as satellite remote sensing and real-time alert systems. While terrestrial volcanoes are often the most visible and studied, submarine volcanic systems also play a crucial role in shaping the Earth system and can produce significant, though less predictable, hazards. Volcanic eruptions can also originate in submarine environments, where they may go unnoticed until their effects manifest at the surface or across large distances. Although less visible, submarine eruptions can produce significant geophysical and climatic consequences. As discussed before, the 2022 eruption of the Hunga Tonga–Hunga Ha'apai volcano in the South Pacific, for instance, generated powerful shockwaves, tsunami waves that reached distant coastlines, and injected unprecedented amounts of water vapor into the stratosphere, raising concerns about potential short

- long term atmospheric effects. These events underscore the fact that even remote or submarine volcanic systems can have far-reaching impacts on both regional stability and the global environment. Whether eruptions occur on land or beneath the ocean, the effectiveness of response efforts depends heavily on how volcanic risk is communicated and managed at multiple societal levels. In this complex scenario, effective communication of volcanic risk plays a central role in minimizing loss of life and ensuring timely evacuations. Scientific monitoring and forecasting, while crucial, must be supported by transparent, accessible, and culturally sensitive communication strategies that reach both authorities and at-risk communities. Past events have shown that communication gaps can lead to confusion, delayed response, and in some cases, public mistrust in scientific institutions. Strengthening the dialogue between scientists, decision-makers, and the public is therefore essential for building societal resilience and enhancing preparedness in volcanic regions. Beyond emergency response, volcanic activity can have broader societal implications that extend well beyond the immediate vicinity of the eruption. The societal impacts of eruptions are not limited to direct human casualties or damage to infrastructure. Ash dispersal can compromise aviation safety and ground flights over large regions, as demonstrated by the Eyjafjallajökull eruption in Iceland in 2010 [44]. Volcanic gases can affect air quality and public health, while disruptions to water supply, agriculture, and transportation networks can have long-term socioeconomic consequences. Over the centuries, human understanding of volcanism has evolved from mythological explanations to a scientific approach rooted in geophysics, geochemistry, and observational technologies. Today, volcano monitoring has become an integral component of disaster risk reduction strategies. The establishment of institutions such as the United States Geological Survey (USGS), the Istituto Nazionale di Geofisica e Vulcanologia (INGV), and global initiatives like the Global Volcano Model (GVM) reflect a growing recognition of the need for coordinated efforts in hazard assessment and early warning. In the modern context, where local volcanic events can have global repercussions, both socioeconomically and climatically, advancing monitoring techniques and predictive capabilities are more crucial than ever. Understanding the historical and present-day relevance of volcanic activity reinforces the societal imperative to integrate scientific knowledge, technological innovation, and community preparedness to mitigate risk and enhance resilience in volcanic regions.

1.2 Volcanic Forcing of Climate: mechanisms and evidence

Volcanic eruptions are among the most significant natural forcings of Earth's climate system, causing impacts that range from short-term cooling episodes to longer-term influences on climate variability. These effects are primarily driven by the injection of gases and particles into the atmosphere. Sulfur dioxide (SO_2) plays a crucial role, as it converts in the stratosphere to sulfuric acid aerosols (H_2SO_4), which reflect solar radiation, increase Earth's albedo, and cause surface cooling. The magnitude and duration of this cooling depend on factors such as the volume of emissions, plume altitude, and the eruption's latitude and season. In addition to direct radiative effects, volcanic aerosols can

alter cloud properties and influence atmospheric circulation patterns, including shifts in the North Atlantic Oscillation (NAO) and changes in monsoon systems, potentially amplifying the cooling signal. Volcanic impacts are well documented in ice cores, tree rings, and sediment records. Historic eruptions such as Laki (1783), Tambora (1815), Krakatoa (1883), and Pinatubo (1991) caused notable global cooling and have been instrumental in calibrating climate models and improving our understanding of climate sensitivity [45], [46]. Modern satellite data since the 1980s have enhanced our ability to monitor volcanic emissions, revealing changes in aerosol levels and reductions in solar radiation [47], [48]. Although volcanic effects are typically short-lived compared to greenhouse gases like CO_2, volcanic forcing remains crucial for understanding natural climate forcing and for evaluating climate models. The climatic impact of volcanic eruptions depends on several factors, whose relative importance may vary depending on the context. These factors, ranging from the magnitude and composition of the erupted material, to the altitude of injection, geographic location, and background atmospheric conditions, will be examined in detail throughout this thesis. Each will be discussed under different scenarios in order to highlight their specific role and relative weight in shaping the short- and long-term climatic consequences of volcanic activity. Assessing whether an eruption can significantly affect the climate requires considering multiple elements. Not all eruptions have the same capacity to influence global climate. The effectiveness of a volcanic eruption as a climate forcing depends on interconnected factors, including the volcano's geographical location, eruption type, Volcanic Explosivity Index (VEI), and the season of occurrence.

1.2.1 Geographical location and seasonality

The geographical location of a volcano is one of the primary factors determining the ability of a volcanic eruption to influence Earth's climate. Atmospheric dynamics and the structure of global circulation dictate how volcanic aerosols, once injected into the stratosphere, can disperse on hemispheric or global scales. Eruptions occurring in the intertropical region (approximately between 23°N and 23°S) have the highest potential for global climatic impact. This is due to the Brewer-Dobson circulation, a key component of stratospheric dynamics, which transports air masses, and thus sulfuric aerosols, upward and poleward [49], [50]. As a result, aerosols emitted by tropical volcanoes can gradually spread into both the Northern and Southern Hemispheres, achieving near-global coverage. A classic example is the 1991 eruption of Mount Pinatubo in the Philippines, located around 15°N. The sulfuric aerosols injected into the upper stratosphere spread across both hemispheres, causing a global mean surface temperature decrease of approximately 0.4–0.6°C over the following two years. Mid-latitudes eruptions (approximately between 23°N and 55°S) can have significant climatic effects, but these tend to be more confined to the hemisphere in which the eruption occurs. This is because the zonal flow in the stratosphere acts as a barrier, limiting the exchange of stratospheric air masses between the two hemispheres, particularly during winter months, when the polar circulation is stronger and more isolated. For example, the eruption of Laki in Iceland (1783–1784), although significant in terms of gas emissions, produced climatic impacts primarily in the Northern Hemisphere, leading to thermal anomalies and socio-economic

consequences in Europe and North America, without significant effects in the Southern Hemisphere. High latitudes eruptions (above 55°) are less likely to exert a significant global climatic impact for several reasons. The tropopause is lower (around 8–9 km), meaning that even a relatively modest eruption can reach the stratosphere. However, the volume of air available for stratospheric transport is more limited. Additionally, the polar stratospheric circulation, especially during winter, tends to confine aerosols within the polar vortex, reducing the likelihood of spreading toward lower latitudes. Moreover, polar regions receive lower seasonal solar insolation, temporarily limiting the photochemical processes necessary for converting SO_2 into sulfuric aerosols. Nevertheless, a significant high-latitude eruption could still produce strong regional impacts, such as shifts in storm tracks or alterations in local atmospheric circulation, but it is unlikely to cause sustained global cooling.

In addition to geographical location, seasonality plays a significant role. During summer months, the tropopause is generally higher, particularly in the tropics, requiring greater eruptive energy for plumes to penetrate the stratosphere. Conversely, eruptions occurring in winter may more easily reach stratospheric altitudes due to a lower tropopause height. Furthermore, seasonal variations in atmospheric circulation influence how efficiently aerosols are transported and dispersed, ultimately affecting the spatial extent and duration of volcanic climatic impacts. In summary, tropical eruptions have the highest potential to produce significant and long-lasting global climate forcing, whereas mid- and high-latitude eruptions tend to generate more localized or hemispheric effects. The intensity of these effects depends on the complex interplay between atmospheric dynamics, seasonality, and the specific characteristics of the eruption itself.

1.2.2 Volcanic Explosivity Index (VEI)

The Volcanic Explosivity Index (VEI) is one of the indices currently used, as it was also in the past, to provide a quantitative measure of eruptive intensity. It is a logarithmic scale ranging from 0 to 8 and considers several parameters, including the volume of erupted material, the height of the eruption column, and the duration of the event. The type of eruption, explosive or effusive, is crucial in determining whether the volcanic plume will reach the stratosphere. Over the years, several indices have been developed as shown in Figure 1.4 and employed to assess the climate changes induced by volcanic eruptions. Among these are the Dust Veil Index (DVI), the index proposed by Mitchell, the VEI, the Sato index, and the Ice Core Volcanic Index (IVI). Each of these indices aims to measure different aspects of volcanic impacts, such as the quantity of aerosols released into the atmosphere, the persistence of stratospheric aerosol layers, or evidence found in paleoclimate records. However, VEI is currently the most widely used and serves as the primary reference for classifying the intensity of volcanic eruptions, thanks to its simplicity and the availability of historical and observational data on a global scale.

1.3 Earth Observation (EO) Technologies

Satellite remote sensing is today one of the most powerful and versatile tools for observing and studying the Earth [51]–[53]. In the fields of volcanology and atmospheric sciences, the use of satellite data has radically transformed the

Name	*Units*	*How Calculated*	*Reference*
Dust veil index (DVI)	Krakatau = 1000	*Sapper* [1917, 1927], sunsets, eruption, and radiation observations	*Lamb* [1970, 1977, 1983]
Mitchell	aerosol mass	based on H. H. Lamb (personal communication, 1970)	*Mitchell* [1970]
Volcanic explosivity index (VEI)	Krakatau = 6	explosivity, from geologic and historical reports	*Newhall and Self* [1982] *Simkin et al.* [1981] *Simkin and Siebert* [1994]
Sato	τ (λ = 0.55 μm)	*Mitchell* [1970], radiation and satellite observations	*Sato et al.* [1993]
Ice core volcanic index (IVI)	τ (λ = 0.55 μm)	average of ice core acidity or sulfate measurements	*Robock and Free* [1995, 1996]

FIGURE 1.4: Picture from [1]. Summary of commonly used volcanic indices employed to quantify the climatic impact of eruptions. The table reports the name of each index, its units of reference, the method of calculation, and the primary literature sources.

scientific community's ability to monitor, analyze, and understand complex phenomena such as volcanic eruptions, the dispersion of gases and aerosols, and their interactions with the climate system [54]–[57]. In the past, volcanic and atmospheric monitoring was severely limited by restricted spatial coverage, the temporal discontinuity of in situ observations, and the limited geographic accessibility of active sites. Remote or inaccessible volcanoes, such as those along mid-ocean ridges, in polar regions, or on remote oceanic islands, were often excluded from systematic observation programs. Today, thanks to satellites, it is possible to obtain global, synoptic, multispectral, and temporally continuous observations, with increasingly high spatial resolutions and reduced revisit times. The current generation of satellite missions, which includes both passive and active sensors (Appendix A.2), provides a broad spectrum of data products that are valuable for both operational volcanology and fundamental scientific research. A key parameter that determines the suitability of a satellite sensor for volcanic and atmospheric monitoring is its spatial resolution. The balance between spatial, spectral, and temporal resolution therefore defines the level of detail attainable in each application and determines the complementary roles of the different missions considered in this thesis. These satellites operate in two main orbital configurations polar orbits, and geosynchronous orbits. Polar-orbiting platforms, such as Sentinel, Terra, Aqua, Landsat, and Suomi-NPP, provide high spatial resolution and global coverage, but with a relatively low revisit frequency (ranging from one to several days). These are particularly suitable for detailed mapping and quantitative analysis of volcanic phenomena. Conversely, geostationary satellites, such as SEVIRI on Meteosat, ABI on GOES-R, and AHI on Himawari-8/9, offer continuous observations over a fixed area with high temporal resolution (up to every 10 minutes), though at the expense of spatial resolution, which is generally on the order of kilometers. These instruments are particularly effective for promptly detecting sudden events, such as Strombolian explosions, dome collapses, or the onset of lava flows, and for tracking the evolution of such phenomena in near real time. The integration of observations from both polar and geostationary orbits, combined with the use of both active and passive sensors, currently provides a synergistic and multi-layered satellite observation system capable of delivering reliable data for volcanic surveillance, event modeling, and risk management. These observational capabilities have not only improved the effectiveness of early warning systems and risk assessment

but also provided new tools to test and validate physical and numerical models at both local and global scales. Satellite remote sensing has become essential for studying the role of volcanic eruptions in the Earth's climate system, a topic of growing interest in scientific literature. At the same time, the massive use of satellite data presents methodological and computational challenges, related to the management of large data volumes, the integrated interpretation of multi-source measurements, and the refinement of spectral inversion algorithms, atmospheric correction, and georeferencing. Within the framework of this thesis, the discussion of satellite observation technologies is not intended as a general overview, but as the foundation for the methodological chain developed throughout the work. Each of the subsequent stages, ranging from forecasting to nowcasting, relies on specific types of satellite data, selected according to their spatial, temporal, and spectral resolution. This connection between sensor characteristics and model design represents a central methodological theme of the thesis. The following sections will present an overview of the main sensors used for volcanic monitoring (Paragraph 1.3.1) and for atmospheric studies (Paragraph 1.3.2), followed by a description of the latest technological developments and next-generation satellite missions (Paragraph 1.3.3), with a special focus on prospects for both research and operational applications.

1.3.1 Advances in Volcanic Monitoring

Satellite remote sensing has assumed a central role in the monitoring of active volcanoes over recent decades. Thanks to space-based observation, it is now possible to systematically detect and analyze multiple signals associated with volcanic activity, even in geographically remote, hard-to-reach areas, or those affected by prohibitive atmospheric conditions [58]. Satellite monitoring allows for the continuous observation of thermal anomalies linked to the presence of magma on the surface, the mapping of morphological transformations of the volcanic edifice, and the measurement of ground deformations associated with the pressurization or drainage of magmatic reservoirs [59]–[62]. These capabilities are made possible by a wide range of currently operating sensors, each designed to observe a specific portion of the electromagnetic spectrum or to detect topographic and dynamic changes on the Earth's surface. One of the most established applications of remote sensing in volcanology concerns monitoring thermal emissions as reported in Paragraph 1.1.1. Sensors operating in the thermal infrared (TIR), mid-infrared (MIR), and shortwave infrared (SWIR) regions enable the continuous and global detection and characterization of volcanic thermal signals [6]. A variety of instruments, including ASTER (aboard Terra), MODIS (Terra and Aqua), VIIRS (Suomi-NPP and NOAA-20), Landsat OLI/TIRS, Sentinel-2 MSI, and ECOSTRESS (aboard the ISS), provide data at different spatial and temporal resolutions. These datasets can be tailored to diverse monitoring objectives, ranging from real-time surveillance of ongoing eruptions to post-eruptive deposit mapping, as well as to the reconstruction of a volcano's thermal evolution over extended periods [62], [63]. The goal of monitoring thermal emissions is not limited to identifying hot sources but also quantitatively characterizing eruptive activity [57], [64], [65]. Among the available approaches, one of the most widely adopted is the estimation of the Volcanic Radiative Power (VRP), which represents the thermal power emitted by active

lava surfaces (Appendix A.2). VRP is derived from radiance measurements in the MIR and TIR spectral channels and can be further exploited to calculate the Time-Averaged Discharge Rate (TADR), i.e., the mean lava effusion rate over a given time interval. These parameters, derived from methodologies initially developed for wildfire monitoring (known as Fire Radiative Power), are now an integral part of operational satellite thermal analysis methods and provide a quantitative measure of the intensity of effusive eruptions. Beyond temperature, multispectral sensors make it possible to analyze the spectral signature of volcanic materials, that is, the variation in reflectance or emissivity depending on the wavelength. This property depends on the material's chemical composition, physical state, and age and can be used to identify and distinguish between different types of volcanic surfaces. For example, recent lava flows tend to show strong responses in the visible and near-infrared ranges, while older and more weathered flows reflect lighter due to surface oxidation. A representative application of this approach is the study conducted by [66] on Mount Etna, which defined reference spectral signatures for different volcanic material classes, such as fresh lava, ancient lava, dry tephra, and wet tephra. This type of information is essential for mapping eruptive products, monitoring the volcano's morphological evolution, and evaluating post-eruption residual activity.

Gas emission monitoring, instead, has become a crucial component of volcano surveillance, as changes in the rate and composition of volcanic gases often precede eruptive activity. Ground-based techniques, such as DOAS, Multi-GAS sensors, and FTIR spectroscopy, allow continuous measurements of SO_2, CO_2, and other species directly at the volcano. Meanwhile, satellite-based instruments like TROPOMI and OMI provide regional to global coverage, enabling the detection of large-scale degassing events and the tracking of volcanic plumes across thousands of kilometers [67]–[69]. These measurements are integrated with seismic, deformation, and thermal data in a multiparameter approach that allows early detection of changes in the magmatic system. In addition to optical and thermal sensors, radar remote sensing also plays a key role in satellite-based volcanic monitoring, particularly for detecting ground deformation. Synthetic Aperture Radar (SAR) systems, such as those on-board Sentinel-1 (ESA), ALOS-2 (JAXA), and TerraSAR-X (DLR), allow for the acquisition of interferometric images that can be processed using InSAR techniques to measure vertical changes in the surface with sub-centimeter precision [70]–[72]. The high spatial resolution and regular revisit frequency of SAR observations, especially those from Sentinel-1, which offers global coverage every 6 to 12 days, make it possible to systematically monitor active volcanoes worldwide, even in the absence of visible or thermal signals. Several studies also take advantage of the different polarization modes of SAR data to extract more detailed information [73]–[75]. Despite the many advantages offered by satellite remote sensing, it is important to acknowledge the operational and methodological limitations that may affect its effectiveness, particularly in the context of active volcanic monitoring. One of the main limiting factors is atmospheric conditions, which can significantly impact the quality of data acquired by optical and thermal sensors. Passive sensors, such as those operating in the visible (VIS), shortwave (SWIR), mid-infrared (MIR), and thermal infrared (TIR) bands, depend on the transparency of the atmosphere along the line of sight between the sensor and the Earth's surface. Dense cloud cover, ash-laden eruption plumes, condensed

vapors, or even simply high humidity levels can completely obstruct the view or alter the recorded signal. In particular, while the SWIR and MIR bands are highly sensitive to high temperatures, they are also highly susceptible to water vapor absorption and scattering effects. Although TIR is generally more robust under moderately cloudy conditions, it too can be affected by atmospheric attenuation or diffuse reflection in the presence of aerosols. In explosive scenarios, the volcanic plume itself can act as an optical barrier, preventing the detection of thermal features at the crater base. In such cases, surface thermal anomalies may be either underestimated or overestimated, depending on the specific situation. To mitigate these limitations, it is essential to adopt a multi-source strategy (Paragraph 2.1.2), which includes integrating observations from sensors operating in other spectral ranges (such as nighttime thermal infrared or microwave sensors), as well as using active sensors like SAR, which are unaffected by cloud cover and can provide observations under adverse weather conditions or in the absence of sunlight.

1.3.2 Advances in Atmospheric Studies

Atmospheric monitoring through satellite observations has become an essential component in understanding the processes governing chemistry, physics, and dynamics of the atmosphere on local, regional, and global scales [76]–[78]. This capability is particularly important for studying the interaction between volcanic activity and climate, as it allows for continuous observation, with high spectral detail, of the distribution and evolution of numerous volcanic gases and aerosols in the atmosphere [79], [80]. Satellites dedicated to atmospheric studies are designed to detect trace gases (such as SO_2, CO_2, NO_2, O_3), fine particulate matter, aerosols, volcanic clouds, and radiative parameters that influence Earth's energy balance [81]. Over the past decades, the development of sensors dedicated to atmospheric composition has led to the creation of high-resolution global datasets in both time and spectrum. These are fundamental for climate modeling [82], the validation of dispersion models [83], and the assessment of the radiative impact of volcanoes. Atmospheric sensors rely primarily on passive spectroscopic techniques in the Ultra Violet (UV), visible, near-infrared (NIR), and thermal infrared (TIR) bands. Depending on the observed wavelengths and the sensor architecture, they can be optimized to observe total column abundances in the troposphere and stratosphere, or even to derive more detailed vertical profiles [84]. Among the reference instruments for monitoring volcanic gases in the atmosphere are sensors onboard the Sentinel, Aura, MetOp, and Suomi-NPP missions. These can observe major greenhouse gases and sulfur species by measuring the differential absorption of solar radiation.

Key sensors include: TROPOMI (Tropospheric Monitoring Instrument, Sentinel-5P, ESA), currently the most advanced instrument for detecting SO_2, NO_2, CO_2, CH_4, and aerosols. Its spatial resolution (up to 5.5 km $\times$ 3.5 km) and spectral sensitivity make it a global benchmark for the study of both volcanic and anthropogenic emissions [67], [85]. OMI (Ozone Monitoring Instrument, NASA Aura), which has provided SO_2 and O_3 measurements since 2004, contributing essential time series for studying the impact of major explosive eruptions such as Kasatochi (2008) and Calbuco (2015) [79]. GOME-2 (Global Ozone Monitoring Experiment-2, EUMETSAT/MetOp), operational on three satellites, has been widely used to study stratospheric SO_2 plumes

[86] due to its long mission duration and broad spectral coverage. IASI (Infrared Atmospheric Sounding Interferometer, MetOp), an interferometric sensor operating in the TIR, is capable of detecting gases even at night and under moderate cloud cover. It is one of the few instruments able to determine the altitude of volcanic plumes and characterize their sulfuric acid content and other condensed species [87], [88]. These instruments are particularly effective at identifying large SO_2 emissions that reach the stratosphere. In addition to gases, many satellites are designed to observe suspended solid and liquid particles, including sulfate aerosols, volcanic ash, and clouds of condensed vapor [1], [89]. These observations are crucial for assessing aerosol optical depth (AOD), composition, and vertical distribution, as well as for evaluating their radiative impact and interactions with solar and infrared radiation. Among the key instruments, CALIPSO (Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations) stands out as a lidar mission capable of providing vertical atmospheric profiles with meter-scale resolution, making it one of the most valuable tools for studying aerosol properties and their role in atmospheric processes [89]. MODIS, VIIRS, and SLSTR, which use reflected and thermal radiation to estimate AOD, cloud temperature, and the radiative properties of volcanic plumes [83]. EarthCARE (ESA–JAXA, launched in 2024), which includes both active and passive sensors, such as ATLID (Atmospheric Lidar) and CPR (Cloud Profiling Radar), specifically designed for 3D analysis of clouds and aerosols, improving the detection of thin or long-lasting volcanic plumes. Satellite-based atmospheric observation plays a fundamental role not only in tracking the evolution of volcanic plumes but also in quantifying the potential climatic impact of an eruption. The injection of SO_2 into the stratosphere, for example, can lead to the formation of sulfate aerosols that remain suspended for months or even years, with well-documented global cooling effects, as seen following eruptions like Pinatubo (1991) and El Chichón (1982) [1], [90], [91].

1.3.3 Recent Developments and Upcoming Missions

In recent years, the growing demand for high spatiotemporal resolution observations, combined with the need to effectively monitor atmospheric and volcanic processes on a global scale, has driven the development of a new generation of satellite missions. These technological advances are significantly expanding the capabilities for detecting and analyzing volcanic phenomena and their effects on Earth's atmosphere. One of the most relevant trends is the enhancement of temporal resolution, made possible by sensors in geostationary orbit capable of acquiring data at sub-hourly intervals. This is particularly important for monitoring volcanic gases and the near-real-time evolution of plumes, aerosols, and volcanic clouds. Missions such as Sentinel-4 (ESA/EUMETSAT), GeoCARB (NASA), and TEMPO (NASA/NOAA) are specifically designed for this purpose: to provide continuous atmospheric composition observations over continental regions, with temporal resolutions as high as one observation per hour or less [92], [93]. The ability to detect rapid changes in the distribution of SO_2, NO_2, O_3, and CH_4 is crucial, for instance, to identify and track impulsive volcanic degassing events or variations in the fluxes of precursor gases.

In the field of active remote sensing, a major leap forward is represented

by the launch of the EarthCARE mission (ESA–JAXA, 2024), which combines active sensors (ATLID – Atmospheric Lidar; CPR – Cloud Profiling Radar) with passive instruments (Broadband Radiometer and Multi-Spectral Imager). EarthCARE is designed to deliver a three-dimensional representation of clouds and aerosols, contributing significantly to our understanding of Earth's radiative balance and aerosol–cloud interactions. Thanks to its vertical sensitivity, EarthCARE will be particularly valuable in analyzing the structure and extent of volcanic plumes, especially those reaching the stratosphere [94]. Other instruments, either under development or recently deployed, are introducing advanced hyperspectral capabilities. Notably, EnMAP (DLR) and PRISMA (ASI) can acquire high spectral resolution data in the VNIR and SWIR ranges. These instruments allow for detailed characterization of volcanic materials, as well as the estimation of physicochemical surface parameters and the detection of geochemical alterations related to hydrothermal phenomena [95]. In this evolving framework, upcoming missions such as NASA's Surface Biology and Geology (SBG), ESA's Land Surface Temperature Monitoring (LSTM), and the joint CNES–ISRO Thermal infrared Imaging Satellite for High-resolution Natural resource Assessment (TRISHNA) will further enhance our capacity to monitor volcanic activity. SBG will combine multispectral and hyperspectral imaging with thermal infrared observations, enabling simultaneous assessment of surface composition, temperature anomalies, and gas–aerosol interactions. LSTM and TRISHNA, both focused on high-resolution thermal infrared measurements, will be instrumental in detecting subtle thermal precursors of eruptive activity, mapping lava flow temperatures, and quantifying heat fluxes from fumarolic fields, even under variable atmospheric conditions. The synergy between these missions and existing atmospheric composition sensors will open new perspectives for integrated monitoring of volcanic degassing, thermal anomalies, and surface changes, from local to global scales. At the same time, in the radar sector, significant progress is being made in InSAR time-series techniques, driven by the high acquisition frequency of Sentinel-1 and the upcoming NISAR constellation (NASA/ISRO). NISAR will provide dual-polarization data in both the L and S bands, increasing the penetration through vegetated terrain and enhancing sensitivity to millimetric surface deformations [96]. This will have a substantial impact on the ability to monitor magmatic pressurization or the migration of hydrothermal fluids, even in complex environments subject to anthropogenic noise. Overall, these missions and instruments represent a paradigm shift in monitoring the interactions between volcanic activity and the atmosphere. They will not only improve early warning systems but also enable more realistic integration of observational data into numerical models, both for forecasting plume evolution and for evaluating the radiative and climatic impacts of volcanic emissions. However, important challenges remain, managing the enormous volume of data, developing more robust retrieval algorithms under complex conditions (e.g., simultaneous presence of ash, gases, and water vapor), and performing cross-validation with in situ observations. Looking ahead, the integration of geostationary and polar-orbiting observations, the synergy between active and passive sensors, and the fusion of satellite data with advanced modeling will form the foundation of a global, integrated, and predictive volcanic surveillance system, capable of addressing not only the challenge of volcanic risk but also the broader issue of climate change.

1.4 Observational and Modeling Tools in Volcanology and Climate Science

Progress in research within the fields of volcanology and climate science has been made possible thanks to the evolution of increasingly sophisticated observational and modeling tools. These tools enable the analysis of complex, multidisciplinary, and interconnected phenomena, such as volcanic eruptions and their impact on the climate system. The importance of having observational tools that, for example, allow for near real-time visualization, represents a major step forward not only for the volcanological community but also for the public. Indeed, ordinary users can now easily access additional information about nearby volcanoes, thereby developing a greater understanding and, above all, increased awareness of the volcano itself and the phenomena that may affect it. Several operational systems have been developed to support real-time monitoring and risk assessment.

For instance, MIROVA (Middle InfraRed Observation of Volcanic Activity) utilize data from MODIS and VIIRS to provide near-real-time thermal monitoring of volcanoes worldwide, enabling rapid detection of changes in volcanic heat flux [63]. MOUNTS (Monitoring Unrest from Space) integrate thermal, gas, and deformation data from multiple satellite missions, offering multiparametric analyses to support surveillance of active volcanoes [97]. VolcHazard, developed by TechnoLab research group of INGV, combines multi-satellite for volcanic hazard assessment and decision support - VolcHazard. COMET uses satellite measurements alongside ground-based observations and geophysical models to study earthquakes and volcanoes - COMET. Other tools such as, NHI (Normalized Hot-spot Indices) allows the investigation of high-temperature volcanic features under GEE (Google Earth Engine), through time series analyses [98]. V-STAR (Volcanic Satellite Thermal Anomalies Recognition), that is part of this thesis, focuses on the detection and quantification of volcanic thermal anomalies to improve early warning capabilities [99]. These platforms are essential for transforming large volumes of raw data into actionable information for scientists, civil protection agencies, and decision-makers, enabling rapid situational awareness and supporting timely risk mitigation measures. In the field of volcanology, observational tools include remote sensing technologies, geophysical monitoring networks (such as seismology, ground deformation, and volcanic gas emissions), as well as in situ measurements. These data streams are often used as inputs for numerical models, which simulate and predict the various phenomena associated with volcanic activity. In many cases, modeling tools rely on data collected from external sources, which are then integrated into the models through a process known as data assimilation [100]–[102]. This approach is essential not only to initialize the models accurately, but also to continuously correct and refine them, ensuring that the simulations remain consistent with the observed behavior of the volcanic system. Alongside the observational tools mentioned above, a range of modeling tools has been developed over the years to serve different purposes, particularly in hazard assessment and management. These models are used to simulate lava flow propagation, volcanic plume dispersion, tephra fallout, and pyroclastic density currents. Due to the intricate nature of these physical phenomena, mathematical modeling plays a key role in both fundamental and applied sciences, providing

critical insights and supporting accurate predictions of volcanic system behavior. These tools are essential for anticipating potential scenarios and for informing timely decision-making by civil protection authorities and scientists alike. In both volcanology and climatology, these models are based on a mathematical representation of physical processes, solved using computational techniques. The effectiveness of a model largely depends on how space and time are discretized, and on its ability to integrate observational data through processes such as data assimilation. There are several approaches to numerical modeling, each with specific characteristics that make it more suitable for certain types of phenomena. A first distinction concerns the reference frame used: Eulerian, Lagrangian, or a combination of both (the ALE – Arbitrary Lagrangian-Eulerian approach). Another classification is based on whether the model uses spatial grids, distinguishing between mesh-based and mesh-free methods. Lastly, there are alternative models such as Cellular Automata, which are based on local rules rather than differential equations. In the Eulerian model, observation occurs at fixed points in space: the system (e.g., a fluid or volcanic cloud) evolves by moving through these points.

A notable example in volcanology is FALL3D, a model used to simulate the atmospheric dispersion of volcanic ash following an eruption [103]. Another example is PlumeRise, developed by the University of Bristol, which simulates the rise of eruptive columns during explosive eruptions in stratified atmospheric conditions [104]. In the Lagrangian model, on the other hand, the observation follows individual particles or volumes of material as they move. In volcanology, Tephra2 is used to calculate the spatial distribution of volcanic deposits [105], while ASH3D, developed by the USGS, offers more advanced 3D capabilities in the same context [106]. ALE (Arbitrary Lagrangian-Eulerian) models combine the two previous approaches, allowing greater flexibility in handling phenomena involving large deformations while maintaining good spatial resolution. They are often applied to simulate pyroclastic density currents (PDCs), using ALE solvers available in platforms such as OpenFOAM [107]. For instance, [108] used this approach to simulate PDC propagation on volcanic slopes.

Another important distinction concerns the spatial structure of the model. In mesh-based methods, the domain is represented by a fixed or adaptive grid that defines connections between points. This approach is highly accurate but can be complex to manage when dealing with significant deformations or irregular geometries. It is often used in Finite Element Method (FEM) models, such as those simulating ground deformation due to magma pressure. By contrast, mesh-free methods do not use grids but represent space with a set of independent particles. These models are particularly effective for simulating large deformations, free surfaces, or complex domains. One example is the use of SPH (Smoothed Particle Hydrodynamics) to simulate lava flows, as applied in works by [109], [110]. Finally, an alternative approach widely used in volcanology is that of Cellular Automata (CA). These are discrete models in which space is represented by a grid of cells, each of which evolves according to simple, often empirical, local rules. Although not directly based on physical equations, CAs are extremely effective in simulating territorial and diffusive processes, thanks to their simplicity and computational efficiency. Notable examples include MAGFLOW, developed to improve the understanding of lava emplacement [111], and SCIARA, designed to simulate the propagation of lava flows on Mt. Etna [112]. In summary, the choice of numerical model

depends on the nature of the phenomenon being studied, the required level of detail, and the available computational resources. Each approach has its strengths and limitations, but all contribute significantly to understanding and managing volcanic hazards.

The same fundamental concepts, such as the type of discretization, the use of Eulerian or Lagrangian approaches, and the implementation of mesh-based or mesh-free models, are also found in climate models, adapted to the scale and complexity of atmospheric and oceanic processes. Numerical modeling in climatology is used to simulate the dynamics of physical systems, particularly the atmosphere, oceans, and interactions with the land surface, biosphere, and cryosphere. Climate models, while built on complex physical equations (such as the Navier-Stokes equations and the laws of thermodynamics), make use of numerical approaches that fall into the same categories previously described. A relevant example of a Lagrangian model in the climate domain is HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory). In climatology, it is employed to simulate air mass trajectories, the dispersion of atmospheric pollutants, or the deposition of aerosols (including volcanic aerosols) on a regional or global scale [113]. The model tracks the path of virtual particles immersed in a three-dimensional atmospheric wind field, interpolated over a mesh-based grid. This overview naturally leads us to global climate models (GCMs – Global Climate Models), which represent the primary tools for studying and forecasting the evolution of Earth's climate [114], [115]. GCMs integrate atmospheric, oceanic, land surface, and cryospheric processes on a global scale, solving complex equations over three-dimensional grids. They represent a more advanced evolution of the numerical models discussed so far. Since Global Climate Models (GCMs) operate on very large grids (typically 100–300 km in resolution), they are unable to capture regional details such as mountain ranges, coastlines, or local microclimates. This is why Regional Climate Models (RCMs) are used, through a technique known as dynamical downscaling: the outputs of a GCM are used to drive a finer-scale model (typically 1–50 km), which incorporates local features to simulate climatic phenomena with higher spatial detail [116], [117]. The Coupled Model Intercomparison Project (CMIP), currently in its sixth phase (CMIP6), provides an international framework for the standardized comparison of different climate models [118]. CMIP results are widely used in IPCC reports and form a critical foundation for climate research, including that related to volcanic impacts. Alongside forward-looking models, climate reanalyses, such as ERA5, MERRA-2, and JRA-55, play a key role [119]–[121]. These datasets combine numerical models with historical observations through data assimilation techniques, producing a consistent and continuous reconstruction of past atmospheric conditions. Reanalyses are crucial tools for assessing the climatic impact of past volcanic eruptions, offering information on atmospheric parameters before and after such events. In recent decades, research has increasingly focused on the climatic impact of volcanic eruptions, particularly explosive events that reach the stratosphere. Eruptions injecting sulfate aerosols into the stratosphere can cause negative radiative forcing, resulting in global cooling effects lasting from several months to a few years. Climate models participating in CMIP, when forced with eruption-specific input data (e.g., SO_2 estimates, plume height, aerosol size distribution), can simulate these impacts. However, the representation of volcano-climate interactions remains a key source of uncertainty in climate models, especially

regarding aerosol microphysics, the latitudinal and longitudinal distribution of volcanic clouds, and their interaction with atmospheric dynamics. Recent efforts within CMIP6, particularly through the VolMIP (Volcanic Model Intercomparison Project) and ISA-MIP (Interactive Stratospheric Aerosol MIP) initiatives, have aimed to standardize volcanic injection scenarios and compare the climate responses simulated by different models. These initiatives seek to improve the understanding of feedback mechanisms between volcanism and climate, ultimately providing more reliable projections of the future impacts of major eruptions.

1.5 Overview of Artificial Intelligence in Earth System Science

In recent years, artificial intelligence (AI) has assumed an increasingly strategic role in the field of Earth System Sciences, playing a decisive part in the processing, integration, and interpretation of geospatial data on a global scale. This progress aligns with the so-called "Fourth Scientific Revolution", marked by exponential growth in the availability of data from heterogeneous sources, such as satellite observations, ground-based sensors, and numerical models, which has promoted a transition toward a new scientific paradigm: data-driven science [122], [123]. In this context, AI emerges as a particularly powerful tool, capable of building predictive models, identifying latent patterns in data, automating complex analytical processes, and producing near real-time estimates. This represents a paradigm shift from traditional methods based on fixed thresholds or explicit rules, which are often inadequate to capture the non-linear and dynamic complexity of natural systems [124]. In the field of volcanology, a discipline historically grounded in observational approaches and terrestrial monitoring networks, AI has found application in response to the need for continuous and systematic analysis of multiparametric signals. The scientific community has progressively adopted Machine Learning (ML) and Deep Learning (DL) techniques, which can autonomously learn from datasets and generalize to new scenarios. Numerous recent studies have employed AI for detecting unrest signals [125], identifying eruption onsets [11], and tracking the spatial and temporal evolution of eruptive phenomena [126], [127]. For example, [61] used a UNet architecture to analyze thermal satellite images, achieving 93% accuracy in detecting anomalies, including early signals during the 2021 degassing crisis at Vulcano (Italy). Similarly, [27], inspired by the mechanisms of the human visual system, used Gabor features to detect precursor thermal variations in 13 out of 16 eruptions up to a week in advance.

These examples demonstrate the potential of AI to handle complex and large-scale datasets. Traditional thresholding, rule-based, or statistical regression methods remain highly effective in many operational contexts, particularly where physical parameters are well constrained and noise levels are low. In contrast, AI models tend to outperform classical methods mainly under conditions of high data heterogeneity, non-linearity, or when prior knowledge is incomplete. Therefore, the added value of AI lies less in universally superior accuracy, and more in its flexibility, scalability, and ability to generalize across diverse volcanoes, sensors, and environmental conditions. Moreover, the operational adoption of AI-based tools in observatories remains limited.

This is due to several factors: (i) the need for extensive and high-quality labeled datasets; and (ii) the computational resources required for both training and real-time deployment. As a result, while AI has introduced significant conceptual and methodological advances, its integration into routine monitoring workflows still represents an ongoing challenge rather than an established standard.

In the field of thermal anomalies, Convolutional Neural Networks (CNNs) have outperformed traditional methods like MODVOLC [128], offering greater sensitivity and robustness in detecting weak or diffuse signals. A rapidly growing area is the application of AI to InSAR data. In a pioneering study, [10] applied a CNN with transfer learning (AlexNet) to over 900 volcanoes, automatically classifying interferometric fringes to identify deformation, even in interferograms not corrected for atmospheric effects. [129] improved this approach using a neural network trained on synthetic data, capable of detecting subtle volcanic deformation in multi-temporal stacks and distinguishing between different deformation styles [130].

In the atmospheric domain, ML techniques are used to estimate gas and aerosol concentrations, model radiative processes, and improve satellite observation products. In particular, the identification and quantification of volcanic SO_2 has attracted significant interest due to its relevance in eruptive processes and global radiative balance. [131] proposed using neural networks to detect SO_2 from AVHRR data, while [69] developed a Random Forest algorithm using TROPOMI UV data, integrating plume height information from auxiliary products to improve gas mass estimates. The model, fully implemented on Google Earth Engine, was validated on several active volcanoes and showed robustness and scalability. For explosive eruptions, [126] proposed a VGG16-UNet architecture for automatic segmentation of volcanic clouds containing ash and SO_2, using SEVIRI Ash RGB imagery. At the same time, AI is also being applied to climate simulation. Emulative models, such as FourCastNet [132], can replicate atmospheric dynamics in significantly reduced computation times, opening new possibilities for analyzing interactions between volcanic aerosols and radiation. However, the global adoption of AI presents significant infrastructural challenges. Earth observation datasets require high computational capacity and scalable architecture.

In this regard, cloud computing has emerged as a key enabler for large-scale data processing and model training, providing scalable computational resources and facilitating global collaboration [133], [134]. At the same time, the emerging field of quantum computing holds the potential to further accelerate the training and optimization of complex AI models, particularly through quantum-enhanced algorithms for optimization and pattern recognition [135], [136]. Increasing attention is also being directed toward approaches such as Explainable AI (XAI), self-supervised learning, and transfer learning, which aim to make models more transparent, adaptable, and generalizable—critical requirements in high-risk operational contexts. Nevertheless, significant challenges remain concerning data quality, representativeness, and bias management [137].

In recent years, a new generation of models known as Foundation Models has emerged, designed to learn general representations from extremely large datasets and then be adapted to various specific tasks through fine-tuning [138]. Initially developed in the fields of natural language processing and computer vision, these models are rapidly finding application in EO due to their

ability to handle multispectral, temporal, and geographically diverse data. Foundation Models help overcome data scarcity by leveraging self-supervised and transfer learning, allowing the creation of generalist models adaptable to different spatial and temporal scales. A notable example is SatMAE [139], a self-supervised model for multispectral satellite imagery, capable of extracting informative features about environmental changes, vegetation, land cover, and natural disaster signals. Similarly, models like ClimateGPT, FourCastNet, or Pangu-Weather [140] demonstrate how Foundation Models can be applied to high-resolution weather and climate forecasting, representing a promising alternative to traditional numerical models, especially for intensive simulations and near real-time operations. In volcanology, these models could enable simultaneous analysis of thermal, SAR, spectroscopic, and environmental data, improving predictive capabilities and the recognition of precursory patterns. The integration of Foundation Models with multi-modal architectures, combining visual, textual, and temporal inputs, opens new possibilities for building intelligent systems capable of understanding and synthesizing complex phenomena, such as the interaction between volcanic activity, the atmosphere, and climate. While this is still an emerging field, the transformative potential of Foundation Models in Earth Sciences is substantial, and their use will likely be at the center of future research and monitoring strategies. In conclusion, volcanology and climate science, traditionally based on direct observation and expert interpretation, is undergoing a radical transformation. Artificial intelligence not only automates monitoring but amplifies its interpretive power, enabling the discovery of latent relationships and the understanding of complex dynamics. This convergence of data, physical models, and computational techniques marks a decisive step toward a new era in volcano science and in the analysis of interactions between volcanic activity and the atmosphere.

1.6 Objectives and Structure of the thesis

This thesis aims to explore and develop advanced techniques for volcanic monitoring across all its key phases, forecasting, detecting, tracking, quantifying, and nowcasting, to construct a comprehensive picture of how a volcanic eruption unfolds over time. Particular attention is given to how each of these phases can enhance our understanding of the interactions between volcanic activity and the Earth's climate system. Specifically, the work focuses on the integration of observational technologies, remote sensing data at multiple spatial and temporal resolutions, and artificial intelligence techniques. This approach is intended to contribute not only to hazard mitigation but also to a deeper understanding of the role of volcanic emissions in altering atmospheric composition and radiative forcing. The primary challenge addressed is the development of integrated and efficient monitoring systems capable of operating on the global scale, supported by tools that can generate scenario-based analyses as eruptive conditions evolve, an essential component of volcanic hazard and climate impact assessment.

The thesis is structured as follows:
Chapter 1 provides a scientific background on volcanic systems, types of eruptions, and the mechanisms through which volcanic activity influences the atmosphere and climate. It also discusses observational and modeling tools,

advances in satellite remote sensing, and recent applications of AI in Earth system science.
Chapter 2 presents an overview of the current state of the art in satellite remote sensing and AI technologies applied to volcanic and environmental monitoring. It describes the main EO instruments, data fusion strategies, and preprocessing techniques, as well as recent developments in AI and ML for Earth Observation.
Chapter 3 introduces a multi-step AI-based framework for monitoring volcanic activity using satellite data. It focuses on forecasting an eruption, detecting eruption onset, tracking spatio-temporal evolution, and quantifying volcanic products relevant to hazard assessment and nowcasting.
Chapter 4 introduces the selected case studies and demonstrates how the proposed framework (from forecasting to nowcasting) is applied to real-world data, integrating multi-source observations with AI-driven analysis to assess eruptive behavior and potential climatic implications, including the 2022 Hunga Tonga–Hunga Ha'apai event.
Chapter 5 concludes the thesis by summarizing the key findings, discussing their implications, and suggesting directions for future research in volcanic monitoring and climate-related risk assessment. Through this structure, the thesis aims to provide a comprehensive and interdisciplinary contribution to the fields of volcanology, climate science, engineering, and environmental monitoring, offering tools and insights that are both scientifically robust and operationally relevant.

Contribution Statement

This thesis provides multiple and complementary contributions that advance both volcanic hazard monitoring and the understanding of volcano–climate interactions. The research follows the entire analytical chain, from forecasting to nowcasting, demonstrating how Artificial Intelligence (AI) can support the monitoring of different volcanic phenomena across multiple regions and spectral domains of the electromagnetic spectrum. Through this multi-step workflow, the thesis enables the detection, classification, tracking, and quantification of volcanic activity, thereby enhancing operational capabilities for hazard assessment. Beyond hazard monitoring, the thesis extends the same forecasting-to-nowcasting framework to the study of volcanic–climatic interactions. By integrating reanalysis products, it explores how volcanic forcing perturbs atmospheric dynamics, bridging methodologies from volcanology and climate science within a unified data-driven perspective. The main methodological innovations concern the implementation of feature-engineering strategies and the design of ad hoc AI architectures tailored to each specific phenomenon under study. These developments, combined with the effective use of diverse satellite datasets, demonstrate both advanced geospatial data handling and a deep understanding of the physical and technological principles underlying volcanic processes and Earth observation systems. From a scientific standpoint, this work contributes novel insights into the understanding of volcanic activity by linking physical processes and remote-sensing observables through explainable AI. It demonstrates how data-driven models can complement traditional physics-based approaches, enabling faster, scalable,

and more interpretable analyses. Moreover, by extending volcanic monitoring techniques into the climate domain, the thesis opens new pathways for interdisciplinary research, fostering the integration of geophysical and atmospheric sciences within a unified analytical framework.

Chapter 2

State of the Art and Technological Background

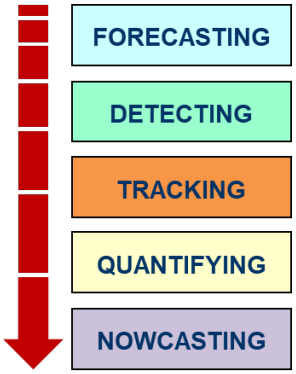

FIGURE 2.1: Schematic representation of the main stages in volcanic monitoring and hazard assessment: forecasting, detecting, tracking, quantifying, and nowcasting.

The primary uses of satellite data in volcanology are forecasting, detecting, tracking, quantifying, and nowcasting volcanic activity as shown in Figure 2.1. The *forecasting* phase involves identifying early indicators of unrest, such as subtle increases in surface temperature, deformation patterns, or changes in volcanic gas emissions, that are detectable from space days to weeks before an eruption. These precursors, when integrated with ground-based observations, improve the capacity to anticipate potential eruptive events and reduce uncertainties in hazard scenarios. *Detection* refers to the rapid recognition of eruptive onset, often achieved through the automated analysis of satellite imagery. Thermal infrared sensors can capture sudden and localized thermal anomalies associated with fresh lava or dome growth, while multispectral and hyperspectral instruments detect ash and gas signatures (e.g., SO_2) in the

atmosphere. Early detection is essential for issuing timely alerts to civil protection agencies. Once an eruption is underway, *tracking* tools use time-series satellite imagery to monitor the evolution of volcanic phenomena. This includes mapping the emplacement and progression of lava flows, quantifying the extent of pyroclastic deposits, characterizing ash plume morphology and altitude, and detecting shifts in eruptive style or intensity. Beyond tracking, *quantification* focuses on deriving precise numerical estimates from satellite data, such as SO_2 fluxes, ash mass loading, thermal output, and the areal extent of volcanic deposits. These parameters are key inputs for dispersal models, climate impact assessments, and long-term volcanic budget calculations. The most recent addition to the operational chain is *nowcasting*, which combines near-real-time satellite observations with advanced modeling and machine learning to provide an up-to-date, dynamic images of ongoing activity. By integrating continuous satellite data streams, potentially from both polar-orbiting and geostationary platforms, nowcasting enables rapid updates on plume dispersion, lava flow advance, or changes in thermal intensity, facilitating immediate decision-making.
Despite the increasing availability of satellite data for forecasting, detection, and tracking, most operational systems still rely on threshold-based or semi-empirical approaches. These methods often lack adaptability across different volcanic settings and sensor modalities, leading to inconsistent detection performance and delayed response. Moreover, the integration between near-real-time satellite observations and modeling frameworks remains limited, which constrains the capacity for dynamic hazard nowcasting. These gaps motivate the development of AI-based, data-fusion-driven methodologies explored in the following chapters.
In this context, part of the current research efforts are focusing on developing web-based platforms that leverage artificial intelligence applied to multispectral satellite imagery to improve hazard monitoring. The workflow presented here will be central to the discussion in Chapter 3, where it will be applied to volcanic hazard monitoring, and will be further extended in Chapter 4 to address the evaluation of volcanic eruption impacts on climate.

2.1 Satellite Remote Sensing

Satellite remote sensing data, as initially distributed by space agencies, are organized according to standardized processing levels that indicate how far the data have been transformed from their raw form. At the most basic stage, Level 0 contains the unprocessed measurements recorded by the satellite instruments, cleaned only of communication artifacts and still expressed in instrument counts. These are then converted into Level 1 products, in which the measurements are radiometrically calibrated and time-referenced, with geolocation metadata included. Depending on the mission, Level 1 may appear as Level 1A (annotated but not yet radiometrically converted), Level 1B (calibrated to radiance or brightness temperature), or Level 1C (such as for Sentinel-2, already corrected for geometric distortions and expressed as Top-Of-Atmosphere reflectance). Level 2 products go a step further, transforming these calibrated measurements into geophysical variables such as surface reflectance, sea surface temperature, or atmospheric gas concentrations, usually at the original spatial resolution of the sensor. From here, Level 3

products aggregate and reproject the data into uniform spatial and temporal grids, enabling large-scale or long-term analysis through daily, monthly, or seasonal composites. At the highest stage, Level 4 products are generated by combining satellite observations with models or ancillary datasets, producing value-added variables such as interpolated global fields, climate indices, or derived hazard maps.
Understanding these distinctions is essential for selecting the most appropriate dataset for a given scientific objective. For instance, Sentinel-2 imagery is provided in both Level-1C and Level-2A formats, the latter being atmospherically corrected and thus immediately suitable for quantitative analysis. In the Synthetic Aperture Radar (SAR) domain, Level-0 products contain raw, unfocused radar data; Level-1 products consist of focused imagery such as Single Look Complex (SLC) or Ground Range Detected (GRD) data; Level-2 products deliver application-ready geophysical parameters; while higher-level products may integrate multi-source datasets to support advanced geophysical modeling and analysis. This hierarchy reflects the progression from raw sensor measurements to analysis-ready and model-enhanced products, guiding both the preprocessing needs and the analytical possibilities for each mission. However, the diversity of processing levels and data formats across missions also introduces challenges for interoperability [141]. Differences in spatial resolution, calibration standards, and cloud-masking algorithms can lead to inconsistencies that propagate into downstream analyses, limiting comparability across datasets and hindering large-scale, multi-sensor studies.

2.1.1 Instruments

The analysis conducted in this thesis was based on the integration of data from several satellite sensors, each with specific characteristics in terms of spatial resolution, temporal resolution, and spectral coverage, as well as operational advantages and constraints that influence their use. Among the main instruments, Sentinel-2, equipped with the MSI multispectral sensor, provides high-resolution imagery: 10 metres for bands in the visible and near-infrared, 20 metres for red-edge and SWIR bands, and 60 metres for atmospheric correction bands, with a revisit time of about five days when combining Sentinel-2A and Sentinel-2B. In this study, particular emphasis was placed on the shortwave infrared bands, specifically B11 (1.610 µm) and B12 (2.190 µm), which are sensitive to high surface temperatures and therefore useful for detecting very hot bodies such as lava flows or wildfires. The main advantage of Sentinel-2 lies in its high spatial and spectral resolution, which enables detailed analysis even of localized phenomena; however, as a polar-orbiting optical sensor, data quality is strongly affected by cloud cover and the availability of daytime acquisitions. Landsat 8 and Landsat 9, equipped with the OLI and TIRS sensors, offer a spatial resolution of 30 meters for multispectral bands, 15 meters for the panchromatic band, and 100 meters for thermal bands, with a revisit time of 16 days, reduced to eight days when combining both platforms. Thermal bands B10 (10.60–11.19 µm) and B11 (11.50–12.51 µm) allow the estimation of surface temperature, while SWIR bands B6 (1.57–1.65 µm) and B7 (2.11–2.29 µm) enhance the detection of very high-temperature heat sources. The key strength of Landsat lies in its long historical archive, which enables multi-decadal analyses and long-term change detection; limitations include the sensitivity to cloud cover typical of optical

sensors, the lower revisit frequency compared to Sentinel-2, and the coarser thermal resolution. For atmospheric monitoring, particularly of sulphur dioxide emissions, TROPOMI onboard Sentinel-5P was employed, providing data with a spatial resolution of up to 3.5 × 5.5 km and daily global coverage. SO_2 detection is based on the gas's characteristic absorption in the ultraviolet range, between 310 and 330 nm, with peaks in sensitivity around 313, 325, and 332 nm, allowing the identification and quantification of volcanic and anthropogenic emissions on a large scale. The main advantage of TROPOMI is its ability to monitor vast areas and detect diffuse plumes, while its primary limitation is the relatively low spatial resolution, which makes it difficult to pinpoint the exact source in complex areas or where multiple emission points are close together. Near real-time monitoring was ensured by the geostationary sensors SEVIRI, onboard second-generation Meteosat satellites, and FCI, onboard the third generation. SEVIRI offers a resolution of 1 km for visible and near-infrared bands and 3 km for infrared bands, with a temporal resolution of 15 minutes, and includes thermal channels at 3.9 µm, 10.8 µm, and 12.0 µm that are useful for surface temperature estimation. FCI enhances these capabilities, providing resolutions of 0.5–1 km in the visible and near-infrared and 1–2 km in the infrared, with a 10-minute revisit time over the Euro-African area and a greater number of SWIR and IR bands. The most evident advantage of geostationary sensors is their rapid acquisition rate, which allows the evolution of dynamic phenomena to be followed almost in real time; however, their main drawback remains the lower spatial resolution compared to polar-orbiting sensors. Additionally, data from other reference sensors were considered, such as MODIS, onboard Terra and Aqua, which with its 36 bands and spatial resolution between 250 and 1000 meters ensures daily global coverage and is useful both for detecting thermal anomalies and producing atmospheric products; VIIRS, onboard Suomi-NPP and NOAA-20, with a resolution of 375–750 meters and the unique Day/Night band that enables the detection of light and thermal sources even at night; and ASTER, onboard Terra, with a resolution of 15 meters in the visible and near-infrared, 30 meters in the SWIR, and 90 meters in the thermal range, suitable for high-precision thermal and geological studies.

These complementary instruments extend temporal and spectral coverage but remain limited by acquisition availability and operational scheduling. Their integration combines high-resolution detail with broad temporal coverage, mitigating individual sensor constraints and enabling more continuous observation of the studied phenomena.

Nonetheless, such integration remains non-trivial. Differences in spatial and temporal sampling often cause misalignment and inconsistencies in multisensor datasets. Moreover, active and passive sensors exhibit varying sensitivity to atmospheric conditions, illumination geometry, and acquisition time, which may introduce systematic biases when comparing or merging products. These challenges underline the importance of advanced data fusion strategies capable of harmonizing heterogeneous observations [142]–[144]. The need to merge heterogeneous information in terms of spatial, temporal, and spectral resolution has therefore made the adoption of data fusion approaches essential, enabling the integration and harmonization of data from different platforms, maximizing their informational content and improving the reliability of the analyses. In the Earth observation domain, the effective combination of datasets with varying spatial, temporal, and spectral properties has proven

invaluable for improving detection capabilities, enhancing the accuracy of environmental monitoring, and extending the scope of analyses beyond the limits of any single instrument. The following section will therefore explore in greater detail the significance of data fusion in satellite applications, highlighting its benefits and presenting selected examples from scientific literature where its adoption has yielded particularly successful results.

2.1.2 Data Fusion

In the field of satellite-based Earth observation, data fusion refers to the integration of data from different sources, platforms, or acquisition modes to produce a dataset that is richer and more informative than could be obtained from a single source alone. This approach is justified by the intrinsic trade-offs of individual satellite systems: a sensor with high spatial resolution may have a low revisit frequency, while one with high temporal resolution may lack the spatial detail required for certain analyses; some instruments are optimized for surface observation, others for atmospheric composition; some cover specific spectral ranges, while others offer broader coverage but at coarser resolution. By combining complementary datasets, data fusion helps to mitigate these limitations, improving both the accuracy and robustness of the final products. The importance of data fusion is well established in the literature, where it has been shown to enhance monitoring capabilities across a wide range of applications [144]–[146]. Numerous fusion techniques have been proposed, from simple pixel-based blending to advanced deep learning approaches, yet their performance largely depends on the spatial, temporal, and spectral compatibility of the input datasets. Early linear methods, though computationally efficient, often fail to maintain physical consistency, whereas model-based and machine learning approaches generally yield improved accuracy at the expense of interpretability and higher data requirements. Moreover, most existing fusion frameworks remain sensor-specific, limiting their transferability across different missions and application domains.
In volcanic monitoring, for example, integrating high-resolution optical data (such as Sentinel-2 or Landsat) with the high temporal frequency of geostationary platforms (e.g., SEVIRI or FCI) enables detailed spatial characterization alongside near real-time tracking of dynamic phenomena. Likewise, combining thermal anomaly detection from multispectral imagery with gas emission measurements (e.g., SO_2) from atmospheric spectrometers such as TROPOMI provides a simultaneous view of eruptive activity and related emissions [147], [148]. A notable example is the 2021 study of the Cumbre Vieja volcano eruption (La Palma), where multi-sensor fusion enabled more temporally and spatially precise mapping of SO_2 plumes by integrating data from TROPOMI, CrIS (Cross-track Infrared Sounder), and ABI (Advanced Baseline Imager) [149]. Similarly, more recently, the RSDF (Remote Sensing Data Fusion) methodology developed for near real-time volcanic radiative power (VRP) analysis has exploited data from MODIS, SLSTR, VIIRS, and SEVIRI, demonstrating greater sensitivity in detecting thermal sources compared to standard products [57]. Beyond these application-specific cases, the specialized literature also offers comprehensive methodological reviews. In particular, [150] analyze several spatiotemporal fusion techniques developed to generate imagery with both fine spatial detail and high temporal frequency by combining infrequent high-resolution images with frequent but

coarse-resolution ones. Furthermore, [151] provides an overview of the challenges and advances in multi-source and multi-temporal fusion, which now enables the manipulation of data in four dimensions (3 spatial + time) to extract complex information. For a broader perspective, [145] reviewed machine learning techniques applied to data fusion in remote sensing, highlighting the potential of new technologies to exploit the large heterogeneity of available data.

These limitations reveal that, despite substantial progress, there is still no universal or operationally standardized fusion framework for volcanic monitoring. The challenge lies in balancing spatial detail, temporal continuity, and computational efficiency, an aspect that will be addressed in the data-driven integration approach proposed in Chapter 3.

In summary, data fusion proves to be crucial not only for overcoming the operational limitations of individual sensors but also for building consistent time series, expanding the informational spectrum, and producing more reliable results. In Chapter 3, after the presentation of a series of methodologies and case studies, a possible application of data fusion in the volcanological field will be discussed. This approach, aimed at volcanic hazard monitoring, seeks to intelligently integrate the previously described datasets through AI techniques, ensuring both consistency and continuity in volcanic risk management.

Beyond its technical role in enhancing resolution and coverage, data fusion also has a deeper conceptual value: it enables a more integrated understanding of the physical mechanisms driving volcanic activity. Each satellite observation, thermal, spectral, or interferometric, captures a different manifestation of the same underlying processes, such as magma ascent, degassing, or heat transfer. When these datasets are jointly analyzed, it becomes possible to disentangle their respective contributions and to connect surface expressions with subsurface dynamics. For instance, simultaneous increases in thermal emission (from optical sensors) and SO_2 flux (from spectrometers) can indicate enhanced magma degassing at shallow levels, while coeval deformation signals detected by SAR may reveal pressurization within the plumbing system. In this sense, data fusion does not merely generate more observations; it provides a framework for causal inference, linking measurable signals to their physical origins. Such an integrative perspective is essential for developing more accurate conceptual and numerical models of volcanic systems, and it represents the bridge between empirical monitoring and process-based understanding.

2.1.3 Preprocessing

Before such data can be used effectively for scientific analysis or machine learning applications, they require a series of preprocessing steps to correct, clean, and structure the information into usable, analysis-ready datasets. A primary challenge lies in applying radiometric and atmospheric corrections to optical data, ensuring that variations in illumination, atmospheric conditions, and sensor performance do not bias the results. Sentinel-2 Level-1C products, requires external tools or algorithms to achieve surface reflectance values comparable across time and space. Another critical step is quality control using QA (Quality Assurance) flags or scores provided in the product metadata. For example, TROPOMI (Sentinel-5P) Level-2 SO_2 products

include a qa_value field (0 = unreliable, 1 = high confidence); filtering with thresholds (e.g., qa_value > 0.5) is essential to remove measurements affected by clouds, sunglint, or sensor noise.
Equally important is geospatial harmonization: satellite data are distributed in varying coordinate systems, projections, and resolutions. For multi-source integration or machine learning pipelines, it is often necessary to reproject, resample, and align datasets onto a common, regular grid. This facilitates pixel-level fusion, time-series analysis, and efficient storage in formats such as NetCDF, GeoTIFF, or multidimensional arrays (e.g., xarray, zarr). These preprocessing operations, although crucial, are often computationally intensive when applied at scale or across multiple sensors. For this reason, cloud-based geospatial platforms such as Google Earth Engine (GEE) are becoming increasingly relevant. GEE offers on-demand access to petabytes of data from missions like Sentinel, Landsat, and MODIS, many already processed to Level-2 or higher. It also automates common preprocessing tasks, cloud masking, terrain correction, temporal compositing, enabling researchers to focus on analysis. In this thesis, both strategies were adopted: GEE was used for rapid prototyping and scalable access (e.g., Sentinel-2 and TROPOMI), while raw downloads from ESA and NASA archives were processed manually when fine control over preprocessing (e.g., QA filtering, spatial alignment) was needed for precise model training and validation.

2.2 AI Technologies

Artificial intelligence (AI) provides innovative and effective tools for pattern recognition, automated monitoring, and forecasting of eruptive events for volcanic phenomena. This thesis adopts an integrated methodological approach that combines different classes of AI algorithms, selected based on the nature of the available data and the scientific objectives to be achieved.
This methodological shift reflects a broader epistemological transition within the geosciences, where deterministic, rule-based systems are increasingly being complemented by probabilistic, data-driven approaches rather than being replaced outright. This evolution has been documented in several domains, including climate modeling, hydrology, and geohazard forecasting, where hybrid frameworks now integrate physical constraints with data-driven inference [123], [152], [153]. In such approaches, machine learning does not substitute physical understanding but extends it, allowing models to capture non-linear dependencies and uncertainties that are difficult to represent explicitly in traditional formulations. In this sense, AI provides a functional framework that complements physical modeling by enabling the discovery of patterns and relationships that may not be explicitly represented in current process-based formulations. Rather than replacing mechanistic understanding, these data-driven insights can highlight latent connections or precursory behaviors that warrant further physical interpretation. This synergy between data-driven inference and process-level modeling is essential in volcanology, where systems evolve over time and purely statistical approaches risk failing under changing conditions. Integrating AI with physically grounded models therefore represents a pathway toward more resilient and explanatory frameworks for volcanic monitoring.
The works presented in Chapter 3 demonstrate how the use of AI allows

for overcoming the limitations historically imposed by rigid, threshold-based methods, shifting towards information that is more deeply grounded in both data knowledge and operational context.
However, the effectiveness of AI methods in EO strongly depends on the quality and representativeness of training data. Overfitting, label scarcity, and domain shift across geographic regions remain persistent challenges, often constraining model transferability. Furthermore, the interpretability of deep learning outputs poses a significant limitation for scientific validation, especially in domains such as volcanology where physical causation must be clearly understood [154], [155]. The analytical process follows a multi-step paradigm, consistent with the temporal evolution of volcanic phenomena. This includes phases such as forecasting, detection, tracking, quantification, and finally nowcasting, aligning with recent developments in geophysical monitoring. For the segmentation of thermal anomalies, supervised machine learning algorithms, like Random Forest, were implemented to enhance the discrimination between background and volcanic anomalies, even in cases where thermal variations are especially subtle. These techniques have shown superior performance compared to conventional methods, particularly due to their ability to adapt to the intrinsic variability of satellite data [156]. In the detection phase, Convolutional Neural Networks (CNNs) were employed, leveraging their layered architecture to perform spatial and spatio-temporal analyses on imagery acquired in the SWIR (Short-Wave Infrared) and NIR (Near-Infrared) bands. This approach enabled more precise identification of anomalies and accurate scene classification, confirming the effectiveness of CNNs in Earth observation applications. Special attention was devoted to the classification of the different phases of a volcanic eruption. Various machine learning techniques, such as Support Vector Machines, K-Nearest Neighbors, Decision Trees, and again Random Forest, were tested not only for classification tasks but also to identify the most relevant features for predictive modeling [157]. Further methodological advancement is represented by the integration of foundation models based on Vision Transformer (ViT) architectures [158], pre-trained on large-scale Earth Observation datasets such as BigEarthNet [159] and SEN12MS [160]. Two main applications were explored: the first involved fine-tuning for the semantic segmentation of thermal anomalies, improving quantification capabilities and overcoming known obstacles such as halos and refractive artifacts. The second application adopted a zero-shot approach, testing the foundation model's generalization capabilities for the segmentation of sulfur dioxide (SO_2) plumes without additional task-specific training [158]. Model performance was assessed using both classical metrics (e.g., accuracy, F1-score, precision, recall) and geospatially aware indicators such as Intersection over Union (IoU) and area under the ROC curve (AUC). Special emphasis was placed on temporal consistency and false positive rates, which are critical when developing operational alert systems for volcanic hazards.
Despite their promising results, the application of foundation models to EO and volcanology is not without limitations. These models are computationally intensive and often require substantial resources for fine-tuning, raising questions about their operational feasibility in low-latency, field-deployed scenarios. Moreover, their high capacity to generalize comes at the expense of reduced task-specific interpretability, a crucial aspect in scientific domains where physical causality remains essential. Nevertheless, their use opens up

new avenues for zero-shot learning, domain adaptation, and multi-task learning, offering the possibility of simultaneously addressing detection, segmentation, and classification in a unified framework. The exploration of these models within this thesis should therefore be seen as a first step toward operationalizing large-scale, pre-trained AI systems in EO for volcanic monitoring. This methodological diversification enabled the proposed approach to be effectively adapted to the varying spatial and temporal scales of volcanic phenomena. It allowed for a flexible response to common challenges in the field, such as the scarcity of reliable labels, sensor heterogeneity (optical and radar), and the need for robust spatial generalization across different geographic contexts [123].

2.2.1 Machine Learning for Satellite Image Time Series (SITS)

In the context of Earth Observation (EO), a Satellite Image Time Series (SITS) is defined as a chronologically ordered sequence of images acquired by one or more satellite sensors observing the same geographic area at regular or irregular time intervals. The analysis of such sequences enables the study and quantification of the temporal evolution of physical, biophysical, or anthropogenic characteristics of the Earth's surface, providing a fundamental tool for understanding and monitoring dynamic phenomena such as volcanic activity. SITS, typically composed of multispectral and radar observations from missions like Sentinel-1 (SAR radar) and Sentinel-2 (multispectral optical), allow continuous observation of pre-eruptive changes, eruptive activity, and post-eruptive landscape transformations.
SITS have been extensively used in various application domains, including precision agriculture, forest monitoring, climate impact assessment, urban planning, and disaster management, due to their ability to track spatial and temporal variations on a large scale [161]–[163]. However, despite their potential, satellite time series also pose significant analytical challenges. The complexity of the data, resulting from high spatial, spectral, and temporal dimensionality, is compounded by issues related to observation quality, such as frequent cloud cover in optical bands, and by the heterogeneity in temporal sampling, often influenced by external factors such as data availability or sensor malfunctions. There are several ways to construct and process a satellite image time series. In the simplest cases, a time series may consist of the bi-temporal difference between two images acquired at distinct dates, useful for detecting abrupt changes [164]. A more informative approach involves extracting, for each pixel, the value of selected spectral bands or indices (e.g., NDVI, NBR, NDWI) over time, building curves that describe the evolution of one or more variables [165], [166]. For Land Use/Land Cover (LULC) classification, the temporal profile of NDVI is frequently used to distinguish crops or vegetation types based on their seasonal behavior [167]. As data complexity and analytical needs have grown, there has been a progressive shift from descriptive methods based on fixed empirical thresholds to more advanced methodologies based on artificial intelligence, capable of capturing non-linear patterns and complex dependencies across space and time. Among these, the use of three-dimensional convolutional neural networks (3D-CNN), or recurrent models such as Long Short-Term Memory (LSTM) networks, has enabled more effective modeling of temporal and spatial dynamics in EO data, improving the ability to forecast and classify complex geophysical events [162],

[163].
In volcanology, SITS play a particularly strategic role. They are used to monitor a wide range of geophysical and geochemical parameters, including thermal anomalies observed in the SWIR (Short-Wave Infrared) and TIR (Thermal Infrared) bands, the presence and dispersion of sulfur dioxide (SO_2) detected by sensors such as Sentinel-5P, and surface deformation derived from interferometric techniques applied to radar data (InSAR), particularly from Sentinel-1. In this thesis, the vast majority of satellite-based information has been structured in the form of time series, demonstrating how temporal information is essential for a deeper understanding of complex phenomena such as volcanic processes. The dynamics that precede, accompany, and follow an eruption cannot be fully understood through static analyses alone; rather, it is the temporal evolution of thermal, spectral, and morphological signals that provides crucial insights into the progression of volcanic activity. The temporal representation of data has thus proven to be more suitable than isolated snapshots for supporting forecasting, detection, classification, tracking, and nowcasting within operational monitoring and volcanic risk assessment systems. The methodological foundations outlined in this section provide the framework for the applications presented in Chapters 3 and 4, where different AI models are applied to satellite image time series (SITS) to detect and interpret volcanic phenomena across diverse sites and eruptive contexts.

2.2.2 AI Foundation Model for Earth Observation (EO)

In recent years, AI foundation models, large-scale, pre-trained models designed to be adaptable across a wide range of tasks, have begun to reshape the landscape of Earth Observation (EO). Originating from the field of natural language processing (e.g., BERT, GPT) and later adopted in computer vision (e.g., CLIP, ViT), these models are now finding powerful applications in remote sensing and geospatial analysis [168]–[170]. Their central premise is to learn general-purpose, high-dimensional representations from massive, diverse datasets, which can then be fine-tuned or adapted for specific downstream tasks such as classification, segmentation, change detection, or anomaly identification. In the EO domain, this paradigm is being increasingly adopted due to the growing availability of petabyte-scale satellite imagery from missions such as Sentinel, Landsat, and MODIS, and the pressing need for scalable, transferable AI systems [170]. Recent examples include TerraMind, an open-access foundation model trained on over 1.2 million multispectral satellite image tiles, which has demonstrated strong generalization capabilities across a wide range of geospatial tasks [171]. TerraMind leverages a Vision Transformer (ViT) architecture combined with self-supervised pretraining objectives, such as masked autoencoding and contrastive learning, making it robust to label scarcity and adaptable across sensor types and geographic domains. Other notable examples include SatMAE [139], a masked autoencoder trained on Sentinel-2 data, and Prithvi [172], which focuses on climate-relevant tasks. These models not only outperform traditional CNN-based models in accuracy but also show significantly better transfer learning performance, enabling rapid adaptation to new, data-scarce regions without the need for extensive retraining. The emergence of such models represents a substantial step forward for volcanology, a field where timely and reliable interpretation of EO data is essential but often hindered by challenges such as cloud cover, data

imbalance, limited in-situ observations, and the spatio-temporal variability of eruptive events.
This thesis presents two applications involving the use of a foundation model for semantic segmentation known as SAM2 (Segment Anything Model 2), which operates under a prompt-based paradigm [173]. The model was evaluated in two different modes, aimed at exploring its flexibility and adaptability in the context of volcanic phenomena monitoring through satellite Earth Observation (EO) data. The first mode involves using the model in a zero-shot inference configuration, meaning that no prior training was performed on the specific downstream task. In this scenario, SAM2 was tested directly on new EO scenes to assess its ability to generalize the segmentation task to previously unseen volcanic contexts. This approach aligns with recent trends in the use of large vision models (e.g., CLIP, SAM, Segment Anything) trained on generic datasets and later transferred to geospatial domains [174]. The second mode employed the model with fine-tuning, aiming to refine its performance by optimizing the prompt in relation to the specific characteristics of the satellite imagery under analysis. This mode corresponds to a growing body of research applying task-specific adaptation to foundation models, such as the BigEarthNet-MM pretraining approach proposed by [175], which fine-tuned convolutional and transformer-based architectures for land cover classification, or TempCNN and Transformer-based time series models fine-tuned for vegetation and crop dynamics [162]. Similarly, Vision Transformers adapted to EO tasks, as shown in [158] and further validated in geospatial contexts by [176], demonstrate the effectiveness of fine-tuning large-scale vision models to enhance performance in semantic segmentation and change detection using multispectral satellite imagery.
These models rely on transfer learning to adapt their generalized pre-trained representations to specialized EO contexts, often with a limited number of annotated samples, following a few-shot or semi-supervised paradigm. Through supervised adaptation, SAM2 was trained to better recognize the elements of interest within each scene, enhancing segmentation quality compared to the zero-shot approach. Notably, prompt tuning has emerged as a lightweight and computationally efficient alternative to full fine-tuning, enabling the adjustment of model behavior by learning task-specific input embeddings without modifying the model's backbone. In both applications, SAM2 demonstrated excellent performance, both in terms of segmentation accuracy and spatial consistency, confirming the potential of prompt-based foundation models to support critical automated spatial analysis tasks in volcanic Earth observation. These findings contribute to the growing evidence that foundation models, when appropriately adapted, can significantly reduce the need for large, labeled datasets while maintaining high performance across geographically diverse and operationally challenging environments.

2.2.3 Evaluation metrics and Validation Framework

The validation of machine learning and deep learning models in Earth Observation (EO) has traditionally relied on the use of well-established benchmarks and standardized evaluation protocols that promote reproducibility and comparability across studies. In the EO domain, datasets such as BigEarthNet [175], SEN12MS [160], and segmentation and classification challenges like

DeepGlobe [177] have served as important milestones for large-scale algorithm development. More recently, within the emerging paradigm of foundation models, benchmark platforms like GEO-Bench have begun to define best practices for training and evaluating generalist models designed for multiple EO tasks [178]. These resources enable consistent performance assessment on classification, segmentation, and change detection tasks across diverse regions and input modalities. By contrast, the domain of volcanology still suffers from a lack of publicly available, standardized benchmarks, particularly for pixel-level tasks such as thermal anomaly segmentation, plume detection, or eruption phase classification. A notable exception is the Hephaestus Minicubes dataset, a global, multimodal EO archive that includes labeled deformation data from 44 active volcanoes [179]. However, such datasets remain rare and underutilized. In the absence of established benchmarks, many studies rely on U-Net and its variants [180] as a baseline architecture for semantic segmentation. U-Net continues to be a widely used reference model thanks to its encoder–decoder structure with skip connections, which preserves spatial information even when trained on limited annotated data. Its proven robustness in remote sensing applications has made it a go-to choice for tasks such as thermal hotspot detection using MODIS, ASTER, or VIIRS imagery, achieving reliable accuracy in segmenting volcanic heat signatures [61], [181]. Similarly, for volcanic deformation, although EO data from global repositories such as COMET and the Smithsonian Global Volcanism Program (GVP) are available, there is still a lack of curated, ML-ready datasets with standardized ground truth [60], which limits rigorous benchmarking across model architectures, including U-Net, transformers, and foundation models. For performance evaluation, standard metrics such as Intersection over Union (IoU), F1-score, Precision, Recall, and Overall Accuracy are commonly used.
These metrics capture complementary aspects of model performance and are essential for assessing methodological robustness. IoU and F1-score quantify spatial accuracy and balance between omission and commission errors, which are critical in pixel-level segmentation of volcanic phenomena. Precision and recall provide insight into the reliability and sensitivity of detection, respectively, while overall accuracy, although widely used, may be biased in the presence of class imbalance. Therefore, model validation in this thesis prioritizes IoU and F1-score, complemented by temporal consistency analysis to ensure the stability of predictions across image sequences. In volcanology, where false positives may lead to unnecessary alerts and false negatives could result in missed hazards, achieving the right balance between precision and recall is essential. Ground truth annotations used in this thesis were generated through a hybrid approach, combining both manual and official sources. Most image-based annotations, such as masks for thermal anomalies, lava flows, or plume extents, were created manually, through expert visual interpretation of satellite imagery, often supported by the application of threshold-based pre-detection methods (e.g., SWIR intensity) which were subsequently verified and refined. In contrast, temporal labels, such as eruption start and end dates or periods of heightened activity, were assigned by referencing official monitoring reports.
Overall, the lack of standardized benchmarks and consistent validation protocols in volcanology remains a major obstacle to methodological comparability. Many studies still report accuracy metrics that are not directly comparable due to differences in sensor type, spatial scale, or labeling criteria. Addressing

these inconsistencies is essential to ensure the reproducibility and scalability of AI models in operational monitoring environments. In the case of Mount Etna, these labels were derived from the weekly bulletins issued by the Istituto Nazionale di Geofisica e Vulcanologia (INGV), while for other volcanoes, the Smithsonian Institution's Global Volcanism Program (GVP) provided authoritative eruption records.

2.3 Reinforcement of the Observational systems of the Earth

The continuous improvement and reinforcement of Earth observational systems is a key requirement for advancing our ability to monitor, understand, and predict natural phenomena, particularly those with rapid onset and high impact such as volcanic activity. In the volcanological domain, the capacity to integrate heterogeneous sources of information, from satellite-based platforms to ground-based stations and airborne systems, is essential for building a complete, multi-scale picture of events as they unfold. A modern, reinforced observational system is inherently multi-sensor and multi-platform. Spaceborne assets in polar orbits, such as Sentinel-2, Landsat 8/9, and TROPOMI, provide high spatial resolution and detailed spectral information, while geostationary satellites like SEVIRI and FCI ensure continuous temporal coverage over fixed regions. These orbital data streams are complemented by ground-based instruments, including seismometers, GNSS networks, infrasound arrays, and gas sensors, as well as airborne platforms such as drones and research aircraft capable of targeted, high-resolution observations. Such capabilities are embedded within observational networks that extend beyond isolated instruments, forming coordinated frameworks where measurements are standardized, interoperable, and jointly analyzed. These networks, whether regional or global, facilitate the sharing of critical parameters in near real time, enabling the cross-validation of satellite observations with in situ measurements and viceversa.
Alongside these networks, experimental and computational infrastructures, including high-performance computing clusters, large-scale data repositories, modelling platforms, and simulation laboratories, provide the analytical backbone required to process, assimilate, and interpret vast volumes of observational data. They also enable the development and testing of new algorithms, such as AI-based detection systems or advanced data fusion methodologies, under controlled and reproducible conditions. Despite these advances, current systems still face limitations. Gaps in temporal coverage from high-resolution polar-orbiting satellites may delay the detection of fast-changing phenomena. As discussed before, cloud cover can obstruct optical and thermal measurements, while calibration differences across sensors may introduce inconsistencies in long-term datasets. Access to data can also be hindered by latency, licensing, or technical constraints, reducing the ability to deliver actionable information in near real time. In this context, the integration of data streams becomes a critical enhancement. Data fusion approaches, ranging from simple co-registration and cross-calibration to advanced artificial intelligence methods, enable the combination of complementary datasets to fill observational gaps, harmonize spectral and spatial characteristics, and improve the accuracy and robustness of derived products. AI-based fusion can further support

the extraction of patterns from complex, multi-source datasets, producing information layers that are both consistent over time and rapidly available for decision-making. On a broader scale, the reinforcement of Earth observation systems also relies on international collaboration. Initiatives such as the Global Earth Observation System of Systems (GEOSS) and the Group on Earth Observations (GEO) promote the sharing of data, infrastructure, and best practices across national and institutional boundaries, creating an interconnected framework where satellite, airborne, and in situ measurements are linked into a seamless global network. For volcanic monitoring in particular, such collaborations are vital to ensure that data collected in one region can be used effectively for both local hazard assessment and global-scale impact evaluation. Ultimately, strengthening observational systems is not solely about increasing the number of instruments in operation, but about maximizing the value of their combined outputs. By leveraging the unique strengths of each sensor type, integrating them within coordinated networks, supporting them with robust experimental and computational infrastructures, and ensuring open and timely access to data, the scientific community can significantly enhance its capacity to detect, track, and model volcanic phenomena. This integrated approach lays the foundation for the near real-time monitoring of volcanic activity and for robust assessments of its climatic impacts, topics which will be explored in the following sections.

2.3.1 Near real time monitoring of Volcanic Activity

Near real-time (NRT) volcanic monitoring is characterized by the need to capture and interpret signals of eruptive activity within minutes to hours of their occurrence, translating them into timely information for civil protection and scientific analysis. Unlike general observational capacity building, which aims at strengthening long-term and structural monitoring systems, NRT operations focus on streamlined workflows and immediate data delivery. In practice, this involves continuous acquisition and automated analysis of multiple indicators of volcanic unrest: thermal anomalies on the edifice or in active lava flows, sudden increases in SO_2 emissions, rapid expansion of ash clouds, or significant deformation of the volcanic surface. Each of these signals can be detected by specific datasets, thermal infrared channels for heat sources, ultraviolet spectrometry for gas plumes, multi-angle or lidar observations for ash distribution, and rapidly processed through dedicated pipelines. A key aspect of modern NRT systems is their interoperability with predictive models. Satellite-derived measurements can be ingested directly into ash dispersion models such as HYSPLIT or FALL3D, as well as lava flow simulation frameworks, to produce short-term forecasts within hours of an eruption onset. This predictive capability allows authorities to anticipate the likely trajectory of volcanic hazards and prepare targeted responses before critical thresholds are reached. The effectiveness of these systems has been demonstrated in recent events. During the 2021 eruption of Mount Nyiragongo, thermal anomaly data from VIIRS were used in near real time to track the rapid advance of lava flows toward populated areas, feeding into hazard maps that guided evacuation planning. Similarly, in the 2022 Hunga Tonga–Hunga

Ha'apai eruption, multi-sensor NRT observations were combined with dispersion models to update aviation warnings as ash plumes reached the stratosphere and began their global spread. Integration with non-satellite monitoring networks further enhances the robustness of NRT alerts [182]–[184]. Ground-based seismic arrays, GNSS stations, and infrasound detectors can confirm or clarify the nature of a satellite-detected anomaly, distinguishing, for example, between a thermal signal caused by an eruption and one caused by a wildfire or anthropogenic activity [185]. This cross-validation increases confidence in automated alerts and reduces the likelihood of false positives. What distinguishes NRT systems is not just the variety of sensors involved, but the tight coupling between acquisition, processing, and dissemination. Satellite raw data streams are ingested into automated detection algorithms, often supported by artificial intelligence to reduce false alarms and prioritize events requiring immediate attention. The outputs, hotspot alerts, gas concentration maps, ash dispersion forecasts, are distributed to operational agencies within strict latency constraints, sometimes less than one hour after acquisition. A persistent challenge in NRT volcanic monitoring is the balance between speed and certainty. Rapid alerts are essential for hazard mitigation but may carry higher uncertainty due to incomplete data or limited cross-verification. Systems must therefore be designed to manage confidence levels explicitly, updating initial assessments as more refined measurements become available. In essence, NRT monitoring transforms the reinforced observational capacity described earlier into operational intelligence, where speed, automation, and targeted communication are as critical as the observations themselves. It is this operational immediacy that enables authorities to act in synchrony with the dynamic evolution of volcanic phenomena, reducing the window of vulnerability for exposed communities.

2.3.2 Volcanic-Climate Impact Assessment

Understanding the complex interactions between volcanic activity and the Earth's climate system has become an increasingly urgent scientific challenge in the context of global environmental change. Tracer constituents of volcanic origin directly affect both the chemical composition and the optical properties of the atmosphere (The Role of Volcanic Activity in Climate and Global Changes, 2021). Large-magnitude eruptions, particularly those with a volcanic explosivity index (VEI) equal to or greater than 4, are among the most powerful natural forcings of climate, capable of injecting vast quantities of aerosols and gases into the stratosphere. These substances can significantly alter the radiative balance of the planet, often resulting in short-term surface cooling, disruption of atmospheric circulation patterns, and changes in precipitation regimes [186]. In the oxidizing atmosphere, SO_2 released from volcanoes is rapidly converted into chemically stable sulfuric acid (H_2SO_4) or sulfate aerosols. In the troposphere (8–15 km above the surface), this transformation occurs within a few days, whereas in the stratosphere (15–50 km), it can take weeks to months [187]. The resulting aerosols interact with both incoming and outgoing infrared radiation, reducing the net solar energy reaching the surface and leading primarily to cooling in the lower troposphere [1]. As discussed in Chapter 1, the magnitude of this impact depends on multiple eruption-specific parameters, including location, season, gas concentration, and magma composition [188], and must also be evaluated alongside

other forcings such as greenhouse gas emissions and anthropogenic aerosols [91]. Furthermore, climate variability itself can influence volcanic systems: for example, an eruption coinciding with an El Niño event can attenuate the expected cooling effect [189], [190]. Historical records highlight the climate-altering potential of major eruptions. The 1815 eruption of Mount Tambora (Indonesia) injected an estimated 60–80 megatons of SO_2 into the stratosphere, which dispersed globally and converted to sulfate aerosols, significantly reducing surface solar radiation [191]. The following year, known as the "Year Without a Summer," saw widespread crop failures, ecosystem disruptions, and altered precipitation patterns in Europe and North America, alongside severe societal impacts such as famine and disease outbreaks. The 1883 Krakatau eruption similarly released vast quantities of SO_2 and ash, producing striking optical effects, such as a blue Moon, due to the scattering of sunlight by fine particles. In the 20th century, the 1991 eruption of Mount Pinatubo (Philippines) injected about 20 megatons of SO_2 into the stratosphere, increasing aerosol optical depth by a factor of 10–100 over pre-eruption levels [192]. This led to a global surface temperature decrease of approximately 0.6 °C over the subsequent 15 months and marked the largest stratospheric SO_2 cloud observed since the advent of satellite monitoring in 1978 [193]. While these examples demonstrate the scale of volcanic climate forcing, they also reveal significant observational challenges. Accurately quantifying the radiative impact of volcanic aerosols requires a coordinated observational architecture capable of capturing both horizontal dispersion and vertical profiles in the stratosphere. Low Earth orbit satellites provide global mapping and detailed spectral information, while geostationary platforms track the temporal evolution of plumes. Lidar missions such as CALIPSO and the upcoming EarthCARE mission are critical for profiling aerosol layers, and hyperspectral radiometers enable the retrieval of microphysical properties. These spaceborne capabilities must be complemented by ground-based atmospheric monitoring networks, such as the Global Atmosphere Watch (GAW) and the Network for the Detection of Atmospheric Composition Change (NDACC), which provide high-frequency measurements for validation and model assimilation.

However, current capabilities are still limited by gaps in vertical sampling, inconsistent calibration between missions, and uneven geographical coverage, particularly in the Southern Hemisphere. Data heterogeneity and latency further complicate integration into climate models. Addressing these limitations requires reinforcing the observational system through harmonized calibration protocols, expanded polar and tropical monitoring, and systematic use of data fusion to merge satellite and in situ observations into coherent, model-ready datasets. Recent advances in artificial intelligence offer a pathway to accelerate this integration. AI-enhanced workflows can automate the detection and characterization of volcanic plumes, harmonize multi-source data streams, and reduce the time between observation and model assimilation. Emerging concepts such as digital twins of the volcanic–climate system promise to merge near real-time observations with high-resolution modelling, enabling scenario testing and rapid impact assessment. The inclusion of these technologies within strengthened observational networks would not only improve the quantification of volcanic radiative forcing but also help identify climatic conditions that may influence volcanic activity, paving the way for a more integrated understanding of volcano–climate feedback.

The examination of volcanic–climate interactions highlight both the complexity of the processes involved and the crucial role of reinforced observational systems in capturing them. By integrating multi-platform satellite data, in situ networks, and advanced modelling infrastructures, it becomes possible to characterize volcanic forcing with greater accuracy, account for its variability, and evaluate its interplay with other natural and anthropogenic climate drivers. Addressing existing observational gaps, whether in vertical aerosol profiling, temporal coverage, or data harmonization, will be essential to improving the reliability of impact assessments and forecasts. Strengthening this observational capacity is therefore not only a scientific priority but also a strategic requirement for global climate resilience. In the broader framework of this chapter, the insights gained here reinforce the central argument: only by enhancing and interconnecting our Earth observation systems can we ensure that rapid-response hazard monitoring and long-term climate assessment operate within a unified, robust, and future-ready framework.

Chapter 3

Volcanic Hazard Monitoring from Space: An AI-Based multi-step approach

In this chapter, and in the following one, the results achieved during the PhD are presented systematically, illustrating the methodologies developed and applied for volcanic hazard monitoring from space. The work follows the operational sequence introduced in the previous chapters, from Forecasting to Nowcasting, but the structure adopted here is not strictly linear. Instead, it reflects the interconnected nature of the different analytical steps, each contributing to a comprehensive and scalable monitoring framework. The chapter begins with a description of the data sources and volcanoes used as test sites for method development and performance evaluation. It then introduces the core methodological components (Section 3.3), which represent the foundation of the entire workflow. These include feature engineering strategies, classical and deep-learning algorithms (Random Forest, SqueezeNet, Cascading ML, and the few-shot Segment Anything Model 2), and the evaluation procedures used to assess their robustness. Each algorithm was conceived as a modular element: while designed for a specific purpose, most can operate across multiple stages of the monitoring chain, from the identification of precursors to near-real-time tracking and quantification.

To provide a synthetic view of this framework, Figure 3.1 illustrates the conceptual workflow implemented in this thesis, encompassing the materials and methods adopted, the main operational phases of volcanic monitoring (from forecasting to nowcasting), and the tools developed as final outcomes of this research. Rather than functioning as a rigid sequence, the system allows cross-interaction between modules: for instance, the same model may support both forecasting and detection tasks, depending on the input data and temporal context.

Following the methodological description, the chapter is organized according to the main operational tasks of volcanic monitoring, forecasting, detection,

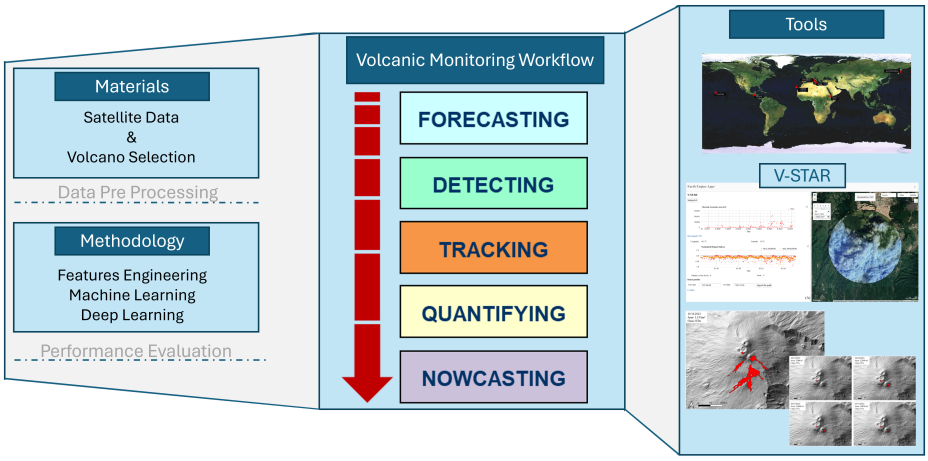

FIGURE 3.1: Conceptual overview of the AI-based workflow for volcanic hazard monitoring from space.

tracking, quantification, and nowcasting. Each section presents the implementation and application of the developed models within the corresponding phase, supported by case studies demonstrating their effectiveness in real scenarios.
Finally, Section 3.9 introduces two operational tools, V-STAR (Volcanic Satellite Thermal Anomalies Recognition) and an AI-based Platform, both conceived as practical outcomes of this research, translating the developed methodologies into accessible, near-real-time applications for hazard management.

3.1 Introduction

Near-real-time monitoring of volcanic activity is today an indispensable requirement for understanding eruptive processes and, above all, for risk mitigation. The ability to track an eruption from the onset of its earliest precursor signals through to short-term forecasting and the characterization of its different phases forms the basis for effective warning systems and timely mitigation strategies. In this context, active volcanoes pose significant challenges to direct observation: the unpredictable nature of eruptive phenomena, the high risk to operators, and the logistical complexity of maintaining ground-based networks in remote and hazardous areas make ground monitoring alone insufficient [194], [195]. Satellite observations thus emerge as an irreplaceable tool for volcano hazard monitoring, thanks to their ability to provide global coverage, high temporal frequency, and multispectral acquisitions across different regions of the electromagnetic spectrum [196], [197]. Polar-orbiting sensors, such as Sentinel-2 MSI, enable the detection of thermal anomalies of varying intensity with high spatial resolution, from persistent lava effusions to subtle degassing features, exploiting the NIR and SWIR bands, which are highly sensitive to elevated surface temperatures [11]. Geostationary sensors, such as SEVIRI aboard Meteosat satellites, instead provide continuous surveillance with a temporal resolution of a few minutes, making it possible to estimate Volcanic Radiative Power (VRP) and to identify characteristic patterns that

precede or accompany specific eruptive styles, thereby enabling nowcasting capabilities.
Over the past two decades, technological progress has made available unprecedented volumes of data from numerous multispectral, thermal, and radar sensors. While this abundance of information represents a valuable resource, it also requires innovative and scalable methods for analysis and interpretation, overcoming the limitations of traditional approaches based on fixed thresholds or manual inspection [6], [196]. Although simple and computationally efficient, these methods prove inadequate in complex conditions, for instance in the presence of clouds, snow, reflective surfaces, or weak precursor signals, resulting in false positives or false negatives [63], [97].
Recent advances in artificial intelligence (AI) offer a response to these challenges. Machine learning (ML) and Deep Learning (DL) techniques have proven capable of detecting and mapping thermal emissions, estimating lava field volumes, and characterizing explosive plumes, while improving sensitivity and specificity compared to static methods [10], [13], [14], [61], [124], [126]. These data-driven approaches, based on large archives of historical data as well as real-time acquisitions, allow adaptation to spectral and environmental variability, thus significantly enhancing monitoring capabilities. Further advancements have been introduced through ensemble and cascading architectures, which combine multiple models to increase accuracy and reduce classification errors, often implemented on cloud computing platforms capable of processing multi-petabyte datasets within timeframes compatible with operational monitoring [198]–[200]. In parallel, the rapid development of Foundation Models, large pre-trained neural networks trained on heterogeneous datasets and capable of performing multiple tasks, is opening new perspectives for volcanic monitoring. Examples such as the Segment Anything Model 2 (SAM2) demonstrate extraordinary potential in zero-shot segmentation [173], although their direct application to multispectral and thermal satellite data requires domain-specific adaptations. If properly calibrated, such models could represent a breakthrough in satellite data processing, enabling more generalist and transferable approaches. A crucial aspect in this context is the accurate identification of the scientific problem and the consequent selection of the most suitable spectral bands and analytical techniques. Lava flows, plumes, gases, intracrater anomalies, or weak precursors each require different approaches and the targeted use of specific spectral regions, as well as analytical methods tailored to capture the unique characteristics of each phenomenon [12].
All these developments converge toward the realization of a new generation of platforms for volcanic monitoring, designed as multi-task, modular, and intelligent systems, capable of integrating multiple data streams and automatically activating procedures dedicated to specific tasks: from short-term forecasting to eruptive style classification, from early anomaly detection to the nowcasting of paroxysmal events. Unlike traditional single-task platforms, these systems are able to make distinct modules interact under a common, data-driven logic, translating the massive flow of satellite data into timely, reliable, and operationally relevant knowledge. In this sense, the combined application of AI and satellite observations represents not merely a technological advancement but a true paradigm shift in the field of volcano hazard monitoring, with direct implications for risk mitigation and for advancing our understanding of volcanic processes.

To facilitate understanding, the chapter is organized into two main parts. The first part (Sections 3.2–3.3) focuses on design choices, describing the data used, the rationale behind the selection of algorithms, and the methodological framework developed for volcanic monitoring. The second part (Sections 3.4–3.9) presents the final multi-step framework, in which the developed modules are applied to the operational phases of forecasting, detecting, tracking, quantifying, and nowcasting, culminating in two practical tools: V-STAR and the AI-based Platform.

3.2 Materials

Below, the different satellite data used and volcanoes taken as case studies will be presented.

3.2.1 Satellite Data

Sentinel 2 (MSI) - The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites placed in the same sun-synchronous orbit, phased at 180° to each other. Sentinel-2A (S2A) and Sentinel-2B (S2B) were launched in 2015 and 2017, respectively. The revisit frequency of one satellite is 10 days, resulting in a global revisit frequency of 5 days. Both satellites are equipped with a MultiSpectral Instrument (MSI) with 13 bands at 10 m spatial resolution in the visible and near-infrared, at 20 m spatial resolution in the red edge and shortwave infrared part of the spectrum, and at 60 m spatial resolution in the atmospheric bands.
Within the tasks presented in the following sections, Sentinel-2 data were exploited through Google Earth Engine (GEE), a cloud platform that allows users to access and process, in near real-time, a variety of products provided by different space agencies. On GEE, Sentinel-2 MSI images are available at different product levels. In the tasks involving Sentinel-2, the Level-1C archive from 2015 was used, containing about 6,000 images available as 13 UINT16 spectral bands representing TOA reflectance scaled by 10,000 (see the Sentinel-2 User Handbook for details). The available band goes from visible to short-infrared (VIS1:B2 with 0.496 μm (S2A)/0.492 μm (S2B) wavelength, VIS2: B3 with 0.56 μm (S2A)/0.55 μm (S2B), and VIS3: B4 with 0.66 μm (S2A, S2B)), near-infrared (NIR: B8 with 0.84 μm (S2A)/0.83 μm (S2B) wavelength), and short-wave infrared (SWIR: B11, with 1.61 μm (S2A, S2B) wavelength, and 12, with 2.20 μm (S2A)/2.19 μm (S2B) wavelength).

SEVIRI (MSG) - The Spinning Enhanced Visible and Infrared Imager (SEVIRI) is the primary optical imaging instrument aboard the geostationary Meteosat Second Generation (MSG) satellite series, operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). Positioned in a geostationary orbit at an altitude of approximately 36,000 km, SEVIRI provides continuous and high-frequency observations of the Earth's atmosphere and surface, enabling near-real-time monitoring of dynamic phenomena. Although its principal mission is the observation of weather systems and climate processes over Europe, Africa, and adjacent oceanic regions, its technical specifications make it particularly well-suited for volcanic monitoring.
SEVIRI is a multispectral radiometer with 12 spectral channels spanning visible, near infrared, and infrared wavelengths, covering a broad range from

0.6 μm to 13.4 μm. This configuration allows for the simultaneous detection of meteorological, environmental, and geophysical phenomena, including high-temperature volcanic features. The instrument achieves a spatial resolution of approximately 3 km at the sub-satellite point and scans the full Earth disk every 15 minutes in all channels. For specific applications, such as rapid volcanic hazard assessment, higher temporal sampling is possible in select channels, with acquisitions every 5 minutes in the rapid-scan mode. The infrared channels are particularly valuable for detecting thermal anomalies, quantifying thermal radiance, and estimating Volcanic Radiative Power (VRP). In the tasks employing this sensor, the VRP was calculated according to the methodology described by [57].

3.2.2 Volcano selection

The algorithms presented in the following sections were trained and tested on a heterogeneous set of volcanoes, selected according to their geographical location, type, and the variety of eruptive phenomena observed. This diversity allowed the methodologies to be evaluated in different contexts, ensuring a more robust and generalizable analysis. Particular attention has been devoted to specific volcanoes that constitute the focus of some of the tasks discussed, since the application of the methodologies to these case studies yielded significant results from both an operational and interpretative perspective. Other case studies, on the other hand, represent scenarios of greater complexity: while providing valuable insights, they also open new research perspectives and lay the groundwork for future methodological and applicative developments. Below, a brief description of the volcanoes is provided, followed by the case studies that are analyzed and discussed in greater detail in the subsequent sections.

Etna, Sicily (Italy) – located on the island of Sicily in Italy, is Europe's highest and most active volcano, known for its frequent eruptions, majestic beauty and historical significance. Its eruptive episodes are mainly characterized by the emission of lava fountains, pyroclastic material and lava flows.

Stromboli, Sicily (Italy) – is located on the small island of Stromboli in the Tyrrhenian Sea, part of the Aeolian Islands in Italy. It is known for its nearly constant explosive eruptions, but lava flows do occur at times. Strombolian activity represents the main hazard for the close to the crater terrace within the Sciara del Fuoco.

Cumbre Vieja (La Palma) - is a volcanic ridge located on the island of La Palma in the Canary Islands, Spain. It is known for its potential volcanic activity and the significant impact it could have on the surrounding region due to its unstable geology and is characterized by explosive activity and lava flows.

Kilauea (Hawaii) - is the youngest volcano in Hawaii and one of the world's most active volcanoes, with frequent summit and flank lava flow eruptions interspersed with periods of long-term lava lake activity at Halemaumau Crater in the summit caldera.

Erta Ale (Ethiopia) - is a unique and active shield volcano located in the Afar Region of Ethiopia. It is renowned for its persistent lava lake and is one of the most continuously active volcanoes in the world.

Pacaya (Guatemala) - is a highly active stratovolcano located in Guatemala, Central America. It is one of the most active volcanoes in the country and is

known for its frequent eruptions and stunning lava displays. It has Strombolian activity, but occasional Plinian eruptions also occur.
Klyuchevskaya (Russia) - is part of the Kamchatka volcanic complex and the larger Pacific Ring of Fire, known for its frequent eruptions and its dynamic impact on the surrounding region. It is characterized by frequent moderate-volume explosive and effusive eruptions. In October 2020, an eruption lasting five months started with Strombolian activity, explosions, lava flows and ash plumes. A new activity phase began again in November 2022, with a new eruptive phase characterized by lava fountains and ash plumes.

3.3 Methodology

This section provides an overview of the models and approaches developed and employed to achieve the research objectives, structured around the different tasks that constitute the operational workflow of volcanic monitoring, ranging from forecasting to nowcasting.
The dataset split ratios differ slightly across models because of the specific data requirements and computational characteristics of each method. Traditional machine learning algorithms (SVM, RF, DT, KNN) use a 60/20/20 split to maintain sufficient validation data for hyperparameter tuning. For CNN-based models (e.g., SqueezeNet), a 60/10/30 split was preferred to enlarge the test set and better evaluate generalization performance. For SAM2, the 70/15/15 configuration was adopted to maximize the training pool, as transfer learning requires a larger number of input samples for stable fine-tuning.

3.3.1 Feature Engineering and Selection strategies for Volcanic Monitoring

The adoption of a feature-engineering-based approach, rather than directly applying end-to-end deep learning on raw VRP time series, was motivated by two main factors. First, the available labeled dataset was limited in size and characterized by strong class imbalance, which can hinder the convergence of deep neural networks. Second, handcrafted statistical features provide interpretable descriptors that can be physically related to eruptive dynamics. This interpretability is crucial for volcanological applications, where understanding the physical meaning behind model decisions supports scientific validation and operational trust. Let's delve into the volcanic activity on Mount Etna, focusing on three main types of eruptive behavior: Strombolian activity, lava fountains, and effusive eruptions.
Strombolian activity: This phase is characterized by the intermittent emission of gas, ash, and incandescent lava fragments from the volcanic vent. These bursts can vary considerably in both intensity and duration.
Lava fountain activity: A more energetic phenomenon in which lava is ejected forcefully from the vent, forming sustained vertical jets that can reach heights of several hundred meters.
Effusive activity: This type of eruption involves the slow and continuous outflow of lava from volcanic vents without explosive activity. The lava descends along the volcano's slopes, forming lava rivers that may extend over great distances.
Understanding these three eruptive styles is fundamental, as they constitute

the target classes for the classification and nowcasting tasks developed in this work.

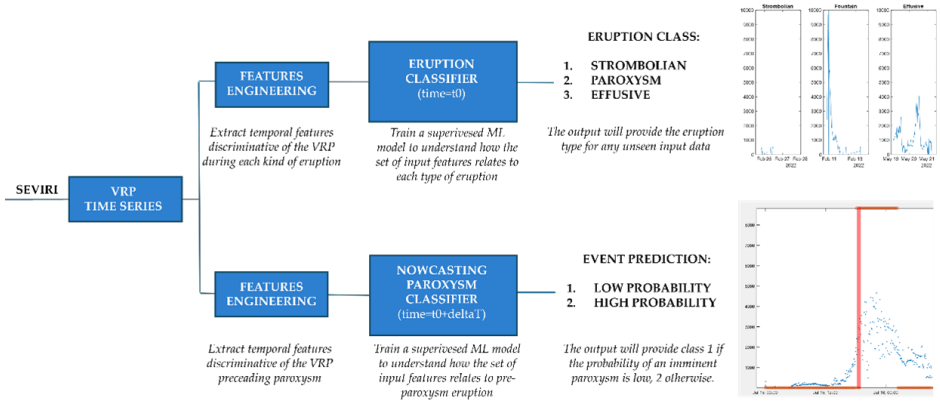

FIGURE 3.2: Workflow Feature Engineering

The workflow developed in this section builds on the analysis of VRP time series derived from SEVIRI data and is articulated into two parallel tasks, as illustrated in Figure 3.2. The first task focuses on eruption classification, aimed at identifying the ongoing eruptive style in near real time, while the second addresses paroxysm nowcasting, with the goal of predicting the likelihood of an imminent paroxysmal or lava fountain event. Both tasks rely on a feature engineering stage, in which statistical features are extracted from the VRP signal, followed by the application of supervised machine learning models to generate operational outputs for volcanic hazard monitoring.

Additional evidence from infrasound studies at Mount Etna further supports this approach. Lava fountain episodes are typically preceded by two phases detectable in infrasound data: an initial phase marked by a gradual increase in amplitude and detection numbers, followed by a violent Strombolian phase in which the signal becomes more intense and frequent as shown on Figure 3.3. This phase anticipates the transition to the lava fountain and the onset of the eruptive column [201]. Incorporating features that capture similar trends in the VRP signal is therefore essential, as they may reflect the same underlying physical processes driving eruptive transitions.

Before applying the machine learning algorithms, it was necessary to construct meaningful features from the raw data. Features are measurable characteristics extracted from the dataset and used as inputs for the algorithms. In this case, features were generated using a sliding window approach: a window of a given size scanned the data with a step of one data point, extracting the relevant statistical descriptors at each step.

The feature engineering process began with the extraction of candidate statistical descriptors from VRP time series, including maximum, mean, minimum, standard deviation, slope, product, variance, median, skewness, kurtosis, and range. These statistics were computed using sliding windows of three different sizes: 10, 30, and 50 data points, resulting in a total of 33 features (11 per window size). The choice of these three temporal windows was motivated by the need to analyze multiple types of events in the classification task, covering phenomena with different durations. The selection of multiple window sizes reflects the multi-scale nature of volcanic activity. Short windows capture

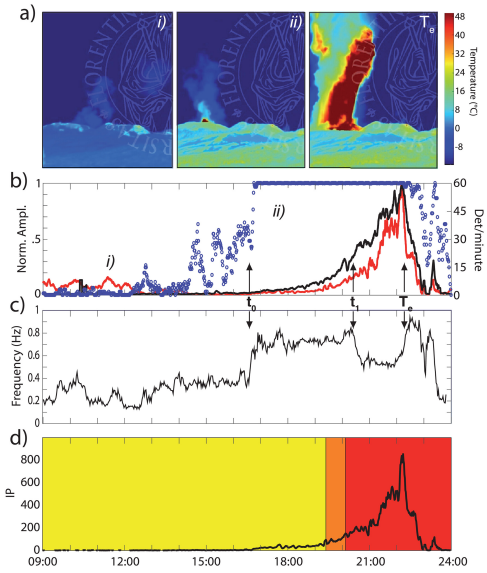

FIGURE 3.3: Picture from [201]. (a) Thermal images at Etna volcano in the absence of activity (i), during the violent strombolian activity (ii), and during the lava fountain (Te). (b) Normalized infrasound (red) and seismic (black) amplitude of the 2 December 2013 lava fountain. Number of infrasound detections per minute (blue dots). (c) Fundamental frequency of infrasound. (d) Infrasound parameter IP of the event reaches values >800 during the paroxysmal phase (Te) and shows a clear increase starting at approximately 16:30 UTC. The EW status is reported by the corresponding colors (EW0, yellow; EW1, orange; and EW2, red) and indicates that the onset of the precursory violent strombolian phase, EW1 (IP > 60), is detected at 19:24 UTC and the onset of the lava fountaining phase, EW2 (IP > 120), at 20:07 UTC.

rapid transients typical of explosive behavior, while longer ones characterize sustained effusive or degassing phases. This multi-temporal feature design ensures that the subsequent classifiers are exposed to temporal structures representative of different eruptive regimes, increasing their capacity to generalize across events of variable duration and intensity. Shorter windows (e.g., 10 samples) are suited to capturing rapid changes in activity such as paroxysms, while longer windows (30 and 50 samples) allow the characterization of more prolonged processes such as effusive eruptions.

For the classification task, it is crucial to select features that do not distort the original signal but rather preserve its inherent characteristics. This ensures that the model can correctly identify when a trend may represent an increase leading to a paroxysm, or an increase–decrease pattern that could be associated with effusive activity. In this context, certain features could be already excluded through visual inspection of the time series after applying transformations such as moving minimum (MovMin), moving median (MovMedian), moving standard deviation (MovStd), feature slope, kurtosis, and moving maximum (MovMax).

A similar rationale applies to the nowcasting task, infact to reduce the risk of overfitting, it is essential to identify the features that best discriminate between the two target classes: precursory increase and no precursory increase in VRP values. The feature selection phase is therefore the most critical step, as it allows for the extraction of descriptors that retain the most relevant

temporal trends. By preserving the signal structures that carry meaningful information for volcanic activity interpretation, the model is subsequently equipped with the most effective inputs to perform the assigned task. Four methods that can be employed for feature selection, have been identified, each providing a complementary perspective on the relative importance of the extracted descriptors: *Correlation matrix analysis*, to measure the linear dependence between features and identify highly correlated variables. *ReliefF* algorithm, which estimates feature relevance by computing distances between each sample and its nearest neighbors, updating feature weights accordingly. *Random Forest*, which measures the impact of each feature on model accuracy by evaluating changes in error after feature permutation. *Decision Tree*, which quantifies the reduction in impurity (Gini index) after splits on each feature.
For the classification task, the ReliefF method was employed to identify the most relevant features. Employing multiple feature selection techniques was

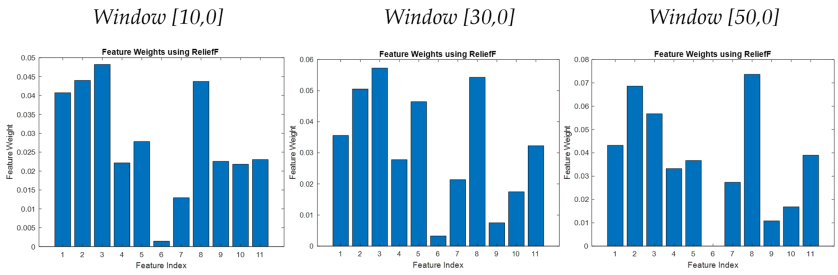

FIGURE 3.4: Feature relevance ranking obtained through the ReliefF algorithm for three different sliding window sizes applied to the VRP time series: (a) 10-point window, (b) 30-point window, and (c) 50-point window. Each bar represents the weight assigned to a candidate statistical descriptor (feature index), reflecting its relative importance in the classification task.

intended to mitigate the biases intrinsic to any single criterion. Correlation-based filtering captures linear redundancy, ReliefF identifies local neighborhood relevance, and tree-based approaches (RF and DT) estimate the global discriminative power of each descriptor through model performance. By combining these complementary perspectives, the selected features represent a balance between statistical relevance, physical interpretability, and model robustness, an essential aspect when features are later used across different tasks such as classification and nowcasting. Specifically, features were ranked according to their weights assigned by the method, and those with the highest weights were selected for each moving window size of 10, 30, and 50 data points (Figure 3.4. To avoid redundancy in feature types, the final selection included the top four features from the 10-point window, one feature from the 30-point window, and three features from the 50-point window. This choice is also motivated by the nature of the volcanic phenomena under investigation. Events such as paroxysms can start, develop, and end within a few hours; given SEVIRI's acquisition interval of 15 minutes, four data points cover approximately one hour, which is often sufficient to capture the dynamics of a paroxysmal episode. Conversely, the 30 and 50 point windows are more suitable for capturing longer-lasting processes, such as effusive activity, which can persist for weeks or even months.

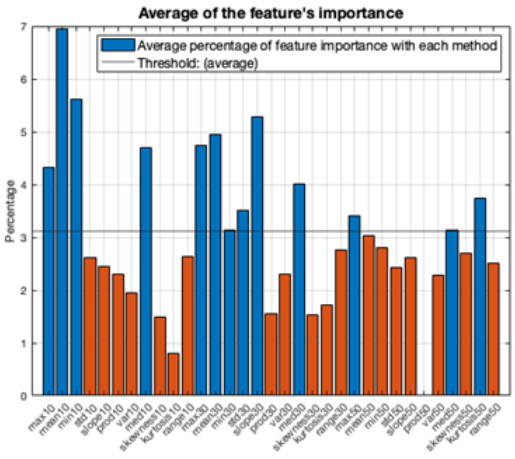

FIGURE 3.5: Average feature importance computed across multiple selection methods for the nowcasting task. The blue bars represent the mean percentage contribution of each feature (grouped by statistical descriptor and window size: 10, 30, and 50 points), while the orange bars indicate the threshold value corresponding to the overall average. Features above this threshold are considered more relevant for discriminating between the two classes (precursory increase vs. no precursory increase in VRP values).

For the nowcasting task, the feature importance scores from the four methods were averaged to obtain a final ranking for each feature. A threshold was then applied to select the most relevant ones. Three thresholds were tested: mean minus standard deviation, mean, and mean plus standard deviation. For each threshold, three selection strategies were explored: selecting features from only the window size with the best average importance, selecting for each feature the window size with the highest importance, or selecting all features that exceeded the threshold. Finally, to ensure the selected features were not overly redundant, a correlation threshold of 0.9 was applied. This combined methodology allowed for the identification of the most informative and non-redundant features, which were then used as inputs for the nowcasting machine learning models. The final subset of features selected for the classification and nowcasting tasks is summarized in Table 3.1. This table highlights which descriptors were retained for each task, reflecting the outcome of the multi-step feature selection procedure described above.

Table 3.1 summarizes the final subset of statistical descriptors retained after the feature selection process for both the classification and nowcasting tasks. Although several descriptors are shared between the two tasks (e.g., *moving minimum*, *moving maximum*, and *moving median* across different window sizes), other features differ, reflecting the distinct temporal and analytical objectives of each application.

For the classification task, the selected descriptors primarily capture short-term variability and intensity fluctuations in the Volcanic Radiative Power (VRP) signal, such as *moving standard deviation*, *moving median*, and *slope*. These features are effective in distinguishing among eruptive styles (e.g., Strombolian, lava fountain, or effusive phases), where rapid changes in radiative output correspond to different dynamical regimes of activity.

In contrast, the nowcasting task relies more heavily on descriptors sensitive to long-term trends and distributional changes, such as *moving mean*, *slope*,

Feature	Classification Task	Nowcasting Task
MovMin [10 0]	X	X
MovMin [30 0]		X
MovMin [50 0]	X	
MovMax [10 0]		X
MovMax [30 0]		X
MovMax [50 0]		X
MovMedian [10 0]	X	
MovMedian [30 0]		X
MovMedian [50 0]	X	
Skewness [10 0]		X
MovStd [10 0]	X	
MovMean [50 0]		X
Slope [30 0]		X
Slope [50 0]	X	
Kurtosis [50 0]		X

TABLE 3.1: Selected features used for the Classification and Nowcasting tasks.

and *kurtosis*. The inclusion of *kurtosis* is particularly important because it quantifies the "peakedness" of the VRP distribution within a time window, thus capturing the emergence of sharp but progressive increases that may act as precursory signatures of an impending paroxysmal event. This distinction highlights how the selected feature sets are tailored to the physical nature and predictive goals of each task: the classification features emphasize instantaneous discrimination of eruptive behavior, while the nowcasting features aim to anticipate temporal transitions in volcanic activity.

For both tasks, the same machine learning models were employed, as they were considered the most appropriate for the objectives of each analysis. These four supervised algorithms were applied to the features obtained after the feature selection phase. These features, derived from VRP time series through signal processing and statistical analysis, serve as quantitative descriptors of different states of volcanic activity. The first algorithm was the Support Vector Machine (SVM), which determines an optimal hyperplane in a high-dimensional space to separate classes with the maximum margin. In this study, a linear classifier was implemented using the Error Correcting Output Codes (ECOC) method with a One-vs-One strategy. The SVM was chosen for its strong generalization ability, even with high-dimensional datasets, and for its robustness in handling complex relationships between variables.

The Random Forest (RF) algorithm, based on an ensemble of decision trees, was also applied. Each tree is trained on a dataset generated via bootstrapping, a resampling technique in which certain samples are randomly removed and others duplicated, ensuring that all trees are trained on datasets of equal size. The outputs of all trees are then aggregated, and the most frequent classification is selected as the final prediction. This approach reduces overfitting and improves classification stability.

The Decision Tree (DT) model was also tested as a single classifier. It is a hierarchical structure in which internal nodes represent decision rules based on feature values and terminal leaves correspond to classification outcomes. In this work, the Gini index was used as the impurity measure. Decision Trees are valued for their interpretability, as they clearly highlight the features most

relevant for distinguishing between different eruptive behaviors.
Finally, the K-Nearest Neighbors (KNN) algorithm was used, which classifies a sample according to the majority class among its nearest neighbors in the feature space. In this study, a K-Dimensional Tree search method was adopted with Euclidean distance and equal distance weighting. KNN was selected for its capacity to identify local structures in the data, a characteristic particularly useful for detecting subtle differences between eruptive styles or precursory patterns.
The decision to test these algorithms stems from the need to address the variability and complexity inherent in volcanic event. SVM provides strong generalization capabilities, reducing the risk of overfitting in the presence of noisy data. DT offers high interpretability, facilitating the physical understanding of classification rules. KNN complements these methods by detecting local patterns that can be crucial for identifying transitions or anomalous events. The comparative evaluation of their performance allows for a comprehensive assessment of the system's reliability and enhances its robustness in an operational near-real-time volcanic monitoring context. The dataset used for these tasks it was divided into three subsets: 60% for training, 20% for validation, and 20% for testing.
Among the four tested algorithms, Random Forest (RF) was selected for a more detailed analysis due to its superior interpretability and overall robustness across both tasks. The other models—Support Vector Machine (SVM), Decision Tree (DT), and K-Nearest Neighbors (KNN), served primarily as benchmarks for comparative evaluation of classification performance. Section 3.3.2 provides an in-depth description of the RF model and its integration into the subsequent deep learning and cascading stages, while Section 3.3.6 and (Table 3.2) summarizes their comparative results.

3.3.2 Random Forest (RF)

Random Forest (RF) is a supervised machine learning algorithm based on an ensemble of decision trees (Figure 3.6). It belongs to the class of ensemble methods, where multiple trees are trained in parallel on randomly selected subsets of the dataset and predictor variables (bagging and random feature selection). Each tree produces an independent prediction, and the final outcome is obtained through majority voting (for classification tasks) or averaging predictions (for regression tasks). This strategy significantly reduces the risk of overfitting typical of individual decision trees, thereby improving the model's generalization capability. Moreover, Random Forest provides measures of variable importance (feature importance), which are particularly valuable for identifying the most relevant predictors in the processes under study. Thanks to its robustness, flexibility, and ability to handle complex and high-dimensional datasets, the algorithm is widely applied in remote sensing and environmental monitoring, including the detection and classification of volcanic phenomena. Random Forest has been widely used as a tool for anomaly detection, as the separation between anomalous and normal behavior is facilitated by the tree structure, which makes it easier to distinguish anomalous instances from ordinary ones.
In this thesis, the RF model was employed both for the classification task and for feature importance analysis. In volcanological applications, fixed-threshold methods are still commonly used, but they are often limiting since

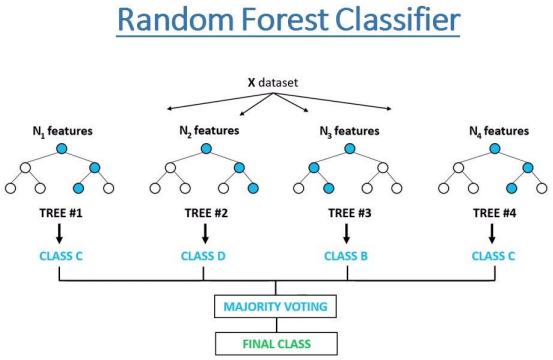

FIGURE 3.6: Conceptual scheme of the Random Forest model

they cannot adapt to the real variability of spectral values. In contrast, a supervised approach such as Random Forest replicates the "if-else" logic typical of rule-based methods, but learns it directly from real data. Thanks to its ensemble structure, RF ensures robustness and generalization capability over variable inputs, allowing the effective detection of thermal anomalies of different intensity levels, including weaker ones that fixed-threshold methods often fail to capture. The algorithm was tested using two different sets of spectral features based on [12]:

- RF1, based on NIR, SWIR, and the NHI (Normalized Hotspot Index), effective in detecting thermal anomalies due to the sensitivity of these bands to temperature variations.
- RF2, which, in addition to NIR and SWIR, also includes the visible spectrum (VIS). It was observed that adding VIS bands makes it possible to exploit the spectral signatures of heated materials even when the anomaly is less intense and the thermal signal is weak. This leads to an increase in the true positive rate and a reduction in false negatives.

As a result, RF2 was selected as the final feature set, since it improves the detection of even subtle anomalies, better capturing the contrast with the background. The choice of Random Forest over other classical machine learning algorithms (e.g., Support Vector Machines or k-Nearest Neighbors) was primarily motivated by its interpretability, computational efficiency, and ability to manage heterogeneous, non-normally distributed variables without requiring extensive parameter tuning. In the context of volcanic monitoring, where the spectral response of thermal anomalies is highly variable and often nonlinear, RF provides a robust balance between accuracy and generalization. Unlike deep learning models, which demand large annotated datasets and high computational resources, Random Forest can achieve competitive performance with limited training data and still provide physical interpretability through feature importance analysis. This makes it particularly suitable for operational applications, where transparency and explainability are essential for scientific validation and decision-making support. Random Forest proves to be a robust and generalizable tool: as an ensemble of decision trees, it is less prone to overfitting and effectively handles the spatial and spectral variability

present in satellite images. Furthermore, in this work, RF was integrated with a CNN within a cascading approach, which confirms and refines the segmentation, thereby contributing to more reliable thermal anomaly detection.
Another important advantage of RF is its ability to provide a robust and interpretable estimate of feature importance already presented in Paragraph 3.3.1. Since each tree in the ensemble splits the data based on the predictors that best reduce classification uncertainty, aggregating these decisions across hundreds of trees makes it possible to rank the relative contribution of each feature. The feature importance analysis enabled the selection of the most informative predictors, which is essential for developing a classification and nowcasting task. By focusing on the features with the highest discriminative power, the model ensures both computational efficiency and scientific interpretability, two aspects that are critical for operational monitoring systems in volcanology.

3.3.3 SqueezeNet Classifier (SN)

SqueezeNet is a deep Convolutional Neural Network (CNN) architecture commonly employed in computer vision tasks [202]. It is specifically designed to drastically reduce the number of parameters, thereby lowering memory requirements while preserving competitive accuracy. Its distinctive feature lies

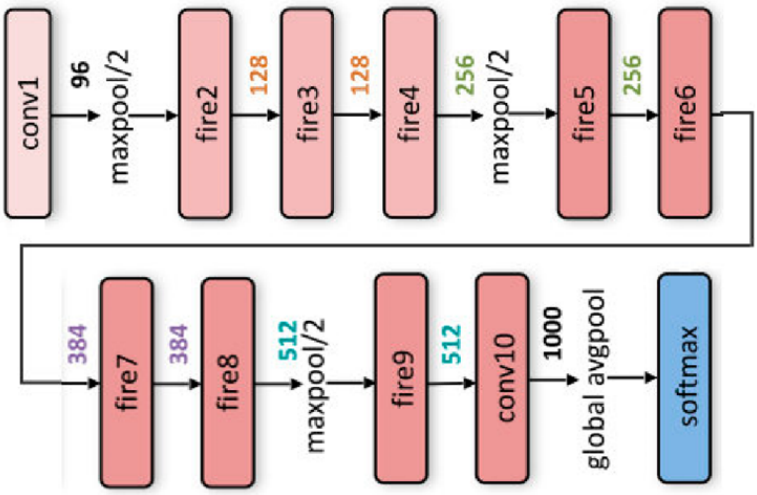

FIGURE 3.7: SqueezeNet Architecture

in the so-called Fire modules, which consist of a squeeze convolution layer with only 1×1 filters, followed by an expand layer that combines 1×1 and 3×3 convolution filters (Figure 3.7) [203]. The model was pre-trained on the ImageNet dataset [204], which contains approximately 14 million annotated images distributed across 1000 classes. Leveraging this large-scale training, SqueezeNet achieves an effective trade-off between training efficiency and classification accuracy [202], [205].
In the presented methodology, a transfer learning approach has been used to readjust the parameters of the pre-trained SqueezeNet model to the volcanic monitoring domain. Transfer learning made it possible to exploit the knowledge previously acquired by the pre-trained model in order to avoid the need for a massive, labeled dataset, thus reducing both training times

and computational costs. This is particularly advantageous in volcanological applications, where labeled data are limited and costly to obtain, while real-time responsiveness is critical. To increase robustness and generalization capability, eleven different SqueezeNet models were trained and combined into a final SqueezeNet Ensemble Classifier. Ensemble learning enhances predictive performance by combining multiple models, mitigating the variance of single runs, and producing more stable and reliable results [206]. To enhance model robustness and minimize sensitivity to initialization or data partitioning, eleven independent SqueezeNet models were trained using different random initializations and data shuffles. This approach ensures that each model learns slightly different feature representations, improving the ensemble's generalization capability when applied to heterogeneous volcanic imagery acquired under variable atmospheric and illumination conditions.
The volcanic dataset used for training was built from false-RGB Sentinel-2 images, each labeled with a volcano-specific class ID, resulting in a balanced dataset covering the 10 volcanoes investigated. The volcanic dataset was randomly split into three independent subsets: a training set (60%), a validation set (10%) and a test set (30%). The training set was generally utilized to train the model; the second one was used to validate the results in the training phase and to fit the hyperparameters. The last set was used to test the model and to calculate the predictive performances of the model. The training procedure consisted of fine-tuning the pre-trained SqueezeNet 1.0 model in Google Colab, using a learning rate $\alpha = 0.001$, a batch size of 25, and 5 epochs. During fine-tuning, all convolutional layers up to the `fire9` module were frozen to preserve the pre-trained ImageNet feature representations, while the final convolutional and classifier layers were retrained on the volcanic dataset. This configuration was selected to balance general visual transferability with task-specific adaptation, avoiding overfitting given the limited training data.
The number of epochs was chosen as a compromise between computational efficiency and performance, since further increasing the epochs did not yield substantial improvements. Each SqueezeNet model was trained independently, and the ensemble strategy combined the outputs of the 11 models: the final predicted class was determined through majority voting among the individual classifiers.
SqueezeNet proved to be particularly suitable for the objective of this research for several reasons. First, its extremely lightweight architecture drastically reduces the number of parameters to be trained while maintaining a competitive level of accuracy compared to more complex networks. This characteristic makes it ideal in contexts such as volcanic monitoring, where large volumes of satellite images need to be processed quickly, often in near real time. Another fundamental aspect is the possibility of exploiting transfer learning. Since SqueezeNet was pre-trained on a large-scale dataset such as ImageNet, it already possesses the ability to recognize general visual patterns and structures that can be effectively adapted to the volcanological domain. This makes it possible to significantly reduce both the need for labeled data, which in our case are limited and difficult to obtain, and the training time of the model. In addition, the combination of multiple SqueezeNet models into an ensemble allowed us to increase the robustness and generalization capability of the system. The ensemble strategy reduces the variability of individual networks and produces more stable and reliable predictions. This

is particularly important when working with heterogeneous and complex volcanological data, where environmental conditions and spectral characteristics can vary significantly from one volcano to another.
Finally, the efficiency and stability of the model make it especially well-suited to support tasks of continuous classification, where the ability to provide timely and updated information on the state of volcanoes is required. Thanks to its low computational cost and solid accuracy, SqueezeNet can be effectively integrated into operational monitoring pipelines, ensuring rapid responses without compromising the reliability of the results.
The choice of SqueezeNet as the core deep-learning model was driven by its architectural efficiency and adaptability to the constraints typical of operational volcanic monitoring. Unlike heavier architectures such as VGG or ResNet, SqueezeNet achieves comparable accuracy with a drastically reduced number of parameters, making it suitable for near-real-time inference on large-scale satellite datasets. Its compact design also facilitates integration within modular workflows, where it can act as a fast, scene-level classifier feeding into subsequent pixel-level models. Furthermore, the use of transfer learning from ImageNet allows the model to leverage pre-trained visual representations and adapt them to the volcanological domain, where labeled data are scarce and heterogeneous. This balance between efficiency, accuracy, and data economy justifies its adoption as the backbone of the deep-learning stage within the proposed framework. As detailed in the following section, the SqueezeNet ensemble operates as the first stage of the cascading architecture, providing a rapid and reliable assessment of volcanic activity that is later refined through the integration of the Random Forest model.

3.3.4 Cascading Machine Learning

A cascading machine learning approach was developed in this work to achieve a more robust and accurate classification of the analyzed scenes, while minimizing the occurrence of false negatives [207]. Using the outputs of the two different models (SN and RF), a more detailed and reliable description of the scene has been built. Since a segmentation model provides finer details (e.g., isolated thermal anomalies can be detected even for scenes containing clouds), a check and revise block integrating a pixel-based classifier RF has been adopted. Therefore, the cascading architecture allowed us to extract the maximum information content from the images by combining both the scene-based and pixel-based classifier schemes.
It consisted of a top-down structure, an information processing strategy also known as stepwise design and stepwise refinement.
A cascading machine learning approach was adopted to automatically detect and characterize volcanic thermal activity from Sentinel-2 MSI imagery, as well as to quantify the spatial extent of thermal anomalies. The advancement proposed here consists in exploiting the capabilities of deep learning (DL) to automatically extract spatial and spectral attributes from the images, thereby distinguishing between different types of volcanic activity recognized as classes. Specifically, isolated volcanic thermal anomalies (ITA), typically associated with intra-crater activity, new vent openings, or explosive events, are discriminated from extended volcanic thermal anomalies (ETA), which include effusive manifestations such as lava lakes, lava domes, and lava flows, non-active volcanic scenes (NVA) include volcanic areas without

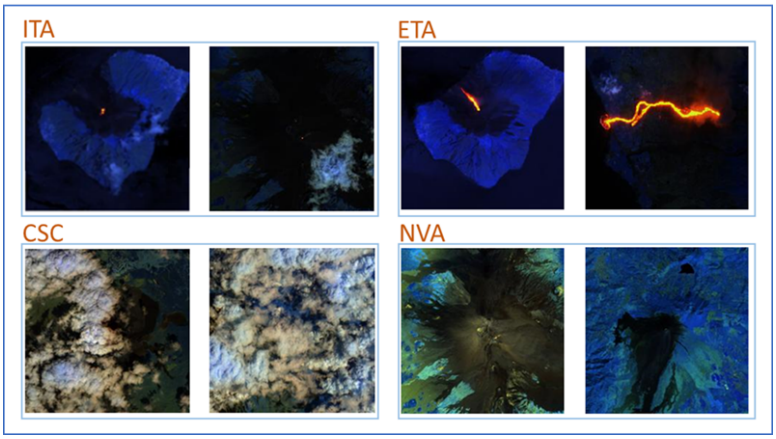

FIGURE 3.8: Sentinel 2- MSI Dataset example. Isolated Volcanic Thermal Anomalies (ITA), Extended Volcanic Thermal anomalies (ETA), No Volcanic Activity (NVA), Cloudy-Sky Condition (CSC).

thermal anomalies and clear-sky conditions and (CSC) include scenes partially/totally obscured by thin/thick clouds where the eventual presence of thermal anomalies associated with eruptive activity was hidden by clouds., as shown in Figure 3.8. This approach leverages thermal infrared satellite observations not only to detect volcanic unrest at a given time, but also to identify and classify volcanic features based on their thermal and geometric properties, effectively differentiating between ITA and ETA classes.

Input preparation - The default region of interest (ROI) for each volcano is one square [1 × 1] km centered at the volcano summit, but, if necessary, the ROI can be expanded. For each acquisition, a false RGB image is created as input for the image classifier using (NIR, SWIR1, SWIR2) TOA data, then normalized using a z-score [208] and converted into a [224,224,3].jpg format (in the 0–255 scale). A subset of 1700 images is selected from the entire collection and then split into training, validation and test dataset, i.e., 60% training set, 10% validation set and 30% test set [11]. The selection criteria were used to build a balanced target dataset with half of its scenes containing volcanic activity related to active flows or erupting vents (scenes with both isolated and extended thermal anomalies) and the other half with no volcanic activity related to a volcano at rest or presence of significant cloud coverage (scenes with no thermal anomalies). This workflow was implemented for the SqueezeNet model. In parallel, for the Random Forest classifier, the model proposed by [12] was adopted. This model was specifically designed to map thermal anomalies ranging from low to extreme intensity by exploiting TOA reflectance from the VIS1, VIS2, VIS3, NIR, SWIR1, and SWIR2 bands. The pre-trained RF classifier was directly employed to support the detection and mapping of volcanic anomalies in this study. The methodology adopted for the Random Forest is thoroughly described in the Paragraph 3.3.2. Our cascading approach involved two steps (Figure 3.9). The first used the volcanic dataset to retrain eleven SqueezeNet models by applying the transfer learning approach and combined their outcomes using an ensemble approach. The second step used a random forest model, pre-trained for detecting volcanic hot spots in Sentinel-2 MSI images [12], to check the reliability of the SqueezeNet outcomes, and quantify spatially the thermal anomalies if any. These two

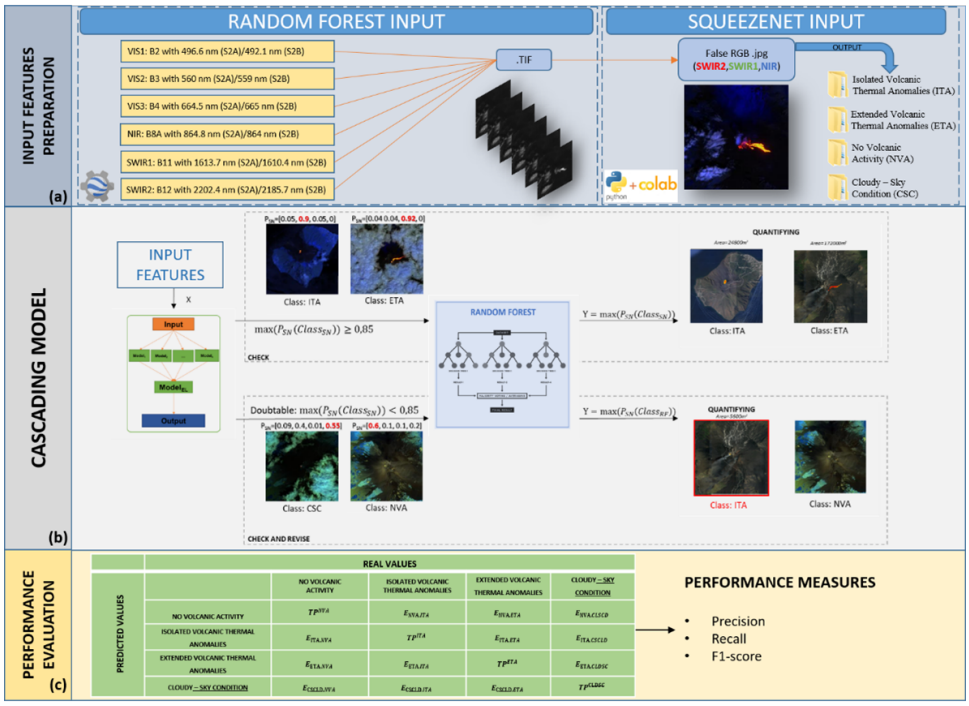

FIGURE 3.9: General scheme representing the three main steps of the cascading machine learning models: (a) Input feature preparation, (b) Cascading model and (c) Performance evaluation.

steps are described in more detail below.

It started with the general big picture and then broke down from there into smaller segments and details. SN assigned a probability score for each class, and the predicted class corresponded to the one with the highest value. It has been assumed that the prediction was reliable if it was greater than a given probability threshold (TH). Therefore, when this happened, the scene classification would be exclusively based on the SN, and the random forest model would only be used to segment the image classified as either *'ITA'* or *'ETA'* and thus quantify the detected area. Conversely, when the SN probability outcome was less than TH, the check and revise block would be activated, and the RF would be used for the scene classification.

If at least one pixel was identified as an anomaly, this means that the scene contained either isolated or extended thermal anomalies, and thus the candidate classes ($Class_{RF}$) would be *'ITA'* and *'ETA'*. On the other hand, if no pixels were detected using RF, the candidate classes would be *'CSC'* and *'NVA'*. Therefore, between the two potential $Class_{RF}$, the final prediction would be the one with the highest SN probability score, i.e., $\max(P_{SN}(Class_{RF}))$. The sensitivity threshold TH was chosen to optimize the combined use of the SN block and the check and revise block, thus guaranteeing a high final cascading accuracy. Higher values meant that the cascading model was mainly reliant on the SN and vice versa; lower values meant that it was mainly reliant on the RF. Thus, it was possible to compute the accuracy of the SN block as the accuracy of the testing subset when only the SN block was active, the accuracy of the RF-based block as the accuracy of the testing subset when

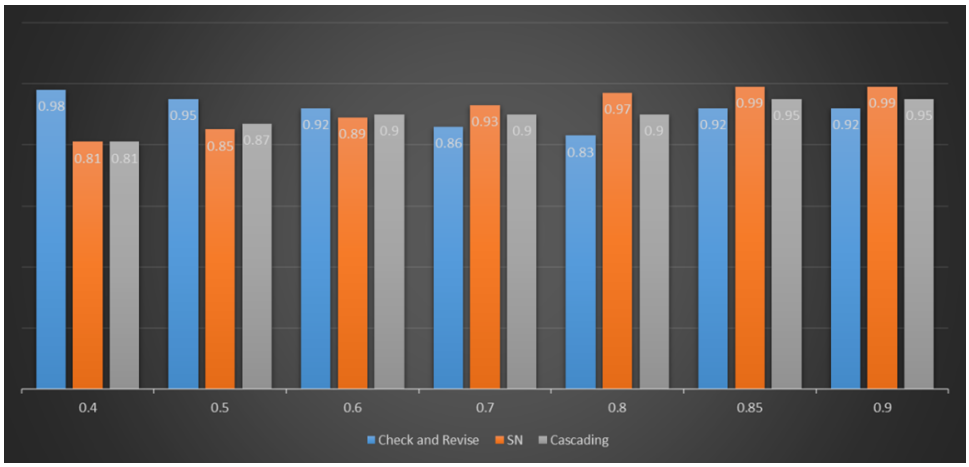

FIGURE 3.10: Values of accuracy with different Thresholds.

the 'check and revise' block was active and finally, the overall accuracy considering the overall cascading model. Figure 3.10 shows how the accuracy of each active block changed by changing the TH value. The value of TH = 0.85 was chosen because it optimized the performances of the two blocks, reaching the maximum accuracy values for both and higher cascading accuracy before a plateau started, i.e., for TH = 0.9 the metrics continued to be the same.
As shown in Figure 3.10, the stand-alone SqueezeNet model achieves the highest overall accuracy values (up to 0.99), slightly exceeding those obtained by the cascading configuration. This outcome, although apparently counterintuitive, is consistent with the design rationale of the proposed framework. The cascading model was not conceived to outperform SqueezeNet in terms of nominal accuracy, but rather to improve the overall reliability and robustness of the classification pipeline under heterogeneous observational conditions. In particular, the inclusion of the RF-based "check and revise" block mitigates the risk of false negatives when the convolutional model's confidence falls below the threshold, such as in scenes affected by partial cloud cover, low-intensity thermal anomalies, or suboptimal illumination. While this additional verification step can slightly reduce global accuracy metrics, it enhances the consistency of detection outcomes and ensures more interpretable, pixel-level confirmation of the classified volcanic activity. Therefore, the cascading strategy represents a trade-off between maximum accuracy and operational robustness, aiming to guarantee stable performance across a wide range of volcanic and atmospheric scenarios.
The cascading design represents the logical and functional bridge between the classification and segmentation components of the monitoring framework. By combining a lightweight CNN (SqueezeNet) with a tree-based ensemble (Random Forest), the model exploits both spatial-spectral pattern recognition and pixel-level statistical reasoning. This hybrid structure compensates for the respective limitations of each algorithm: SqueezeNet provides contextual scene interpretation, while Random Forest ensures local validation and quantification of anomalies. Such an approach was conceived to maximize reliability in operational conditions, where cloud cover, variable illumination, and diverse surface materials can produce ambiguous signals. The cascading structure reduces false negatives and enhances interpretability, ensuring that

the detection of volcanic thermal anomalies is both spatially precise and scientifically consistent.
It is important to note that the cascading architecture integrates only the SqueezeNet and Random Forest (RF) models, while other algorithms tested in the feature-based classification phase (SVM, DT, and KNN) were not included. This choice is motivated by both methodological and operational considerations. From a methodological perspective, RF and SqueezeNet represent complementary paradigms: the former is a pixel-based model capable of exploiting multi-spectral variability and providing interpretable feature importance measures, whereas the latter is a scene-based deep architecture that learns spatial–contextual relationships directly from imagery. Combining them therefore allows the system to leverage the strengths of both feature-driven and image-driven approaches. Conversely, SVM, DT, and KNN did not offer additional complementary advantages: SVM and DT achieved comparable accuracy to RF but with lower interpretability or higher sensitivity to hyperparameters, while KNN exhibited reduced scalability when applied to large satellite datasets. A cross-reference and comparative performance summary for all models are provided in Section 3.3.6 (Table 3.2).

3.3.5 Few-shot Segment Anything Model 2 (SAM2)

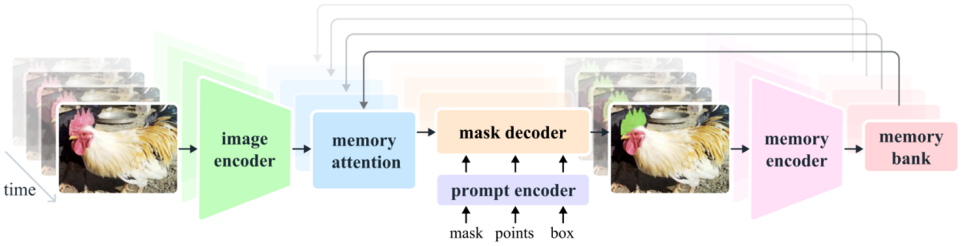

FIGURE 3.11: Segment Anything Model 2 (SAM2) Architecture

The Segment Anything Model 2 (SAM 2) represents a major advancement in the domain of foundation models for visual segmentation, extending the functionality of the original SAM to both images and videos as shown in Figure 3.11. Built on a transformer-based architecture with streaming memory, SAM 2 enables real-time processing by treating videos as continuous frame sequences. Compared to its predecessor, it achieves higher segmentation accuracy, performs static image segmentation up to six times faster, and requires approximately three times fewer user interactions to segment objects in videos.
From an architectural perspective, SAM 2 integrates four key components: Image and video encoder – Based on a hierarchical transformer (pre-trained MAE Hiera), it extracts high-level features from both static images and video frames. The feature pyramid network fuses stride-16 and stride-32 features for frame embeddings, while higher-resolution features (stride-4 and stride-8) are injected through skip connections to enhance segmentation detail.
Prompt encoder – Processes user-provided prompts (points, bounding boxes, masks) to guide segmentation, enabling the model to adapt to user input and target specific objects within a scene. Sparse prompts are embedded with positional encodings and added to the frame representation, while mask prompts

are embedded using convolutional layers.
Memory mechanism – Comprising a memory encoder, a memory bank, and a memory attention module, it stores and reuses information from previously processed frames. This allows SAM 2 to maintain temporal coherence and robust object tracking across long video sequences. Importantly, memory attention conditions the current frame embedding not only on past frames but also on "future" prompted frames when available.
Mask decoder – A lightweight, promptable module that generates final segmentation masks. It stacks two-way transformer blocks that iteratively update both prompt and frame embeddings. In the case of ambiguous prompts (e.g., a single click), multiple masks are predicted to ensure valid outputs, with the most likely (highest IoU) mask propagated through the video.
The memory attention operation plays a central role: each transformer block processes the current frame embedding via self-attention, attends to the memory bank through cross-attention, and applies an MLP layer to refine the output. This design leverages recent improvements in efficient attention kernels while ensuring scalability.
Overall, SAM 2 can be seen as a generalization of SAM to the video domain, with support for multiple prompt types on individual frames and a mask decoder capable of handling both static and dynamic segmentation tasks. Its architecture allows for robust temporal consistency, adaptability to user guidance, and fine-grained segmentation at high resolution, making it one of the most powerful and versatile models currently available for large-scale image and video understanding.

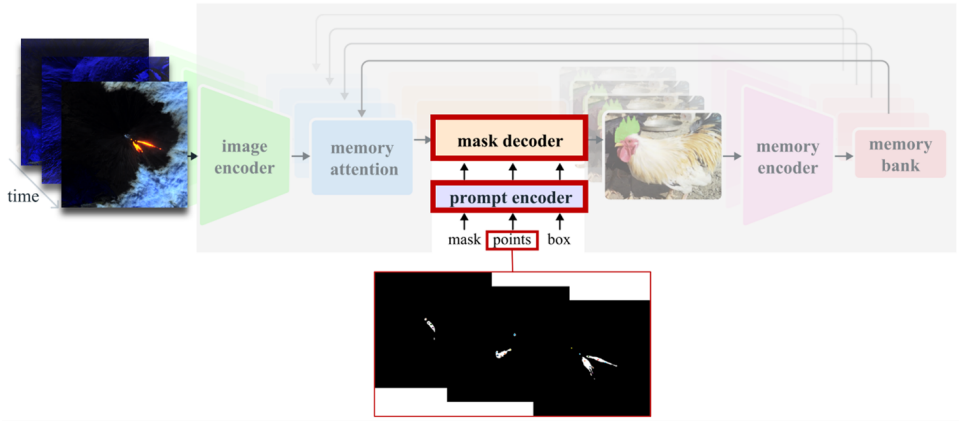

FIGURE 3.12: Workflow proposed

Here, a framework specifically designed for application on satellite imagery has been proposed. Inference is performed exclusively on two components of the model, the mask decoder and the prompt encoder, while the image encoder is kept frozen. This strategy allows the model to be effectively adapted to the domain of volcanic thermal anomalies without compromising the robustness of the visual representations already acquired during pre–training. The proposed framework is schematized in the figure 3.12.
The proposed few-shot model was trained and tested using images from several volcanoes: Mount Etna (Italy), Erta Ale (Ethiopia), and the Kilauea rift zone, specifically the Pu'u 'O'o vent (Hawaii, USA), Shiveluch (Russia)

and Klyuchevskoy (Russia). Their selection was deliberate. Mount Etna was chosen as an exemplary case study due to the abundance and variability of its thermal anomalies, ranging from small to large spatial extents. Conversely, Erta Ale and Pu'u 'O'o were included primarily for their eruptive activity in 2017 and 2016/2017. These latter events represent a greater challenge for the model, as they are characterized by diffuse activity with thermal anomalies that intermittently appear and disappear throughout the eruptive period, sometimes weak and barely detectable, and at other times markedly intensified. Finally, the inclusion of Shiveluch and Klyuchevskoy extends the analysis to two of the most active volcanoes of the Kamchatka Peninsula, both of which are frequently affected by explosive eruptions and effusive events. Their addition broadens the geographical diversity of the dataset and allows testing the model on volcanic systems with markedly different eruptive styles.
During the data preparation phase, a region of interest (ROI) was defined for each of the case studies, delineated using a rectangular geometry in Google Earth Engine (GEE). For Mount Etna, the ROI was centered on the summit area encompassing the four main craters, while for Erta Ale, Pu'u 'Ō'ō, Shiveluch and Klyuchevskoy, the regions were selected to include the full potential extent of thermal anomalies, which, as previously described, occurred at different locations along the eruptive fissures.
Once the regions were defined, a specific temporal interval corresponding to periods of volcanic activity was selected, and images from the Sentinel-2 Harmonized (COPERNICUS/S2_HARMONIZED) collection were filtered according to the ROI and dates of interest. The analysis focused on spectral bands B12 (SWIR), B11 (SWIR), and B8 (NIR), a combination commonly used to generate false-color composites that enhance thermal anomalies and lava flows, improving the distinction between hot areas and surrounding surfaces. For each image meeting the selection criteria, metadata were extracted, and the imagery was converted into NumPy arrays to allow further processing in Python.
During preprocessing, a radiometric stretching was applied between the 2nd and 98th percentiles of the pixel value distribution, reducing the influence of extreme values and enhancing spectral differences between surfaces. Importantly, the conversion was performed with a scale parameter set to 8, which ensured the preservation of the original spatial resolution while also producing images with dimensions compatible with the requirements of the model employed. The normalized values were then scaled to the 0–255 range and converted to uint8 format, making the images compatible with common visualization and processing routines.
Finally, the processed images were saved in PNG format, ready to be used as input for the model. This procedure enabled the creation of standardized and visually interpretable datasets, suitable not only for qualitative analysis but also as consistent inputs for subsequent processing, classification, and training pipelines. The images and their corresponding segmentation masks are stored in dedicated directories. The files contained in these directories are then loaded and verified, ensuring a one-to-one correspondence between each image and its associated mask. The harmonized data are organized into a CSV file, where each image identifier is paired with the respective mask identifier. The dataset is subsequently divided into three distinct subsets, following a 70%–15%–15% split for training, validation, and testing, respectively, with controlled randomization to guarantee reproducibility. Finally,

the script generates Python dictionary lists mapping each image to its segmentation mask, structuring the dataset into a format readily usable for the subsequent training, validation, and testing phases of the model.

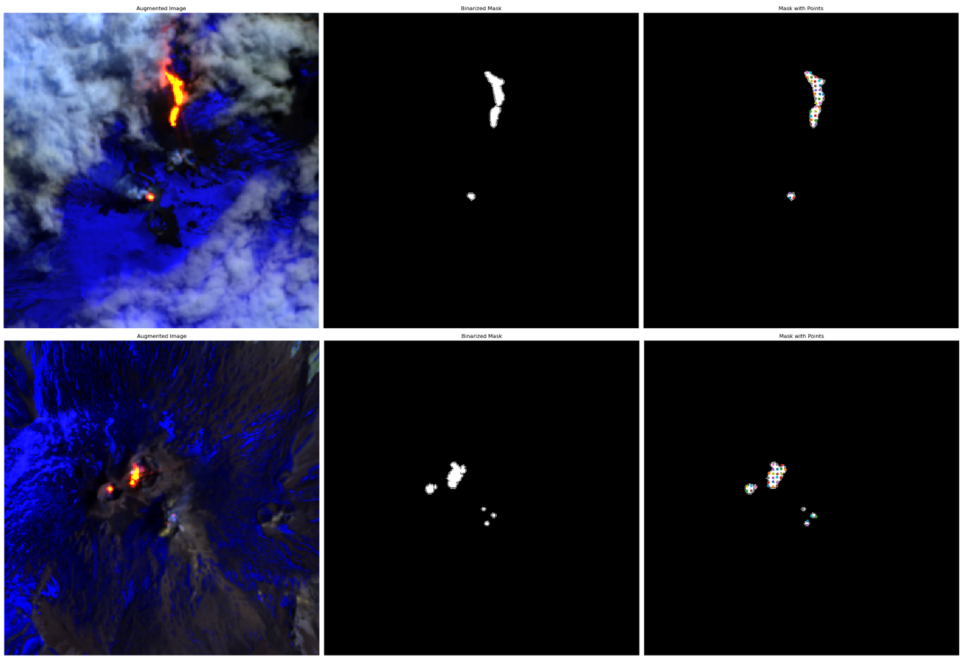

FIGURE 3.13: Points on the image

To increase dataset variability and structure the inputs consistently with the point-based approach of the model, a data augmentation pipeline combined with an innovative spatial sampling strategy has been implemented. After applying random geometric transformations (horizontal flip, vertical flip, and 90° rotation) to both the image and its corresponding annotation mask, the mask is binarized and processed to identify the classes present. It is important to recall that SAM2 is a prompt–based model: in order to operate effectively, it requires an input prompt in the form of a mask, points, or a bounding box. In this specific case, given the high variability of the analyzed images, the points has been selected as the most suitable prompt type for the problem (Figure 3.13). However, to ensure effectiveness, it is necessary to design an automatic strategy capable of placing these points within the region of interest to be segmented.

The main innovation of this approach lies in the subdivision of the mask into a regular grid of predefined size (grid size = 70). Each grid cell represents a spatial sector of the mask and serves as the domain for point sampling. Within each sector containing at least one pixel of the class, a representative point is randomly selected. This strategy ensures that sampling points are distributed more uniformly across the entire mask surface, avoiding excessive clustering in limited areas and reducing the risk of leaving entire regions unrepresented. To maintain scalability and balance the number of examples, a maximum number of points per class is fixed (max_points = 35) as shown in Figure 3.13.

A crucial aspect of this scheme is that it guarantees the presence of at least one valid point within each image. This requirement is essential in the context of using SAM 2 in few-shot point-based prompting: the presence of a labeled

point acts as a guiding prompt for the model, indicating the region on which to focus the segmentation process. Without this spatial sampling mechanism, images with small-scale or diffusely distributed anomalies would carry a high risk of lacking sufficiently informative points, thereby compromising segmentation performance.

In summary, the grid-based approach introduces a robust and systematic method for generating prompts, simultaneously improving the spatial coverage of anomalies and ensuring input consistency across the varying size and morphology of segmented regions. This results in more reliable model outputs and enhances generalization capability in scenarios characterized by complex and scattered anomalies. *Training Phase* - The model was trained with an initial learning rate of 0.0001, using a batch size of 1 image per iteration, for 1000 epochs. Training was conducted in mixed precision mode (torch.cuda.amp.autocast), combined with gradient scaling to stabilize backpropagation on GPUs and to reduce computational costs. Numerical stability was further ensured by gradient clipping (maximum norm = 1.0). Optimization was governed by a StepLR scheduler (step_size = 30, $\gamma = 0.9$), which progressively decreased the learning rate at regular intervals. Training progress was continuously monitored using Weights & Biases (WandB), while periodic model checkpoints (every 20/40/60/100/500 iterations) were saved to allow recovery and ablation analyses.

A crucial aspect of the pipeline concerns the fine-tuning strategy. During training, inference was applied exclusively to the prompt encoder and mask decoder, while the image encoder of SAM 2 was kept frozen. This methodological choice is motivated by two considerations. First, the encoder of SAM 2 has been pre-trained on large-scale multimodal datasets (millions of natural images of high variability), acquiring extremely general and robust visual representations. These features are highly transferable to the satellite image domain, where the geometry and texture of structures (lava, thermal anomalies, lava flows, background) exhibit spectral and spatial regularities that the encoder can already capture effectively. Attempting to retrain the encoder on a limited dataset of satellite images would likely result in severe overfitting and loss of generalization, in addition to prohibitive computational costs. Second, freezing the encoder directs optimization toward the lighter and more specialized stages of the model, namely the prompt encoder and mask decoder. These modules are responsible for interpreting segmentation instructions (the grid-based point prompts) and translating latent features into binary masks. In other words, while the encoder provides a general and stable representation of the image, fine-tuning only the decoders allows the model to adapt to the specific domain of volcanic anomalies under a few-shot setting, improving sensitivity and precision without the need to re-parameterize the entire architecture.

This approach is particularly suited for multispectral satellite imagery, where annotated data are scarce, as it preserves the robustness of the foundational representation while maximizing efficiency and adaptability in the task of thermal anomaly segmentation. Each training iteration begins with batch construction through the read_batch_aug function (defined in the Model Input Preparation section), which returns: (i) an augmented image, (ii) the corresponding binary ground truth mask, (iii) a set of point-based prompts, and (iv) the number of masks. Prompt points are generated via a grid-based sampling strategy (grid size = 70), ensuring spatial coverage and at least one

informative point per image, a crucial condition for employing SAM 2 in few-shot mode with point-based prompting. Points are associated with binary labels (input_labels = 1) and passed to the predictor: the image is set using predictor.set_image(...), prompts are normalized with _prep_prompts, and then encoded into sparse and dense embeddings via the SAM 2 prompt encoder. Prediction is performed using the SAM 2 mask decoder, which outputs, for each prompt, a low-resolution mask along with the corresponding IoU scores.
These masks are then upsampled to the original resolution through the predictor's internal post-processing module and compressed with a sigmoid activation to obtain pixel-wise probabilities. To mitigate bias toward excessively large segments (a frequent issue in diffuse phenomena), the pipeline introduces a ground-truth–guided dimensional filter: the maximum size of the reference object (max_size) is computed, and predicted masks whose size exceeds max_size are discarded.

$$\text{relative\_size} = \frac{\text{segment\_size}}{\text{max\_segmentsize}}$$

Remaining candidates are reweighted to favor smaller segments: for each mask, a relative size (size/max_size) is calculated, and a boost proportional to $1 - \textit{relative size}$ is added, with factor $\alpha = 0.1$.

$$\text{boost} = \alpha \cdot (1 - \text{relative\_size})$$

The final mask per prompt is chosen as the one with the highest adjusted score; a strict threshold (0.80) is applied, and if no mask meets this criterion, the iteration is skipped to avoid noisy updates. The selected masks are ranked by score and aggregated into a probability map (seg_final), representing the consensus across prompts. For operational use and metric computation, this map is binarized at a 0.1 threshold, a choice that increases sensitivity to weak and fragmented signals typical of diffuse thermal anomalies. The loss function is a sigmoid focal loss ($\alpha = 0.5$, $\gamma = 0.75$), computed between the aggregated prediction and the ground truth mask. This formulation addresses the strong class imbalance between the relatively small positive regions (anomalies) and the dominant negative background, while focusing learning on the most difficult pixels.
Model performance is assessed via Intersection over Union (IoU), computed per image, and F1-score at the 0.1 threshold. Overall, the pipeline integrates: (i) spatially aware prompting (grid-based with guaranteed coverage), (ii) size-based filtering and score recalibration to favor coherent segments, (iii) robust mask aggregation across prompts, and (iv) efficient optimization (mixed precision, gradient clipping, adaptive scheduler). This combination proves highly effective in few-shot settings, where a limited number of point prompts and annotations suffice to guide SAM 2 toward accurate segmentation, even in the presence of small, intermittent, or diffuse anomalies.
Testing and Performance evaluation – During the testing phase, the best model checkpoint, selected during training based on performance on the validation set, is loaded. Predictions generated on the test set are then processed to produce candidate segmentation masks. To obtain a coherent final segmentation from the candidate masks, a fusion algorithm based on score

ranking and overlap control was implemented. First, the predicted masks are converted into boolean format and sorted in descending order according to their associated confidence scores, ensuring that the most reliable regions are prioritized in the composition process. Two support structures are then initialized: (i) a segmentation map, which stores the final label assigned to each pixel, and (ii) an occupancy mask, which keeps track of pixels already assigned to a region. The sorted masks are then iteratively integrated into the final map. At each step, the new candidate mask is compared with the occupancy mask: if its overlap with already assigned regions exceeds 40% of its area, it is discarded to avoid duplication and conflicts. Otherwise, the non-overlapping part is added to the final map, the labels are updated, and the occupancy mask is accordingly modified. This procedure allows to: preserve masks with the highest confidence scores, reduce redundancy from multiple predictions of the same region, and ensure that the final segmentation is a coherent, non-overlapping combination of the most informative masks.

3.3.6 Performance evaluation

The evaluation of each approach was carried out in terms of correctness by computing standard statistical measures, namely True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Specifically, TP_i (True Positives) represent the predictions assigned to class i that indeed belong to that class; TN_i (True Negatives) denote the predictions assigned to a class j different from i that actually belong to class j; FP_i (False Positives) correspond to samples incorrectly predicted as belonging to class i but in reality belonging to other classes; and FN_i (False Negatives) indicate the samples incorrectly predicted as belonging to any class other than i but actually belonging to class i. The total number of images in the test set is therefore given by the sum of TP, TN, FP, and FN. These measures are conventionally organized into a confusion matrix, which provides a comprehensive description of classifier performance. The goodness of the classification for the test set was estimated using different model performance indices. The metrics used are shown in Equations (1)–(4). In particular, the F1 score was used to deal with the unbalancing issues, i.e., the different number of samples for each class. Furthermore, micro, macro and weighted F1 scores were used for the multi-class classification problem (Equations (4)–(6)). Table 35 shows the metrics

for the single SN and the cascading models.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$F1 \text{ score} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Micro F1 score} = \frac{TP + TN}{(TP + TN) + (FP + FN)} \tag{4}$$

$$\text{Macro F1 score} = \text{mean}(F1_{\text{score}_1} + F1_{\text{score}_2} + F1_{\text{score}_3} + F1_{\text{score}_4}) \tag{5}$$

$$\text{Weighted F1 score} = \frac{\sum_{i=1}^{N} W_i \cdot F1_{\text{score}_i}}{\sum_{i=1}^{N} W_i} \tag{6}$$

The following section presents all performance evaluations and the results described in the previous methodology section.

Feature Engineering and Selection Strategies for Volcanic Monitoring Classification task – The results obtained show substantial differences among the tested algorithms as shown in Table 3.2. Random Forest (RF) stands out clearly, achieving a global accuracy of 96.84% and F1-scores above 95% across all configurations. This indicates that RF is able to effectively capture the non-linear variability of the VRP signal, ensuring stable performance both in short and intense events, such as lava fountains, and in longer and less energetic ones, such as effusive lava flows. The K-Nearest Neighbors (KNN) model, while reaching good accuracy levels (90%), shows slightly reduced sensitivity for effusive events, where the gradual evolution of the signal makes recognition more complex. Even more pronounced are the limitations of the Support Vector Machine (SVM), which, with accuracy values around 75%, highlights difficulties in representing the complexity and heterogeneity of the different eruptive styles.

	Global Accuracy	**F1-Score Macro**	**F1-Score Micro**	**F1-Score Weighted**
SVM	74.91%	75.40%	74.91%	75.09%
RF	96.84%	95.69%	96.84%	96.84%
KNN	89.71%	87.80%	89.71%	89.75%

TABLE 3.2: Performance comparison of different classifiers (SVM, RF, KNN) in terms of accuracy and F1-scores (macro, micro, weighted).

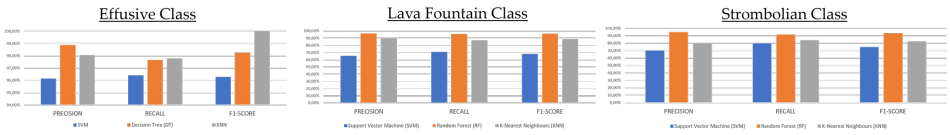

FIGURE 3.14: Performance for each class

A class-by-class analysis confirms that RF maintains high precision and recall

for all eruptive styles, with a low level of bias toward specific categories as shown in Figure 3.14. The few misclassifications detected are mainly concentrated in transitional phases between styles, where thermal signals tend to overlap and generate ambiguity. Overall, the RF model proves to be the most reliable in reconstructing the temporal evolution of volcanic activity, providing a coherent sequence of events.
These preliminary results show that classification based on VRP time series is not only technically feasible but also already provides an accuracy level suitable for operational applications. In particular, the ability to distinguish between Strombolian, paroxysmal, and effusive phases provides information of great value for volcanic risk management and for the development of nowcasting systems, where the timely detection of precursor patterns can translate into earlier warnings and more effective mitigation strategies.

Nowcasting task – The evaluation of the nowcasting framework focused on comparing the performance of three supervised machine learning algorithms, Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), trained on features derived from Volcanic Radiative Power (VRP) time series. The objective was to assess the ability of each model to detect the onset of paroxysmal activity.

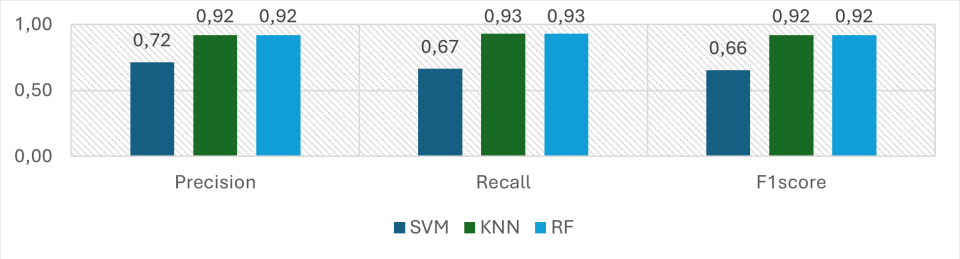

FIGURE 3.15: Performance for Nowcasting task

Across the experiments, Random Forest (RF) achieved the highest overall performance, with precision, recall, and F1-score all equal to 0.92 (Figure 3.15). This balanced performance demonstrates that RF is able to simultaneously minimize false alarms (false positives) and missed detections (false negatives). Such reliability is critical in an operational context, where both overestimating activity (triggering unnecessary alerts) and underestimating paroxysms (failing to anticipate hazardous events) can have severe implications for civil protection.

The K-Nearest Neighbors (KNN) classifier also performed well, with precision = 0.92 and recall = 0.93. These values indicate that KNN is particularly effective in capturing the onset of sudden paroxysmal episodes, where the VRP signal exhibits sharp and well-defined peaks. Figure 3.16 illustrates an example from the 10 March 2021 paroxysm, where KNN successfully associated the steep rise in VRP with a high-probability paroxysm label. This underlines its sensitivity to abrupt changes in the radiative signal. However, KNN also revealed some weaknesses: its predictions were less stable than RF, with a tendency to generate spurious detections during poorly defined precursor phases or in the presence of noisy fluctuations. This behavior reflects the model's local nature, which can overfit to short-term patterns in the signal rather than capturing broader temporal dynamics.

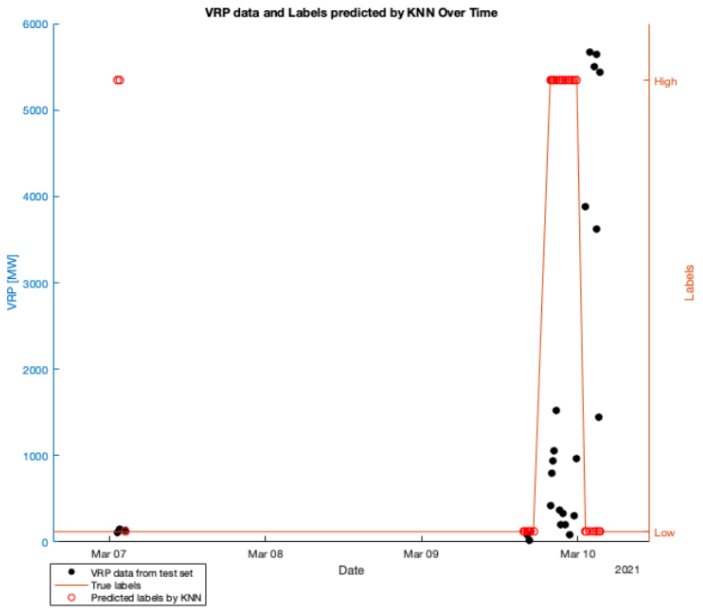

FIGURE 3.16: KNN output

By contrast, the Support Vector Machine (SVM) exhibited the lowest performance, with recall = 0.67 and F1-score = 0.66. While SVMs are powerful classifiers in many domains, their reliance on separating classes with a global decision boundary makes them less effective in handling the high variability and overlapping features of VRP signals across different eruptive styles. The model showed difficulties in consistently capturing paroxysms. The comparative analysis of the three models suggests that Random Forest provides the most reliable and generalizable solution for nowcasting tasks. KNN emerges as a valuable complementary approach, with particular sensitivity to sudden changes, which could be useful in rapidly flagging emerging paroxysms, albeit with a higher risk of false alarms. SVM, despite its weaker performance in this context, still contributes to understanding the limitations of linear and kernel-based models in handling volcanic signals. Taken together, these findings confirm that VRP-derived features, when integrated with supervised learning algorithms, represent a powerful framework for the near-real-time classification and nowcasting of eruptive activity. The ability to reliably identify different eruptive styles and to capture the onset of paroxysmal activity with operationally meaningful lead times opens the way for integrating these tools into monitoring pipelines. Beyond academic validation, the approach holds clear operational relevance: accurate and timely forecasts of eruptive activity provide actionable information for civil protection agencies, improving preparedness and enabling more effective risk mitigation during high-intensity eruptive phases.

Cascading approach - the metrics used to assess model performance are organized within a confusion matrix (Table 3.17). This matrix provides a clear overview of classification outcomes by reporting the number of correctly and incorrectly classified samples for each class. The confusion matrix that can be used to describe the performance of the classifiers (Table 3.17). The confusion matrices for both models are shown in Table 3.18. The dark green diagonal

represents correct predictions, while the other light green cells indicate incorrect predictions (E). As shown, $TP^{N}VA$ is the number of true positive NVA samples in class NVA, while $E_{NVA,ITA}$ is the number of samples from class NVA that were incorrectly classified as class ITA. The false negative total for the class NVA is the sum of $E_{ITA,NVA}$, $E_{NVA,ETA}$, $E_{NVA,CSC}$, which indicates the sum of all class N samples that were incorrectly classified as class ITA, ETA, and CSC. False positive for any predicted class which is located in a row represents the sum of all errors in that row. Moreover, in Table 3.20 shown a confusion matrix for a multi-class classification in only the SqueezeNet case. Table 3.19 shows the metrics for the single SN and the cascading models.

CASCADING	Real values				
Predicted		No Volcanic Activity	Isolated Volcanic Thermal Anomalies	Extended Volcanic Thermal Anomalies	Cloudy – Sky Condition
	No Volcanic Activity	**132** TP^{NVA}	7 $E_{NVA,ITA}$	6 $E_{NVA,ETA}$	2 $E_{NVA,CSC}$
	Isolated Volcanic Thermal Anomalies	7 $E_{ITA,NVA}$	**563** TP^{ITA}	8 $E_{ITA,ETA}$	16 $E_{ITA,CSC}$
	Extended Volcanic Thermal Anomalies	0 $E_{ETA,NVA}$	9 $E_{ETA,ITA}$	**132** TP^{ETA}	0 $E_{ETA,CSC}$
	Cloudy – Sky Condition	7 $E_{CSC,NVA}$	17 $E_{CSC,ITA}$	4 $E_{CSC,ETA}$	**584** TP^{CSC}

FIGURE 3.17: Confusion matrix for a multi-class classification in the cascading case.

	Precision		Recall		F1 - score	
	SN	*CASCADING*	*SN*	*CASCADING*	*SN*	*CASCADING*
No Volcanic Activity	0,48	0,90	0,74	0,90	0,58	0,90
Isolated Volcanic Thermal Anomalies	0,90	0,95	0,59	0,94	0,71	0,95
Extended Volcanic Thermal anomalies	0,92	0,94	0,87	0,88	0,89	0,91
Cloudy – Sky Condition	0,81	0,95	0,99	0,97	0,89	0,96

FIGURE 3.18: Confusion matrix for a multi-class classification in the SqueezeNet case.

Few-shot Segment Anything Model 2 - The predictions obtained on the test set are evaluated using quantitative metrics designed to measure both the overall segmentation accuracy and the model's ability to detect small and spatially irregular thermal anomalies. The primary evaluation metric is the Intersection over Union (IoU), computed at the image level as the ratio

	SN	Cascading
Micro F1-Score	0.79	0.95
Macro F1-Score	0.77	0.93
Weighted F1-Score	0.80	0.95

FIGURE 3.19: Performance metrics for SqueezeNet and cascading.

SN		Real values			
		No Volcanic Activity	Isolated Volcanic Thermal Anomalies	Extended Volcanic Thermal Anomalies	Cloudy – Sky Condition
Predicted	No Volcanic Activity	**108** TP^{NVA}	109 $E_{NVA,ITA}$	1 $E_{NVA,ETA}$	6 $E_{NVA,CSC}$
	Isolated Volcanic Thermal Anomalies	29 $E_{ITA,NVA}$	**353** TP^{ITA}	8 $E_{ITA,ETA}$	3 $E_{ITA,CSC}$
	Extended Volcanic Thermal Anomalies	0 $E_{ETA,NVA}$	12 $E_{ETA,ITA}$	**131** TP^{ETA}	0 $E_{ETA,CSC}$
	Cloudy – Sky Condition	9 $E_{CSC,NVA}$	122 $E_{CSC,ITA}$	10 $E_{CSC,ETA}$	**593** TP^{CSC}

FIGURE 3.20: Performance metrics for SqueezeNet and cascading.

between the intersection and the union of the predicted and reference masks:

$$IoU = \frac{A \cap B}{A \cup B}$$

where A is the predicted area and B is the ground truth. IoU is particularly suitable for assessing the correctness of spatial localization of anomalies, which are often fragmented and irregular in shape. In addition, the F1-score is calculated using a threshold of 0.1, a choice that increases sensitivity to weak or diffuse signals, which are typical of intermittent thermal activity. The F1-score provides a balanced measure of precision and recall, enabling the evaluation of the model not only in terms of its ability to localize anomalies but also in reducing false positives. Although these are preliminary results, the first analysis yields a mean IoU of 0.78 across the test dataset. This value is particularly encouraging, as it demonstrates a high level of agreement between predicted and reference masks, even in challenging conditions characterized by small, irregular, and intermittent anomalies. An IoU close to 0.8 generally indicates good spatial accuracy in segmentation tasks, suggesting that the model is able to capture the shape and extent of anomalous regions with limited loss of information. Nevertheless, it is important to stress that these results are still exploratory. The dataset used is limited in size and variability, and further testing on larger and more diverse case studies will be required to confirm the robustness and generalizability of the methodology. Despite these limitations, the preliminary findings highlight the potential of the approach, showing that the proposed fusion strategy combined with the chosen evaluation metrics can provide reliable segmentation of thermal anomalies, thereby laying the foundation for operational applications in volcanic monitoring and

early warning.

3.4 Forecasting

This section addresses forecasting and presents several case studies. The primary objective of the analysis is twofold: (i) to identify and map thermal anomalies detected in satellite imagery, and (ii) to investigate the thermal dynamics that characterize each volcanic system. The cascading model is validated on three active volcanoes, Etna, Stromboli, and Pacaya, demonstrating its ability to capture the distinctive thermal signatures of each system and to provide insights into their eruptive and pre-eruptive dynamics.

3.4.1 Data Description and Case Studies

The forecasting analysis relies on satellite-based observations from the Sentinel-2 MSI sensor, covering the period from 2017 to 2023. For each volcano, a specific Region of Interest (ROI) was defined around the active craters or summit area, and the cascading model was applied to classify thermal anomalies into four categories: Isolated Thermal Anomalies (ITA), Extended Thermal Anomalies (ETA), Non-Volcanic Activity (NVA), and Cloudy-Sky Conditions (CSC) as described in Section 3.3.4. The workflow was implemented in near real time, with a total processing time of approximately three minutes per scene, including data download, pre-processing, classification, and visualization. The study focuses on three case studies representative of different eruptive behaviors:

- **Mount Etna (Italy):** a persistently active basaltic volcano characterized by frequent effusive and explosive eruptions from multiple summit craters.
- **Stromboli (Italy):** a volcano exhibiting continuous mild explosive activity punctuated by major eruptions and paroxysms.
- **Pacaya (Guatemala):** a tropical stratovolcano combining Strombolian and effusive dynamics, often affected by high cloud cover.

3.4.2 Results and Discussion

Mount Etna

Mount Etna has exhibited persistent degassing and frequent episodes of both explosive and effusive eruptions from its four summit craters over the past two decades [209], [210]. These craters, South-East (SE), North-East (NE), Bocca Nuova (BN), and Voragine (VOR), have contributed variably to Etna's overall volcanic activity. Accurate identification of active craters at any given time and quantification of their thermal emissions are crucial for understanding the dynamics within the volcanic system. In particular, the areal coverage of thermal anomalies provides valuable insights into how each crater's activity contributes to the overall behavior of the volcano. The RF algorithm [12] has been used with the aim of evaluating the contribution of Mount Etna's four summit craters in terms of the areal extent of the thermal anomaly. To achieve this, a Region of Interest (ROI) was defined around each crater, the algorithm was applied, and the areal extent of the thermal

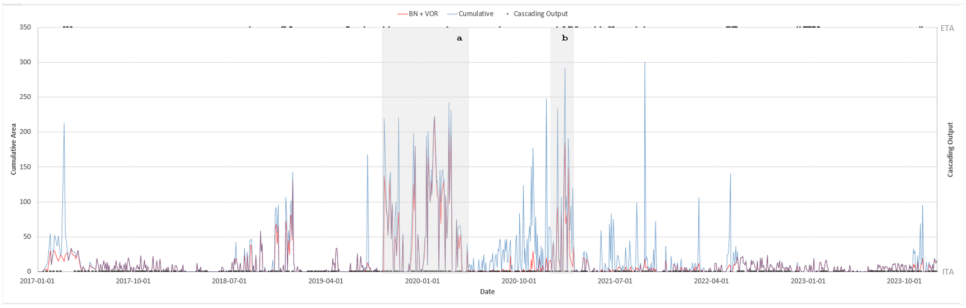

FIGURE 3.21: Thermal activity of the craters (North-East (NE), South-East (SE), Bocca Nuova (BN), and Voragine (VOR)), calculated through Random Forest Algorithm, from January 2017 to January 2023. The cumulative area of all four craters is represented by the blue line, while the contribution specifically related to BN and VOR is shown by the red line.

anomaly associated with each crater was calculated. Figure 3.21 illustrates the temporal evolution of thermal activity at Etna's craters. The blue line represents the Cumulative Area, denoting the total areal coverage of thermal anomalies from all four craters. The red line, labeled BN-VOR Cumulative Area, isolates the thermal contribution from the Bocca Nuova and Voragine craters. This differentiation is essential for understanding the interaction between craters, especially during periods of heightened activity. Figure 3.21 shows the contribution of the two craters (BN and VOR) to the observed overall thermal activity, for instance when only BN and VOR are active, the blue and red lines overlap. In addition, Figure 3.21 displays the results of our cascading model applied to satellite scenes, effectively filtering out samples classified as non-volcanic activity (NVA) or cloud/snow cover (CSC). Scenes identified as either Isolated Volcanic Thermal Anomalies (ITA) or Extended Volcanic Thermal Anomalies (ETA) are marked by gray dots, allowing us to pinpoint periods of significant volcanic thermal activity.

As illustrated in Figure 3.21a, between September 12, 2019, and April 30, 2020, Voragine displayed heightened activity, with effusive eruptions originating from a vent on the eastern side of the intracrater cone. This activity resulted in the formation of a lava flow that advanced toward Bocca Nuova, crossing the saddle between the two craters. As shown in Figure 3.21a, the combined thermal output from BN and VOR during this period contributed significantly to the BN-VOR Cumulative Area, indicating their prominent role in Etna's eruptive dynamics. Starting on December 18, 2020, Etna entered a phase of intense eruptive activity, characterized by 66 paroxysmal eruptions that produced numerous lava flows [211]. Figure 3.21b reveals that prior to the first paroxysm, all four summit craters exhibited marked increases in thermal anomaly coverage. The NE and SE craters showed escalating thermal activity as early as June 2020, while BN and VOR began displaying a similar upward trend in November 2020, signaling the imminent paroxysmal sequence. These thermal anomalies acted as early indicators of volcanic unrest, emphasizing the value of continuous thermal monitoring for forecasting eruptive events. On November 12, 2023, another significant paroxysm occurred. In the lead-up to this event, cascading model outcomes from October 19 to November 5, 2023 (Figure 3.22), revealed persistent thermal anomalies at the southeast crater starting on October 24, 2023. These anomalies were detectable until

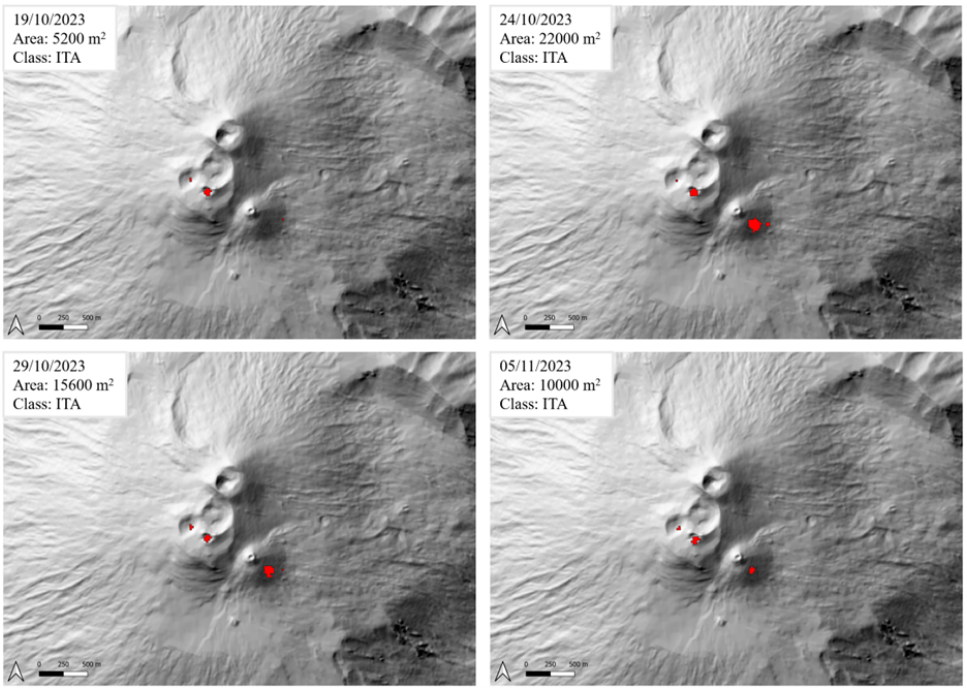

FIGURE 3.22: Volcanic thermal activity detected by the cascading model using S2-MSI data on Etna before the paroxysm on 12th November 2023.

November 5, 2023, the date of the last cloud-free satellite acquisition, just before the lava fountain of November 12, 2023. This paroxysmal event triggered several pyroclastic flows and continued with the emplacement of lava flows in multiple directions.
The cascading model's ability to accurately detect, classify, and quantify thermal anomalies on Mount Etna has proven invaluable for monitoring volcanic activity in near real-time. When referring to the concept of Near Real-Time, we specifically mean the ability to obtain data and information as soon as the Sentinel-2 satellite image becomes available. Once the image is accessible, the cascading model is promptly applied, requiring approximately 3 minutes for the download, processing, model application, and visualization of the results. By continuously tracking changes in thermal emissions across multiple craters, the model not only enhances our understanding of Etna's complex eruptive behavior but also provides critical data for improving hazard assessment and mitigation strategies.

Stromboli

Stromboli is renowned for its persistent explosive activity, which can occasionally escalate into more intense events, including major explosions or paroxysmal eruptions [209], [212]. Such episodes may lead to the effusion of lava flows, marking significant phases of volcanic unrest. One such intense phase of activity began on September 25, 2023, with a major explosive event. This powerful explosion marked the onset of heightened volcanic activity, which culminated in a new eruptive phase starting on September 27, 2023, characterized by multiple lava overflow episodes.
This eruptive activity was preceded by a notable increase in thermal emissions,

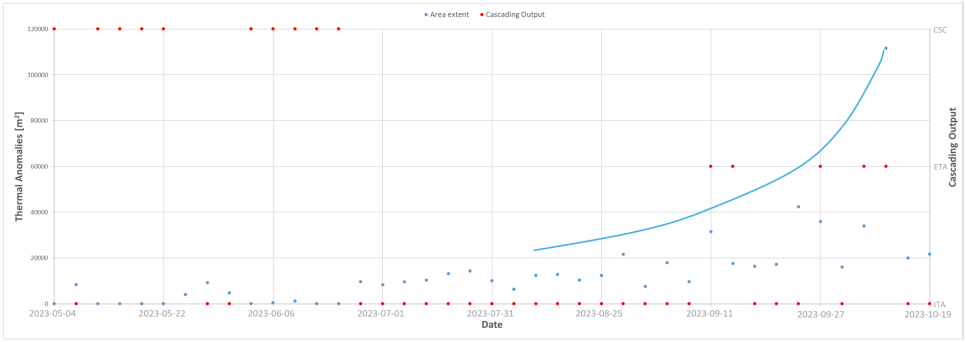

FIGURE 3.23: Volcanic thermal activity detected by the cascading model using S2-MSI data on Stromboli prior to the onset of the activity phase on September 27th, 2023. Red dots indicate the cascading model's output, the curved blue line represents the increasing trend of thermal anomalies over time, and blue dots show the thermal anomalies area extensions.

beginning in late June 2023 (Figure 3.23). This rise in thermal anomaly coverage likely reflects the presence of magma within Stromboli's shallow plumbing system, as well as the accumulation of pyroclastic material. The cascading model was instrumental in monitoring and classifying this thermal activity, allowing for a timely and accurate assessment of evolving volcanic behavior. Figure 3.24 shows the spatial and temporal evolution of thermal anomalies detected on Stromboli in the weeks preceding the major explosion of 27 September 2023. The red areas highlight hot spots identified from satellite data, whose extent varies significantly over time. On 19 October 2023, a relatively small anomaly of about $5,200m^2$ is visible, confined to the summit area. Just a few days later, on 24 October 2023, the anomaly expands markedly, reaching an area of $22,000m^2$, indicating a sudden increase in thermal activity and energy release. In the following week (29 October 2023) the anomaly decreases in intensity but remains significant ($15,600m^2$), suggesting a persistent but more fluctuating activity pattern. Finally, by 5 November 2023, the anomaly shrinks to $10,000m^2$, though still above the initial levels recorded in mid-October. This sequence highlights the dynamic evolution of thermal features, which tend to expand and contract in the days prior to and following explosive phases. The observed variations in the anomaly area are consistent with the pressurization–depressurization cycles of Stromboli's shallow feeding system, reflecting the volcano's high degree of variability and its ability to alternate between quiescent degassing and more energetic eruptive episodes. The cascading model's ability to track the evolution of Stromboli's eruptive activity, particularly in relation to lava flows and thermal emissions, provides critical insights into the behavior of this highly active volcano. By monitoring thermal anomalies in near real-time, the model aids in the early detection of potential hazards, enhancing both scientific understanding and risk mitigation efforts for volcanic activity on Stromboli.

Pacaya

Pacaya a highly active stratovolcano in Guatemala, Central America, is renowned for its frequent eruptions and impressive lava displays. While it typically exhibits Strombolian activity, Pacaya occasionally experiences more

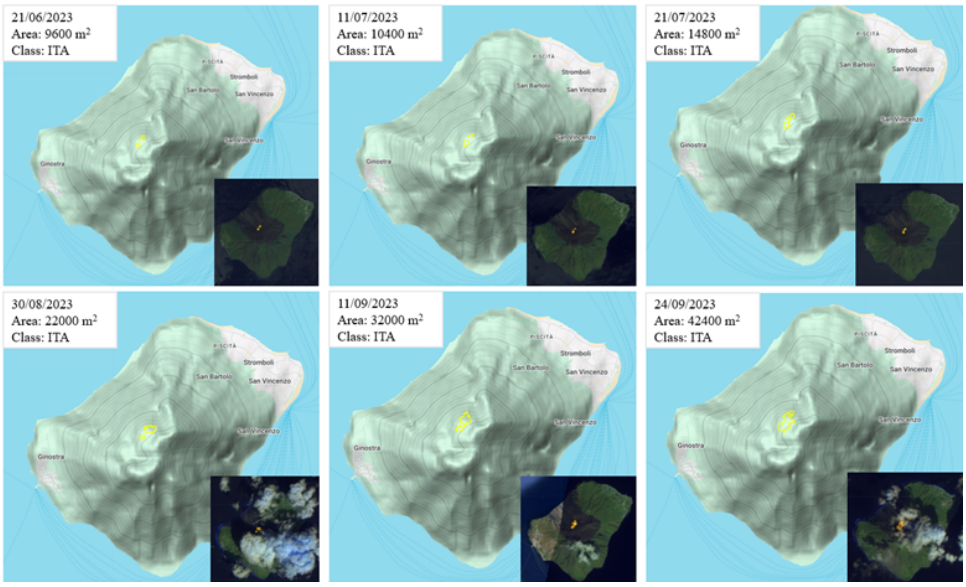

FIGURE 3.24: Increasing thermal activity on Stromboli before major explosion on September 27th 2023. Reference system used WGS84.

violent Plinian eruptions, making it a critical focus for volcanic monitoring in the region [213], [214]. The volcano entered a heightened phase of activity in early 2021, characterized by both effusive (lava flows) and explosive eruptions. This phase had been building up since mid-2015 when the volcano began showing signs of increasing unrest, marked by intensified lava effusion and ash production. The growing volcanic hazard prompted evacuations in affected areas to protect the local population [213].

The first indicators of heightened activity were detected in February 2015, with an uptick in seismicity, ash plume emissions, and increased degassing [Bulletin INSIVUMEH]. These signs of volcanic unrest persisted throughout 2016, as Pacaya exhibited intermittent Strombolian eruptions, accompanied by ongoing lava flows and ash emissions, signaling the persistence of elevated volcanic activity. The cascading model was applied to the 2017–2022 period to assess its effectiveness in monitoring and evaluating Pacaya's activity. Utilizing advanced machine learning techniques, the cascading pipeline effectively classified satellite images of Pacaya's thermal behavior. The SN model [207] within the cascading framework was particularly adapted at filtering out environmental noise, such as cloud cover and solar radiance on local slopes in the daytime images, ensuring accurate detection of volcanic activity. The RF model [Corradino et al., 2022] further reconstructed a continuous time series of thermal anomaly maps, providing a comprehensive view of the volcano's thermal dynamics.

A clear upward trend in volcanic activity is evident from August 2016 onward, coinciding with the availability of Sentinel-2 Harmonized data, which aligns well with field observations [Bulletin INSIVUMEH; Global Volcanism Program, 2017. Report on Pacaya (Guatemala) (Crafford, A.E., and Venzke, E., eds.). Bulletin of the Global Volcanism Network, 42:5. Smithsonian Institution (Figure 3.25). In February 2017, increased activity led to the formation of lava flows extending approximately 300 meters along Pacaya's northwest flank. The cascading model identifies approximately 55% of the

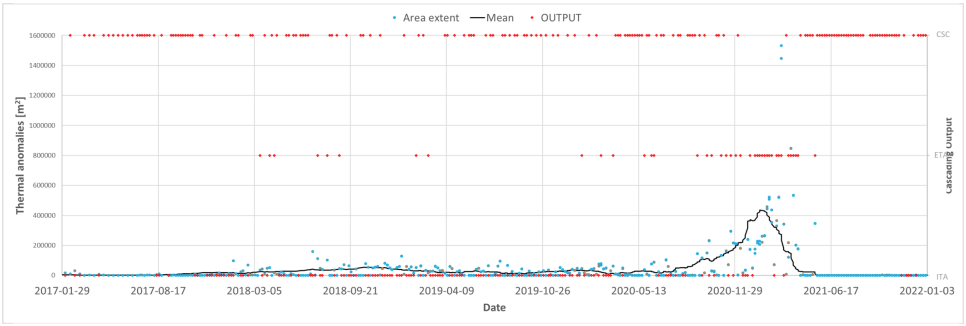

FIGURE 3.25: Temporal analysis of Pacaya's volcanic activity from August 2016 to the present, utilizing the cascading model for thermal anomaly classification. The orange line represents the mean thermal anomaly over time, highlighting the overall increasing trend in volcanic activity. The red output indicates the classifications derived from the cascading model: Isolated Volcanic Thermal Anomalies (ITA), Extended Volcanic Thermal Anomalies (ETA), and Cloudy-Sky Condition (CSC). Non-volcanic activity (NVA) has been filtered out to enhance visualization, allowing for a clearer interpretation of Pacaya's eruptive behavior and thermal dynamics throughout this period.

entire dataset as cloudy scenes, thus the system successfully filtered out much of this environmental interference. However, due to the frequent and dense cloud cover, thermal anomalies may sometimes be obscured. From 2018 to 2020, Pacaya's volcanic activity remained sustained, with multiple lava flows, persistent Strombolian eruptions, and continuous gas emissions. Figure 3.26 presents a series of Sentinel-2 Harmonized images alongside the anomaly maps generated by the cascading algorithm, depicting these key phases of activity. Notably, there was a brief pause in effusive activity during July and August 2020, followed by a renewed and intensified eruptive phase in early 2021. The early months of 2021 marked some of the most intense volcanic activity observed at Pacaya. Frequent explosions and significant lava emissions resulted in lava flows that extended into surrounding areas, necessitating road closures and evacuations to ensure public safety. Authorities issued warnings to nearby communities, highlighting the growing volcanic hazard [Bulletin INSIVUMEH].

Figure 3.27 presents a sequence of Sentinel-2 images highlighting significant increases in activity as identified by the cascading model. Some images exhibit disturbing effects (i.e., clouds coverage, diffraction spikes, blurring and thermal halo [215]. These optical phenomena occur due to the interaction of light with atmospheric particles or the sensor's optical components, causing light to scatter and produce halo effects or refract, leading to image distortions. Such effects can artificially enlarge or intensify the appearance of thermal anomalies, potentially misrepresenting the actual extent of volcanic activity. In [215] discusses how halo effects and cloud refraction can affect thermal detection and outlines methods to effectively isolate genuine thermal anomalies from such interferences. The effusive activity reached its peak in March 2021, with three lava flows, each approximately 1 km in length, advancing along multiple branching paths on different flanks of the volcano. This heightened activity is corroborated by MIROVA radiative power data [63], indicating a substantial increase in intensity through mid-April. Another spike in activity occurred in late April and early May, characterized by additional explosions

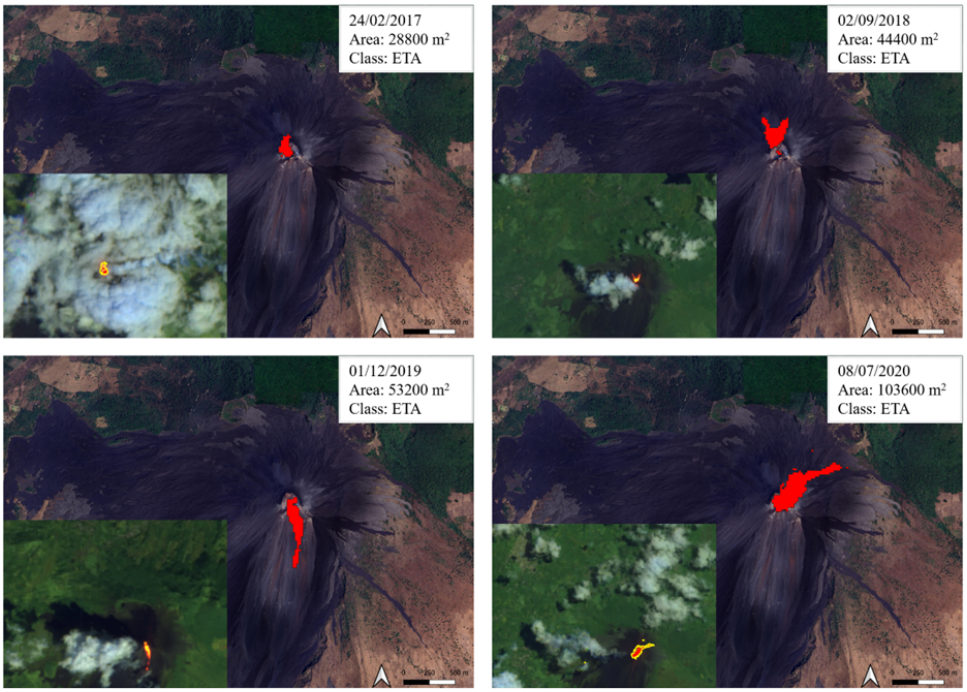

FIGURE 3.26: Lava flow episodes occur on February 24, 2017, September 2, 2018, December 1, 2019, and July 8, 2020. Reference system used WGS84.

and lava flows. However, by the third week of May, thermal activity began to decline significantly, suggesting a shift in Pacaya's eruptive behavior.
This work directly addresses some of the future challenges highlighted by [63], particularly regarding cloud filtering with the MIROVA system. The cascading model we propose incorporates a robust cloud classification step that effectively isolates cloudy scenes, identifying approximately 55% of the dataset as cloud-covered for Pacaya volcano. This is especially relevant considering the complex challenges posed by tropical volcanoes, which often experience frequent and dense cloud cover. Our model not only detects cloud-affected scenes, regardless of cloud density, but also provides a probabilistic assessment for each prediction. This feature enables the investigation of potential thermal anomalies even in images classified as cloudy.
In fact, even when anomalies are partially obscured by clouds, our model maintains the ability to identify and quantify them. Additionally, leveraging high spatial resolution data from the MSI sensor on Sentinel-2 enhances our capability to distinguish genuine thermal signals from artifacts caused by cloud presence or other atmospheric factors. This advanced approach ensures a more reliable assessment of volcanic activity, particularly in environments where traditional monitoring methods might struggle due to adverse atmospheric conditions. The integration of cloud filtering and anomaly detection within our workflow represents a significant step forward in addressing the operational challenges of monitoring tropical volcanoes like Pacaya. Overall, the forecasting results demonstrate that the cascading model effectively identifies and quantifies thermal anomalies preceding major eruptive episodes across diverse volcanic systems. The ability to detect subtle variations in radiative output and to filter cloud-covered scenes ensures a robust operational framework for near-real-time volcanic hazard forecasting.

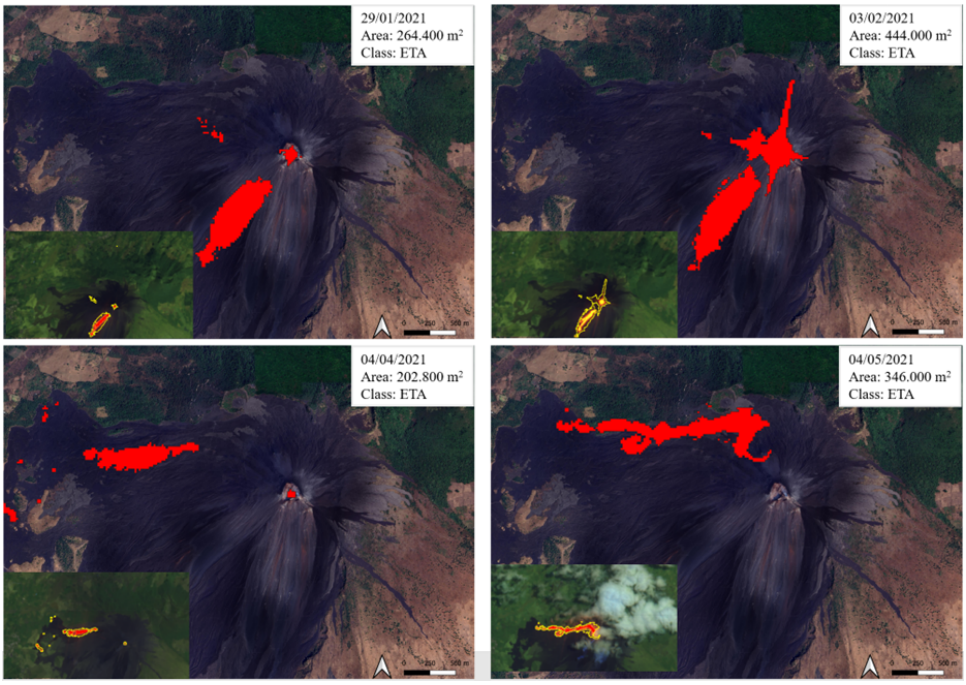

FIGURE 3.27: A sequence of images corresponding to the identified points of increase in the Figure 3.25. This series visually illustrates the key moments of heightened volcanic activity, capturing the dynamics of the eruptions and the development of thermal anomalies over time. Reference system used WGS84.

3.5 Detecting

This section focuses on the detection of eruptive phenomena. Three complementary approaches are presented: (i) the application of the cascading model to Kīlauea volcano; (ii) a novel few-shot segmentation strategy based on the Segment Anything Model 2 (SAM2) applied to Erta Ale; and (iii) the temporal classification of eruptive styles using Volcanic Radiative Power (VRP) time series derived from SEVIRI data.

3.5.1 Data Description and Case Studies

The detection framework relies optical satellite datasets acquired from multiple sensors. Sentinel-2 MSI imagery was used to detect and map active thermal anomalies at high spatial resolution, while SEVIRI data provided continuous temporal monitoring of volcanic radiative output. Each case study was selected to represent a distinct eruptive behavior and data environment:

- **Kīlauea (Hawaii):** a persistently active basaltic shield volcano characterized by alternating phases of effusive and quiescent activity. Sentinel-2 MSI data were analyzed using the cascading model to detect transitions between rest and active phases.
- **Erta Ale (Ethiopia):** a lava lake system and one of the few volcanoes with persistent thermal activity, used to test the SAM2 model for few-shot segmentation of sparse and low-intensity anomalies.

- **VRP-based detection:** SEVIRI time series from multiple volcanoes were analyzed to classify eruptive styles (Strombolian, lava fountain, effusive) based on temporal patterns of Volcanic Radiative Power.

3.5.2 Results and Discussion

Cascading Model Applied to Kīlauea

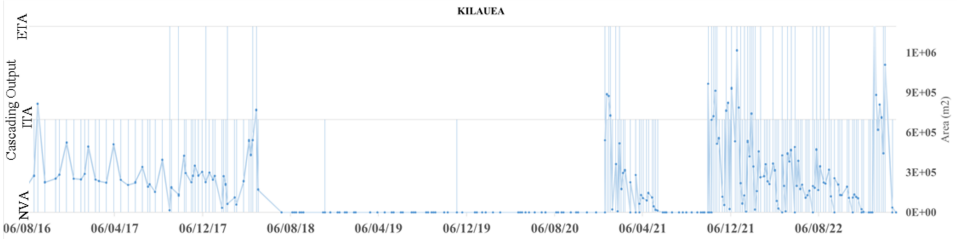

FIGURE 3.28: Temporal analysis of Kilauea's volcanic activity.

The cascading model could reveal volcanic behaviors such as the change from the rest phase to the active phase, and from active to rest to active again, as in Kilauea (Figure 3.28).
Kilauea erupted almost continuously from 1983 to 2018, and again from 2020 onward. Figure 3.28 shows 'NVA' during this resting phase, except for two false positives detected on 9 November of the years 2018 and 2019. Interestingly, these detections occurred at the same time of the year, and they were related to human activities. Moreover, from January to May 2021, it was possible to appreciate the transition from lava lake (i.e., $877,200m^2$) to intra-crater activity (i.e., $47,200m^2$). The almost persistent activity of varying intensities of Kilauea from October 2021 to May 2022 continued through November 2022, although thermal intensities diminished after mid-June 2022. Across all performance metrics, the cascading model classified each class with an accuracy greater than 0.90. The lowest performance was observed in the 'NVA' class, mainly due to a slight underestimation tendency of the cascading model, as indicated by a higher precision relative to the F1 score. Conversely, the 'ITA' class showed a slight overestimation tendency, with recall values marginally higher than the F1 score. This behavior is likely attributable to false positive pixel detections introduced when the check-and-revise block was active. Although the cascading model generally performed very well, occasional errors occurred due to false positives detected by the RF, since even a single misclassified pixel could alter the SN outcome. Nonetheless, given the high accuracy of RF (>0.90 [13]), such cases remained rare. The SN model proved effective in classifying 'ETA' and 'CSC' scenes, both yielding F1 scores above 0.87, but tended to underestimate 'ITA' scenes characterized by small anomalies. In these cases, precision exceeded the F1 score, as some scenes were misclassified as either 'NVA' or 'CSC'. For Kīlauea, this was particularly evident in the analysis of a lava lake, which was correctly recognized as an 'ETA' scene and subsequently segmented to quantify the spatial extent of the detected anomaly (Figure 3.29g–h). This validation highlights the ability of the cascading approach to capture and characterize thermal activity at

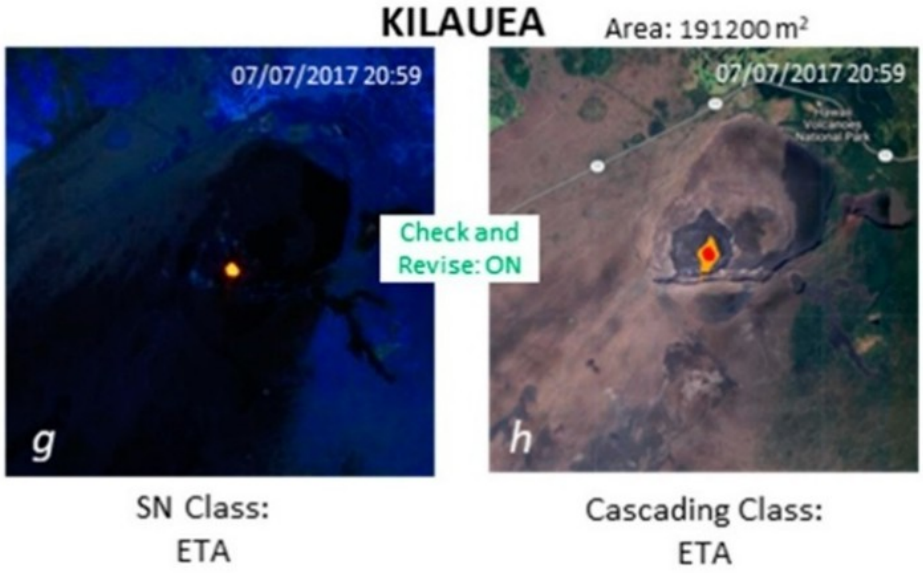

FIGURE 3.29: Cascading output for Kilauea's volcanic activity. g) show the Sentinel 2 image, h) show the segmented thermal anomaly.

Kīlauea with a high degree of reliability, offering a valuable tool for monitoring eruptive dynamics.

Few-shot SAM2 Detection at Erta Ale

The second detection approach was developed through the implementation of a novel framework based on SAM2, specifically adapted for the segmentation of satellite imagery. This adaptation was necessary because the model, originally designed and trained on natural RGB images, exhibits limitations when applied to multispectral and thermal data, which are characterized by different spectral distributions and often more complex signals. By integrating dedicated pre-processing procedures and calibrating the model to the specific domain of satellite observations, it has been possible to exploit the potential of SAM2 in this context as well, enabling its use for the automatic identification of thermal anomalies and eruptive phenomena. Here, the training technique adopted in this work, specifically based on the use of the maximum size of the summed segments, represents a crucial element in the few-shot learning framework. This strategy is particularly important for guiding the training process in scenarios where small anomalies are present within the image. Here, "small" refers both to their limited spatial extent and to their low thermal intensity.
In such cases, starting from a point-prompt located inside the anomaly, the algorithm may otherwise erroneously include disproportionately large portions of the image, leading to the identification of an incorrect and excessively broad area (see Fig. 3.30). This situation typically arises when anomalies are spatially sparse across the surface and exhibit weak intensity, as observed in the cases of Erta Ale.

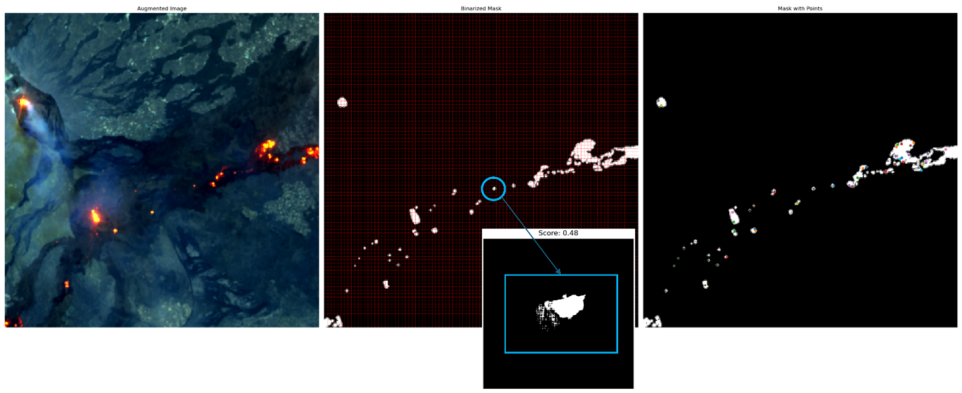

FIGURE 3.30: Original image from Sentinel 2, associated with its mask and points prompts.

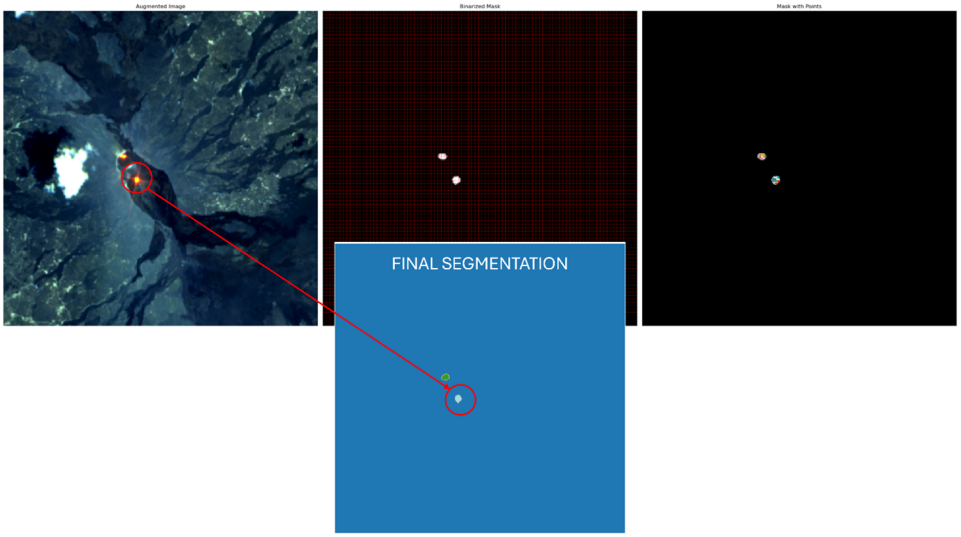

FIGURE 3.31: Halons and refraction errors with SAM2

Another common issue encountered when working with satellite imagery, particularly in the optical domain, is the presence of artifacts caused by signal refraction, excessive reflection, or atmospheric interference. These disturbances often appear as continuous lines or streaks across the area to be segmented, potentially leading the algorithm to misinterpret them as real thermal anomalies. The few-shot SAM2 approach effectively overcomes these issues: during training, the model learns to recognize such artifacts as systematic errors rather than actual regions of interest as shown in Figure 3.31. This provides a major advantage for volcanic monitoring, as it prevents overestimation of the areal extent of anomalies. The resulting segmentation is therefore more accurate, free from atmospheric or instrumental errors, and closer to the true physical phenomenon.
The model also demonstrates remarkable robustness in the presence of cloud cover, which represents one of the major challenges in the analysis of satellite imagery. Clouds can not only partially obscure thermal anomalies, reducing their visibility, but also mislead the segmentation process by generating false

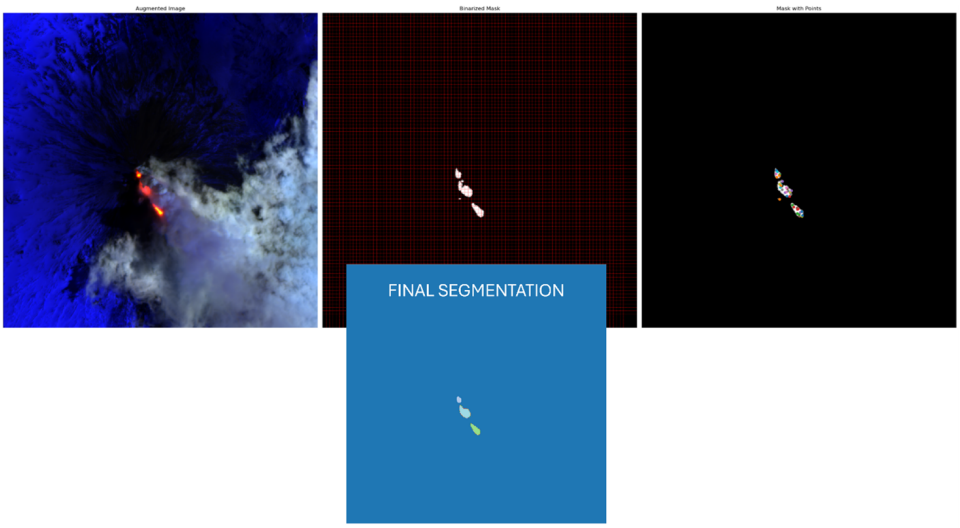

FIGURE 3.32: SAM2 in presense of cloud cover

positives. To mitigate this issue, several training cases explicitly included scenes with cloud coverage or partial obstructions. This strategy enabled the model to learn how to discriminate between true anomalies and perturbations introduced by clouds, thereby improving the reliability of the predictions and reducing the likelihood of misclassification as shown in Figure 3.32.

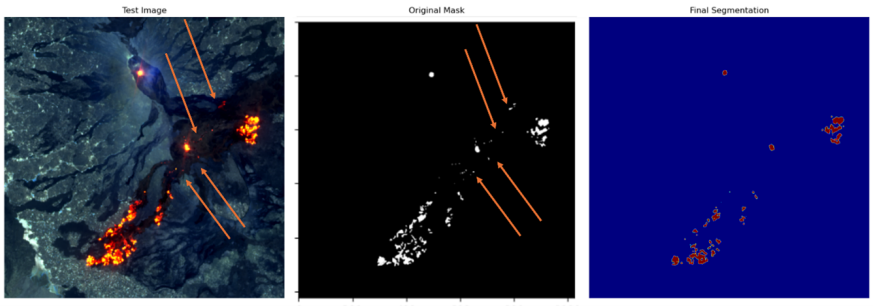

FIGURE 3.33: SAM2 in presense of sparse thermal anomalies

The model demonstrates strong segmentation capabilities, effectively avoiding many of the common issues typically encountered in this type of task, as previously discussed. However, the segmentation of small anomalies remains a more challenging aspect. Despite the introduction of corrective strategies, such as boosting and dimensional control of masks, these anomalies are still harder to detect accurately due to both their limited spatial extent and the weak intensity of their signal as shown in Figure 3.33. This limitation is particularly evident in cases such as the Erta Ale (Ethiopia) and Pu‘u ‘Ō‘ō (Hawaii) eruptions, where thermal anomalies are often sparse, fragmented, and intermittently weak. In such scenarios, the model may still struggle to fully capture the true extent of the activity, highlighting the intrinsic difficulty of segmenting diffuse and small-scale anomalies in satellite imagery.
A possible improvement could be achieved by expanding the training dataset, specifically by incorporating a larger number of examples of this typology.

Increasing the representation of sparse and fragmented anomalies would provide the model with more opportunities to learn the spatial and spectral characteristics of such features, potentially enhancing its ability to detect them consistently during inference. In conclusion, it is important to emphasize that the results presented here are still preliminary, as the algorithm is currently under testing and validation. Nevertheless, the performance achieved so far is promising: the model reaches an average Intersection over Union (IoU) of approximately 0.78, already demonstrating a solid segmentation capability in complex scenarios characterized by volcanic anomalies of varying intensity and spatial distribution.
While these results are preliminary, they open up several interesting directions for future developments. A natural extension of the framework could be its application to imagery from Landsat 8/9. Unlike Sentinel-2, these sensors provide access to the Thermal Infrared (TIR) band, which would allow the segmentation process to incorporate information not only about high-temperature anomalies but also about areas undergoing cooling. If SAM2 proves capable of generalizing effectively to this type of data, it would enable a much more comprehensive and multi-level representation of volcanic activity, with significant potential impact for both monitoring applications and early warning systems.

VRP-based Detection and Classification

The third detection approach is the analysis of Volcanic Radiative Power (VRP). The analysis of Volcanic Radiative Power (VRP) time series represents a fundamental step in linking satellite observations with the understanding of different eruptive styles. Unlike purely spatial methodologies, which focus on the detection and segmentation of thermal anomalies, the temporal approach captures the dynamic evolution of volcanic activity, translating variations in the thermal signal into interpretable eruptive categories. In this way, it is possible to move from a simple detection of anomalies to a functional classification of events, distinguishing between Strombolian activity, lava fountains, and effusive phases. This approach is particularly relevant from an operational monitoring perspective, as it not only allows for a retrospective description of eruption characteristics, but also provides input for predictive models and nowcasting systems, based on the recognition of precursor patterns in VRP data. The figure 3.34 present representative examples of the application of the Volcanic Radiative Power (VRP) time series classification model, where eruptive sequences are divided among the three styles considered: Strombolian, lava fountain, and effusive activity. In the Figure 3.34a, two distinct paroxysmal episodes characterized by sharp VRP peaks can be observed. The model successfully identifies the lava fountain phases, highlighted in red, with good consistency with the actual signal evolution. The classification produces a well-structured temporal sequence, distinguishing periods of intense activity from phases of lower emission (Strombolian or background). Minor discrepancies occur mainly during transitional phases, where the signal shows rapid and ambiguous variations that may lead the model to slightly anticipate or extend the classification compared to reality.
The figure 3.34b illustrates a case of prolonged effusive activity, characterized by sustained high VRP values and numerous internal fluctuations due to

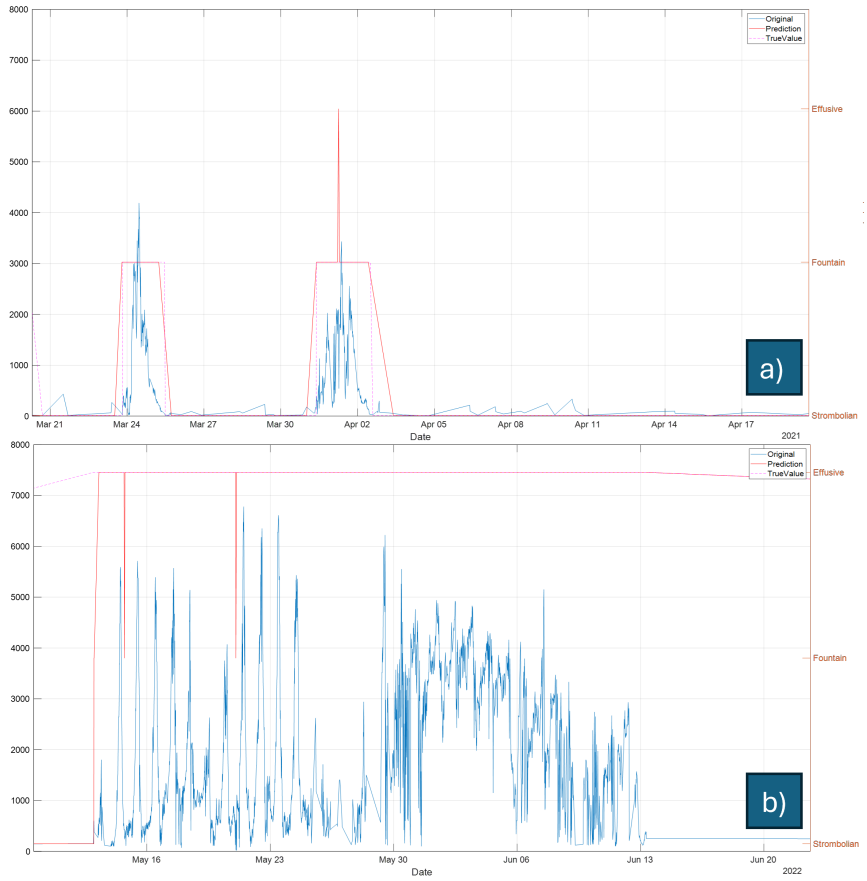

FIGURE 3.34: Output of Random Forest, for the classification Task

degassing episodes and minor explosions. The model is able to capture the continuous nature of the event, correctly classifying it as an extended effusive phase. However, the presence of irregular and overlapping signals sometimes leads to small oscillations in the assigned labels, reflecting the intrinsic complexity of this type of eruption compared to more clear-cut paroxysmal events. Overall, the results show that the model can faithfully reproduce the temporal dynamics of different eruptive styles, recognizing both short-lived explosive episodes and extended effusive phases. The main uncertainties are concentrated in the transitional periods between eruptive styles, where the radiative signal displays mixed characteristics that are more difficult to interpret. These observations confirm the potential of the approach for operational applications, while also highlighting the need for further optimization to improve the handling of intermediate phases.

The time series classification plots further confirm the robustness of the Random Forest model, which accurately reproduced the temporal occurrence of different eruptive styles, including short-lived lava fountain events and extended effusive phases. Misclassifications appear minimal and are likely linked to transitional periods between eruptive phases, where the thermal signal exhibits overlapping characteristics. Overall, these results demonstrate that Random Forest is particularly suited for operational eruption classification using VRP-derived features, offering both high accuracy and stability across

different volcanic activity types. The detailed performance of the classification models is reported in the dedicated Performance Evaluation section.
Overall, the combination of these three approaches demonstrates the versatility of the detection framework. The cascading model ensures operational reliability in optical data, SAM2 enhances spatial precision in segmentation, and the VRP-based Random Forest classification introduces a temporal dimension that captures evolving eruptive dynamics. Together, these methods establish a robust multi-sensor system for the automated detection and interpretation of volcanic activity.

3.6 Tracking

The tracking task focuses on monitoring the spatial and temporal evolution of volcanic activity, with the aim of quantifying changes in the extent and intensity of thermal anomalies over time. This analysis is particularly relevant for eruptions characterized by prolonged effusive activity and complex surface dynamics. The 2021 eruption of Cumbre Vieja (La Palma, Canary Islands) was selected as a representative case study to validate the ability of the cascading model to track thermal features during long-lasting and high-impact eruptive crises.

3.6.1 Data Description and Case Study

The eruption of Cumbre Vieja began on 19 September 2021 on the southwestern flank of the volcano and persisted for nearly three months, represents one of the most significant volcanic crises in the Canary Islands in recent decades. Characterized by the opening of multiple fissures and vents, the eruption produced sustained lava fountains, extensive lava flows, and voluminous ash plumes, causing widespread damage to infrastructure, agricultural land, and residential areas. In particular, lava flows traveled several kilometers westward, eventually reaching and extending into the ocean, forming new coastal lava deltas. This complex and long-lasting activity provides an ideal case study to test and validate satellite-based monitoring methodologies for detecting, mapping, and characterizing volcanic thermal anomalies over time.

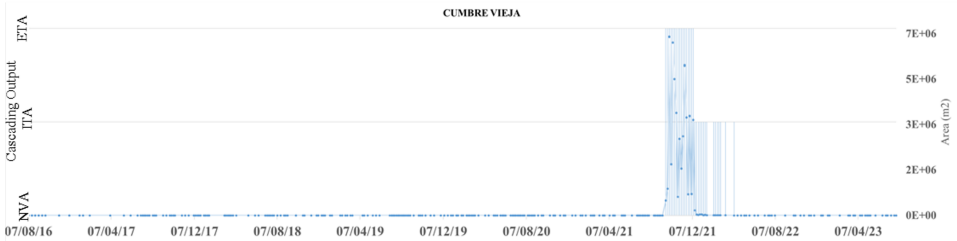

FIGURE 3.35: Temporal analysis of Kilauea's volcanic activity

For this analysis, a sequence of Sentinel-2 MSI images covering the entire eruption period (September–December 2021) was processed using the cascading model. Each acquisition was automatically classified into one of the four target classes, Extended Volcanic Thermal Anomalies (ETA), Isolated Thermal Anomalies (ITA), Cloudy-Sky Conditions (CSC), and Non-Volcanic Activity (NVA), and the areal extent of thermal anomalies was computed for

each scene. Scenes affected by excessive cloud cover were filtered out to ensure accurate quantification. The combination of high revisit frequency and near-real-time processing enabled the continuous tracking of the eruption's evolution, as shown in Figure 3.35.

3.6.2 Results and Discussion

In Figure 3.35, the onset of the eruptive activity of Cumbre Vieja volcano, which lasted for about three months beginning in September 2021, is clearly illustrated. On 19 September 2021, a new eruption started from the SW flank, and the first satellite image available on 20 September shows a partially cloud-covered scene, which prevented the full quantification of the lava flow ($552,400m^2$). Two fissures opened, and multiple vents produced lava fountains, flows, and ash plumes; the lava flows extended over 5 km westward toward the coastline, eventually reaching the ocean and causing severe damage to infrastructure, buildings, and crops. By 30 September, it was possible to visualize and quantify the entire active lava flow, with an estimated extent of $6,672,400m^2$. Strong effusion persisted during the first half of the month, feeding the lava deltas along the coast. Figure 3.36 also allows us to track the

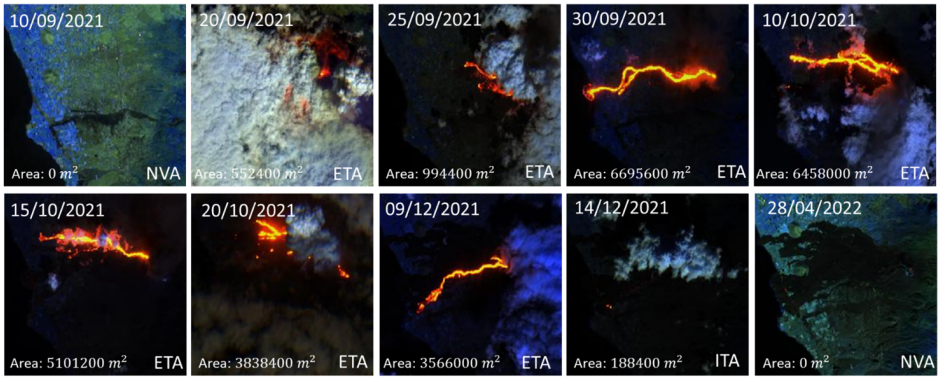

FIGURE 3.36: Image's sequence of Cumbre Vieja's volcanic activity.

complete eruptive sequence, starting from the absence of activity recorded on 10 September 2021 through to the conclusion of the eruption. For each image included in Figure 3.36, the corresponding scene classification is reported, most often as 'Extended Volcanic Thermal Anomalies' (ETA), along with the areal extent of the detected thermal anomaly. Overall, the tracking analysis demonstrates that the cascading framework effectively monitors the temporal evolution of large-scale effusive eruptions, even under challenging observational conditions such as partial cloud cover or complex topography.

3.7 Quantifying

After describing the tracking of thermal anomalies, which makes it possible to follow their position and spatial evolution over time, it is important to introduce the phase of quantification. While tracking provides mainly qualitative and dynamic information, namely where the anomaly is located and how it evolves spatially, quantification aims to measure objective and comparable parameters, such as the affected area, the height and volume. This

approach not only enables the detection of thermal features over time but also allows the assessment of their magnitude and potential correlation with eruptive episodes or changes in the volcano's state. Quantification therefore represents a complementary and essential step to tracking, as it provides the numerical framework needed to interpret the underlying physical processes and to feed nowcasting models.

3.7.1 Data Description and Case Studies

Quantification was performed using Sentinel-2 MSI images and the outputs of the cascading model described in the previous sections. The methodology combines the SqueezeNet (SN) classifier, which labels each scene according to the type of activity (ITA, ETA, NVA, CSC), with the Random Forest (RF) model, which performs pixel-level segmentation to delineate and measure the extent of thermal anomalies. This hybrid structure allows the translation of classification results into quantitative metrics such as hot spot area, lava flow extent, and elevation of the lava fronts.
Two representative case studies were analyzed to validate the quantification approach:

- **Mount Etna (Italy):** quantification of the lava field emplaced during the eruption of 12 November 2023, characterized by intense lava fountains and pyroclastic flows.
- **Stromboli (Italy):** assessment of multiple lava overflow events following the major eruption of 27 September 2023.

3.7.2 Results and Discussion

Mount Etna: Quantification of the November 2023 Lava Flow

After a phase of discontinuous Strombolian activity, volcanic activity at Mount Etna intensified again on 12 November 2023, culminating in a lava fountain. Subsequent pyroclastic flows developed in multiple directions (S, SE, SSW, SW).

The lava flow shown in Figure 3.37 is represented through a false-color composite image created using the NIR, SWIR1, and SWIR2 bands of the Sentinel-2 MSI sensor. This representation highlights the spatial distribution of the lava flows, providing a comprehensive overview of the main eruptive features. The event produced a lava field with an areal extent of approximately $1.15km^2$ and a total lava volume estimated at $2.8Mm^3 \pm 0.8Mm^3$.

Stromboli: Quantification of Lava Overflow Episodes

Following the paroxysm of 27 September 2023, several lava overflow episodes were detected and analyzed (Figure 3.38). On 27 September, the cascading model identified the first overflow, classifying it as an Extended Volcanic Thermal Anomaly (ETA) with a hot area of approximately 36,000 m^2 and a lava front altitude of 300 m a.s.l. A second overflow, initiated on 3 October 2023, was detected on 4 October (during the cooling phase) with a reduced area of 16,000 m^2 and an elevation of 400 m a.s.l, classified as an Isolated Volcanic Thermal Anomaly (ITA). On 6 October, renewed activity was observed: the cascading model classified the event as ETA, identifying a hot

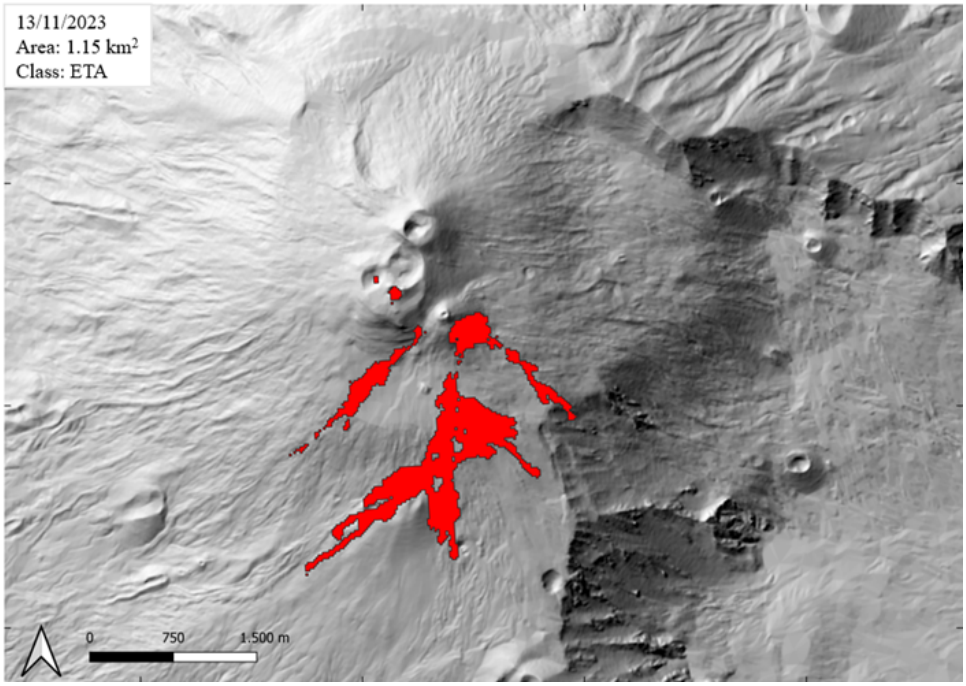

FIGURE 3.37: The lava field was mapped using the bands NIR, SWIR1, and SWIR2 from the image Sentinel 2-MSI 13/11/2023 09:53 UTC (resolution 20 m) and the trained cascading model ML. Its extension is 1.15 km2 and the branch to the southwest reaches the lowest altitude of 2430 m a.s.l. Digital Elevation Model (DEM) (https://tinitaly.pi.ingv.it/). Reference system used WGS84.

area of 34,000 m^2 with a front elevation of 600 m a.s.l. The image available on September 27, 2023, allowed us to detect the first lava flow of the series, for a total of three events (Figure 3.38). On October 3, 6, 7, and 8, four lava overflows were emitted from sector N1 of the N crater area [ref: INGV bulletin (41/2023)]. On September 27, 2023, the cascading algorithm identified the first lava flow of this phase, classifying it as an Extended Volcanic Thermal Anomaly (ETA). The hot area was estimated at approximately $36,000m^2$, with the lava front positioned at an altitude of 300 meters above sea level. The overflow on October 3, which began in the early morning hours, was detected by our model the following day, on October 4 (during the cooling phase), as the data was not available on the same day. The hot area on this date was reduced to about $16,000m^2$, with the lava front advancing to 400 m a.s.l, classified as an Isolated Volcanic Thermal Anomaly (ITA). On October 6, 2023, the activity intensified again, the model detected the lava overflow on the same day as the event, and the algorithm classified the event as ETA, identifying a hot area of approximately $34,000m^2$, with the lava front at an altitude of 600 m a.s.l.

By applying the cascading approach, it is possible not only to detect the onset of eruptive activity but also to quantify the spatial extent of thermal anomalies over time. The combined use of SN for scene classification and RF for pixel-level segmentation provides a robust framework for distinguishing between isolated (ITA) and extended (ETA) anomalies and for translating these classifications into quantitative metrics, such as the hot spot area.

FIGURE 3.38: Overflow episodes, classified by the model, occurred on September 27, 2023, October 4, 2023, and October 6, 2023. Reference system used WGS84.

This quantitative analysis offers crucial insights into the dynamics of eruptive activity, allowing us to track the temporal evolution of volcanic processes and to evaluate the magnitude and persistence of thermal features associated with different eruptive phases. Overall, the quantification analysis highlights the complementarity between scene classification and pixel-based segmentation. The cascading model bridges the gap between qualitative detection and quantitative assessment, enabling consistent estimation of physical parameters such as area and volume across different eruptive contexts.

3.8 Nowcasting

Volcanic paroxysms represent some of the most hazardous manifestations of explosive volcanism, due to their sudden onset and the large amount of energy released in very short timescales. Anticipating such events, even by a few minutes to hours, is of crucial importance for civil protection and risk mitigation. Traditional monitoring networks, while highly effective, may not always guarantee the spatial coverage and temporal continuity needed to capture the precursors of these rapid phenomena. In this context, the integration of satellite-based observations becomes a key component. Due to the rapid and unpredictable nature of eruptions, high-temporal-resolution geostationary satellite data are indispensable for near-real-time monitoring and short-term prediction. The continuous stream of thermal and radiative information provided by geostationary sensors allows for the timely detection of thermal anomalies, sudden increases in radiative power, and other eruptive precursors.

3.8.1 Data Description and Case Study

Building on the preceding phases (forecasting, detecting, tracking, and quantification), the nowcasting task focuses on predicting the probability of imminent paroxysmal activity based on Volcanic Radiative Power (VRP) time series. The approach leverages supervised machine learning models trained

to recognize patterns associated with the onset of explosive activity, enabling near-real-time forecasts of eruptive events. The methodology was tested using VRP data derived from SEVIRI observations of Mount Etna's Voragine (VOR) crater, which has produced some of the volcano's most intense explosive eruptions. Although VOR is generally less active than the other summit craters, it reactivated in mid-2024, producing six lava fountain episodes between July and August, the first such sequence since 2015–2016. This dataset was selected for its clarity and scientific relevance, providing an ideal benchmark for validating the predictive capabilities of the nowcasting framework. For model training and testing, the VRP time series was segmented into fixed-length temporal windows, each characterized by features extracted through the feature engineering strategies described in Section 3.3.1. These include statistical descriptors (mean, variance, slope, kurtosis, skewness, and range) computed over multiple moving-window sizes (10, 30, and 50 samples) to capture both short-term fluctuations and long-term activity trends. The target variable was binary, indicating whether a paroxysmal event occurred within the following temporal window (*precursory increase* vs. *no precursory increase*). The dataset was split into 60% training, 20% validation, and 20% testing subsets.

3.8.2 Results and Discussion

Figure 3.39 presents the time evolution of VRP during the test period, which extended from July 2024 to January 2025 and encompassed multiple paroxysmal episodes. Distinct peaks in radiative power correspond to eruptive phases, while quiescent intervals appear as stable low-level trends, providing a clear structure for predictive modeling.

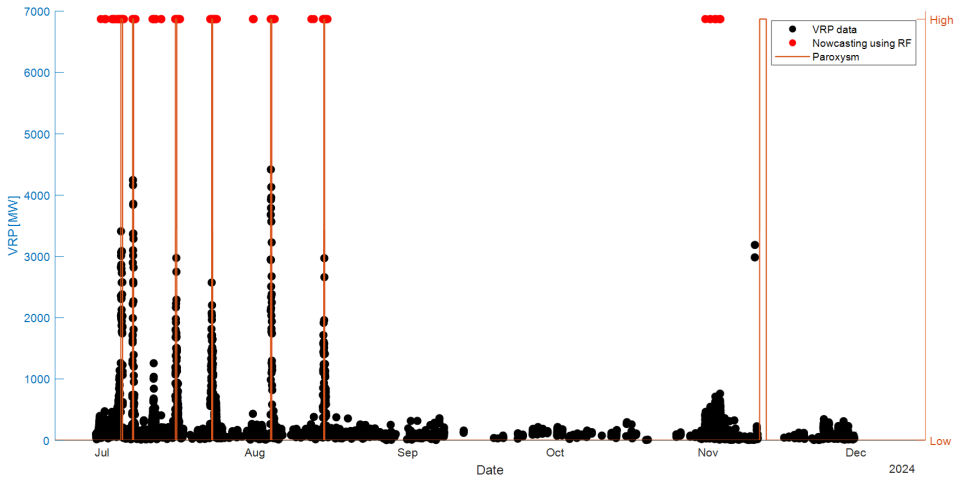

FIGURE 3.39: Random Forest output for the test period (July 2024–January 2025), showing VRP evolution and paroxysmal events.

The Random Forest (RF) model was trained using the features selected through the procedure described in Section 3.3.1. The resulting temporal probability plots (Figure 3.40) show that the model successfully identified the majority of paroxysmal episodes, often predicting their onset several hours in

advance. The predictive lead time varied from less than one hour to over six hours, depending on the intensity and temporal gradient of the precursory VRP increase.

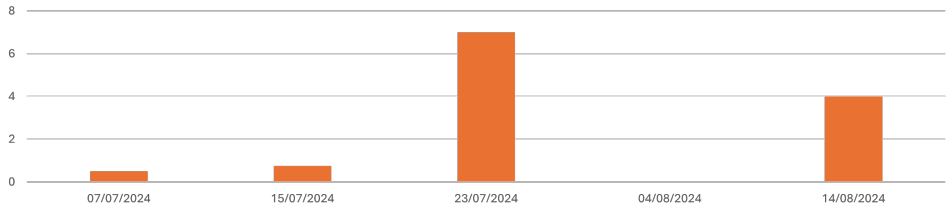

FIGURE 3.40: Temporal probability plot showing the predicted probability of paroxysmal onset over time. High-probability peaks correspond to true paroxysmal events, confirming the model's capability to capture precursory trends in the VRP signal.

For example, the paroxysm of 23 July 2024 was correctly forecasted several hours before onset, demonstrating the model's sensitivity to subtle thermal precursors. Conversely, shorter lead times were observed for weaker or more abrupt VRP variations, which limited the amount of information available for pattern recognition. Despite these challenges, the RF-based nowcasting model achieved high accuracy and reliability across all evaluated events, confirming its suitability for operational applications in near-real-time volcanic monitoring.
Overall, the results demonstrate that satellite-based nowcasting can provide actionable early-warning information by exploiting continuous geostationary observations and data-driven feature extraction. This approach represents a crucial link between quantitative monitoring and real-time decision support, enabling the transition from reactive to proactive volcanic hazard management.

3.9 Tools

The two tools presented in this section build directly upon the methodologies developed earlier in this chapter. The V-STAR algorithm integrates the forecasting, detecting and quantifying tasks to perform real-time monitoring of thermal anomalies. The AI-based Platform, in turn, implements the complete multi-step framework, combining forecasting, detecting, tracking, and quantification into an operational monitoring environment.

3.9.1 V-STAR (Volcanic Satellite Thermal Anomalies Recognition)

Among the tools developed to make some of the previously discussed approaches operationally accessible, a central role is played by V-STAR, an operational and user-friendly system for volcanic monitoring [207]. V-STAR integrates advanced machine learning approaches, specifically adopting the Random Forest algorithm, which enhances both the accuracy and responsiveness of volcanic monitoring. The application is implemented on Google Earth Engine (GEE), leveraging cloud computing to process large volumes of geospatial data in real time. Through an intuitive user interface, the platform

allows for integrated visualization of key volcanic features, while also supporting in-depth analyses that can uncover hidden patterns in thermal data. Thanks to these characteristics, V-STAR emerges as a robust, accessible, and operational tool that combines satellite observations with advanced analytics. In doing so, it contributes not only to improving scientific understanding of eruptive processes but also to strengthening volcanic risk reduction strategies in support of civil protection and emergency management.
The V-STAR application was tested on 14 volcanoes located across different tec-tonic settings and geographic regions: Etna and Stromboli (Italy), Klyuchevskoy and Shiveluch (Russia, Kamchatka Peninsula), Lascar (Chile), Popoca-tépetl (Mexico), Fuego, Pacaya, and Santiaguito (Guatemala), Telica (Nicaragua), Kīlauea (Hawaii, USA), Erta Ale (Ethiopia), Merapi (Indonesia), and Ambrym (Vanu-atu). V-STAR (version: 1.0) offers an intuitive and user-friendly interface that provides all users, including those with limited experience, the essential tools for performing accurate and efficient analyses. The interaction begins by selecting the volcano of interest from a predefined list. Upon selection, the application automatically connects to Google servers, initiates the Random Forest classifier, and retrieves and prepares the relevant data for visualization. Once loading is complete, the interface is divided into two main sections. On the left side (Figure 3.41a), the first graph displays a

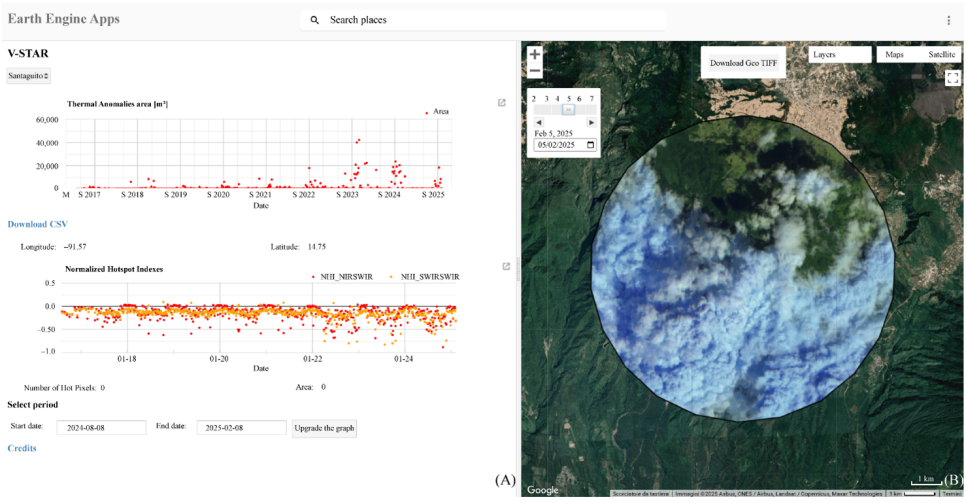

FIGURE 3.41: V-STAR interface. (A) shows the different information for the selected volcano; (B) shows the image from Sentinel-2 MSI related to the selected volcano and the calendar to change the date.

time series of the areal extent of thermal anomalies detected at the selected volcano over the past six months. Users also have the option to download the complete time series, from 2016 to the present, in CSV format for further offline analysis. Below this, the Normalized Hotspot Index (NHI) is plotted over time. The NHI is based on the normalized difference between spectral bands from the Sentinel-2 satellite that are sensitive to thermal anomalies. Specifically, two indices are computed:

- NHI(SWIR–SWIR) $= \frac{B12-B11}{B12+B11}$, where B11 and B12 are both shortwave infrared (SWIR) bands.

- NHI(NIR–SWIR) = $\frac{B11-B8A}{B11+B8A}$, where B8A is the near-infrared (NIR) band and B11 is an SWIR band.

These indices are designed to highlight thermally anomalous pixels by enhancing the radiometric contrast between hot targets (e.g., active lava or fumaroles) and the cooler surrounding background. By default, the NHI is calculated at the center of the crater; however, users can reposition the cursor to any location within the scene to re-calculate the index for the new point. Additionally, the application displays the number of hot pixels identified by the Random Forest classifier and automatically computes the total anomalous area (in square meters) for the selected image. The interface also allows users to define a custom time window for exploring the time series, making it easier to identify temporal patterns, trends, or episodic changes in thermal activity. On the right side of the interface (Figure 3.41b), the most recent Sentinel-2 image available for the selected volcano is displayed. A calendar widget enables users to browse and select specific acquisition dates. If the requested date is not available in the archive, a notification is displayed informing the user. The trained Random Forest model is automatically applied to each available selected image, and when the image contains a thermal anomaly, either intracrateric or more extensive, such as a lava flow or lava lake, the anomaly map is shown. Yellow contours delineate the subtle to high anomaly's extent, while red pixels mark the hottest areas within the segmentation. V-STAR (version: 1.0) also enables the export of classification results. Users can down-load the outputs in GeoTIFF and XML formats, which are compatible with standard geospatial analysis platforms, including GIS software (version: 3.40.3) and Google Earth (version: Pro or other). Some case studies have been reported to demostrate the efficiency of the application. The Ambrym eruption that occurred on 15 December 2018 was investigated. The previous lateral eruptive activity had taken place in 2015, when a minor surface lava flow was recorded within the caldera [216], [217]. The results obtained from the analysis of Sentinel-2 MSI data processed with the RF model reveal the presence of persistent thermal anomalies at the summit craters starting in January 2018, continu-ing up to the onset of the lateral eruption on 15 December 2018 (Figure 3.42). These results align with temperature anomaly data and SO_2 emissions derived from Himawari-8 satellite observations [216], which confirm the initiation of intra-caldera activity. This case study highlights the potential of the V-STAR system, particularly the RF-based classification approach, to track subtle thermal changes during extended periods of apparent quiescence. The ability to detect a progressive increase in thermal anomalies prior to the eruptive event demonstrates the tool's usefulness for monitoring transitions from quiescence to unrest. It also supports the integration of data-driven models and high-resolution satellite imagery in operational early warning systems.

Sentinel-2 MSI imagery combined with the Random Forest model enables effective mapping and monitoring of lava-inundated areas. Two case studies has been presented: the December 2018 eruption of Mount Etna (Italy) and the January–February 2023 eruption of Kīlauea (Hawaii). One of the most recent lateral eruptions of Etna occurred in December 2018 [218]–[220]. From 21 November 2018, thermal anomalies were observed at the Southeast Crater in the summit area. On 8 December 2018, Strombolian activity began at both the Southeast Crater and Bocca Nuova crater in the summit area

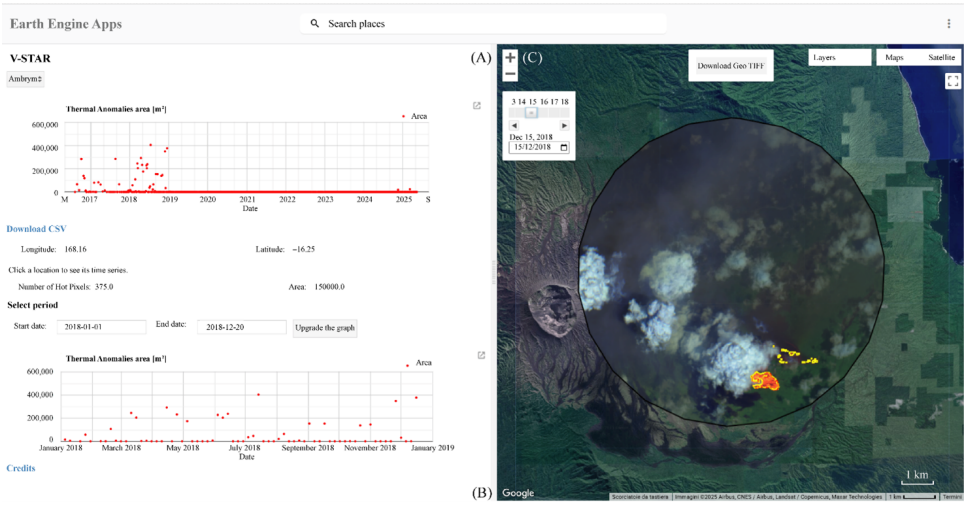

FIGURE 3.42: Thermal anomalies of the Ambrym volcano (Vanuatu). (A) Thermal anomalies area (m2) from 2016 to nowadays; (B) selected period of thermal anomalies; (C) layers menu and thermal anomaly map product from S2-MSI data of 15 December 2018, covering the area of the Ambrym volcano during an intra-calderic effusive eruption.

[INGV-OE, Bollettino Etna settimanale di dicembre 2018; Rep. N. 50/2018 ETNA. Available online: www.ct.ingv.it (accessed on 2 February 2025)]. On the morning of 24 December 2018, a new lateral eruption began. This event was characterized by the intrusion of a magmatic dike into the volcano's upper eastern flank, which triggered a strong seismic swarm and notable ground deformation [INGV-OE, Rep. N. 53/2018 ETNA]. The lateral eruption ended on 27 December 2018, although thermal anomalies associated with the cooling lava flow remained visible until 29 December 2018 (Figure 3.43). V-STAR detected the beginning and end of the lateral eruption. The results show the thermal anomaly area seen on 24 December (0.12 km^2) associated with the emplacement of the lava flow in the first phase of the lateral eruption (Figure 3.43a). The end of the effusive activity was assessed on 27 December, when the cooling lava flow was shown with a residual thermal anomaly area of 0.28 km^2 (Figure 3.43b), indicating rapid flow development. The V-STAR application also allows users to download thermal anomaly maps in GeoTIFF format, which can be directly imported into GIS platforms, such as QGIS.
As shown in Figure 3.44, the lava flow can be visualized over a digital elevation model (DEM) of Etna, enabling users to track its progression over time and produce real-time geological maps. The second case study focuses on the latest reactivation of the lava lake at Halema'uma'u Crater, located at the summit of Kīlauea. The eruption began on 5 January 2023 and ended on 19 February 2023 (Hawaiian Volcano Observatory https://www.usgs.gov/observatories/hvo (accessed on 2 October 2024). The lava coverage was monitored by integrating Sentinel-2 MSI observations with the Random Forest classification.
The first cloud-free image, acquired on 7 January 2023, revealed a thermal anomaly corresponding to the active lava lake, with an estimated surface area of 1.7 km^2 (Figure 3.45a). After five days, on 12 January, the thermal anomaly had reduced to 0.89 km^2 (Figure 3.45b). On 6 February, the last cloud-free image prior to the end of the eruption revealed a renewed increase

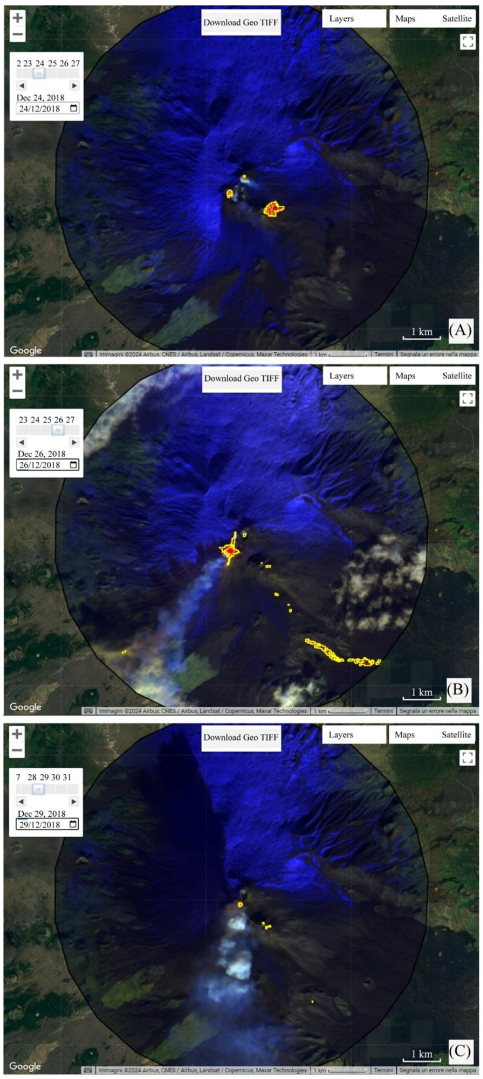

FIGURE 3.43: Thermal anomaly map from Sentinel-2 MSI from (A) 24 December 2018, (B) 26 December 2018, (C) and 29 December 2018 of Etna (Italy).

in activity, with the affected area reaching approximately 1.08 km² (Figure 3.45c). These results confirm the potential of the V-STAR system to provide high-resolution, time-sensitive mapping of lava flows, which is critical for hazard assessment, response planning, and updating volcanic geodatabases.
Accurate localization of active vents, identification of reactivation phases, and assessment of thermal anomalies reported by other systems are among the most relevant capabilities offered by the V-STAR application. This section presents three illustrative cases: the reactivation of Bocca Nuova crater in the summit area of Etna volcano (Italy) in 2018, the reactivation of the Voragine crater (also in the summit area of Etna) in 2024, and the paroxysmal event at Stromboli (Aeolian Islands, Italy) in 2024.
Figure 3.46 illustrates the potential of the system, showing thermal anomaly maps derived from Sentinel-2 MSI data between August and December 2018 for the Mt. Etna summit area. A marked increase in the areal extent of thermal anomalies at Bocca Nuova (BN) is observed starting in August 2018

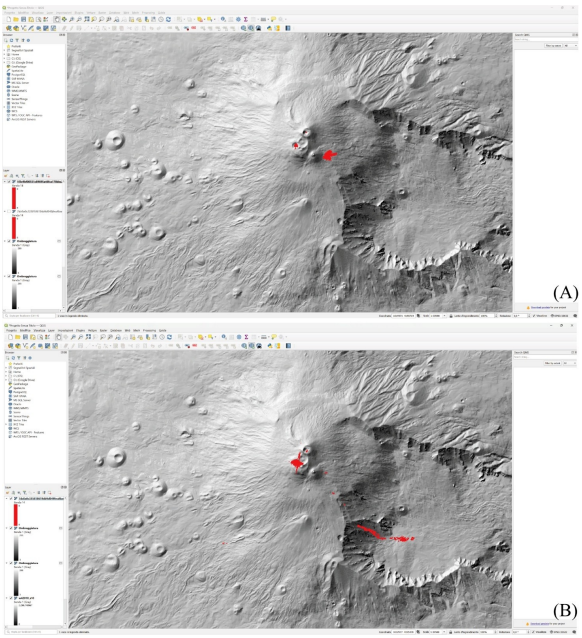

FIGURE 3.44: Geotiff file extracted from the V-STAR App of the (A) 24 December 2018 and (B) 26 December 2018 eruption of Etna (Italy) and plotted on a DEM via QGIS software (version: 3.40.3).

(Figure 3.46a), associated with intensi-fied degassing and intermittent Strombolian activity. At that time, two pit craters, BN1 and BN2, were present inside the Bocca Nuova crater. This Strombolian phase is fol-lowed by the reactivation of BN2 (Figure 3.46b), characterized by sustained gas emissions, discontinuous Strombolian explosions, and several localized high-temperature gas venting points. Subsequently, the fusion of BN1 and BN2 into a single, thermally active Bocca Nuova vent is evident (Figure 3.46c), culminating in increased intracrater activity shortly before the 24 December 2018 eruption (Figure 3.46d). This event was marked by violent Strombolian activity and copious ash emissions [INGV-OE, Bollettino set-timanale di Etna di Dicembre 2018; Rep. N. 53/2018 ETNA. Available online: www.ct.ingv.it (accessed on 2 February 2025)].

Figure 3.47 illustrates the potential of the system, showing thermal anomaly maps derived from Sentinel-2 MSI data in June 2024 for the Mt. Etna summit area. On the night of 13–14 June 2024, activity resumed at the Voragine crater after nearly three years of quiescence (Figure 3.47a). The initial spattering evolved into Strombolian activity over the following days (Figure 3.47b,c). Beginning on 29 June, an intracrateric lava flow was reported within Voragine, progressing toward the interior of Bocca Nuova, specifically toward the BN2 pit crater (Figure 3.47d). On 4 July, the Strombolian activity that had started in mid-June intensified significantly, accompanied by the emergence of two lava flows. These were generated by two distinct vents located on the southeastern and northwestern flanks of the scoria cone inside Voragine. A few hours later, Strombolian activity transitioned into a lava fountain episode.

Figure 3.48 illustrates the potential of the system, showing thermal anomaly

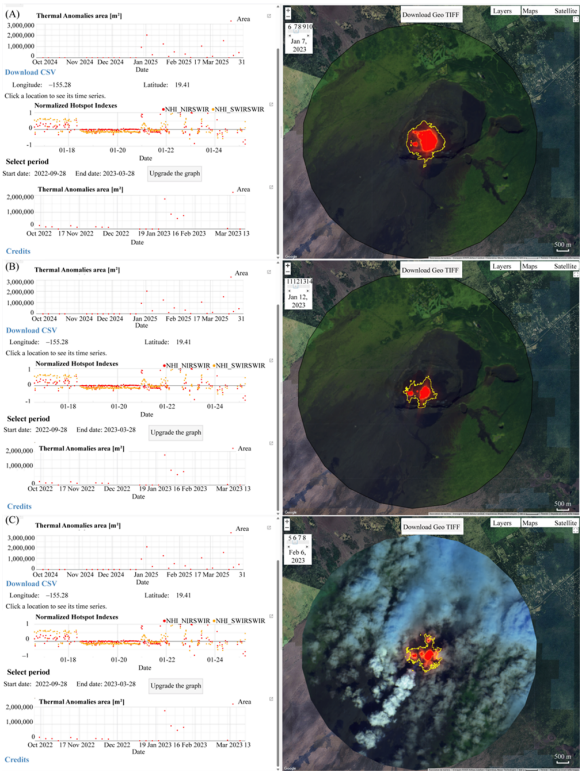

FIGURE 3.45: Thermal anomaly map from S2-MSI from (A) 7 January 2023, (B) 12 January 2023, and (C) 6 February 2023 of Kīlauea lava lake.

maps derived from Sentinel-2 MSI data in June–July 2024 for Stromboli volcano. An intense eruptive phase began with a lava overflow on 23 June 2024, culminating in a paroxysmal event on 11 July 2024, which included multiple overflows, pyroclastic density currents, and debris avalanches [INGV-OE, Bollettino settimanale di Stromboli di Luglio 2024; Rep. N. 29/2024 STROMBOLI. Available online: www.ct.ingv.it (accessed on 2 February 2025)]. Following the brief lava overflow on 23 June, a second overflow was recorded on 28 June (Figure 3.48a,b). On 3 July, intense spattering activity in sector N1 of the northern crater culminated in the collapse of the outer cone margin over-looking the Sciara del Fuoco. This event was followed by another lava overflow from the same N1 vent, accompanied by explosive activity and spattering. On 4 July, effusive activity resumed from two eruptive vents aligned along the Sciara del Fuoco. From that date, both explosive activity and active lava flows remained visible (Figure 3.48c,d). On 11 July, a high-energy explosive event occurred, classifiable as paroxysmal, and was generated by the northern crater terrace area. This explosion produced pyroclastic flows with rapid, widespread expansion toward the sea.

These case studies confirm the capability of the V-STAR system to accurately identify spatially and temporally localized thermal anomalies, track vent reactivation, and support interpretation of eruptive dynamics in near-real time. V-STAR represents an innovative example of how cloud computing, combined with remote sensing and machine learning, can be used effectively for volcanic hazard monitoring. Built on Google Earth Engine (GEE) and employing a supervised Random Forest (RF) algorithm, the system provides accurate and

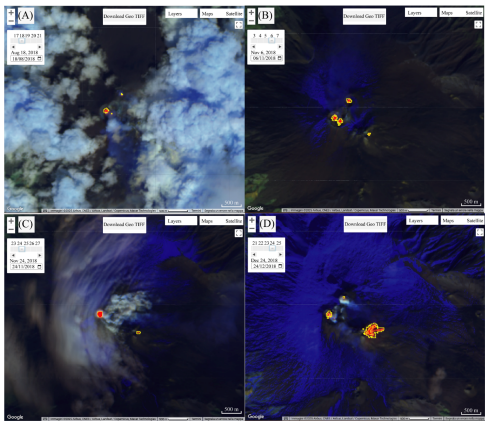

FIGURE 3.46: Mt. Etna (Sicily, Italy) thermal activity of August–December 2018: (A) 18 August 2018 Bocca Nuova (BN) thermal anomalies; (B) 6 November 2018 reactivation of BN2; (C) 24 November 2018 the fusion of BN1 and BN2 into BN; (D) BN intracraterial activity before the 24 December 2018 activity.

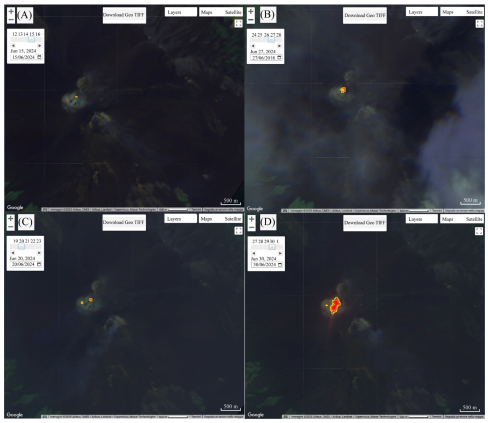

FIGURE 3.47: (A) Activity at the crater Voragine, 13–14 June 2024; (B, C) increasing activity at the Voragine crater; (D) emission of an intracraterial lava flow from the Voragine crater into the interior of the Bocca Nuova pit crater.

timely information on the presence and evolution of thermal anomalies, including the mapping of lava flows and active vents. One of V-STAR's key strengths is its generalization. The model is adaptable to a variety of volcanic systems with different eruptive behaviors, compositions, and geological settings. This versatility makes it a valuable tool for both the scientific community and monitoring authorities, facilitating global-scale assessments and enhancing understanding of complex volcanic phenomena.

3.9.2 The Platform

This final section encompasses and integrates all the algorithms presented in the previous parts of this work, providing a unified vision for the development and operational management of a volcanic monitoring tool [221]. Rather than focusing on a single stage of the monitoring chain, it offers a comprehensive framework that spans the entire temporal continuum, from forecasting to nowcasting, combining the methodological advances described earlier into a

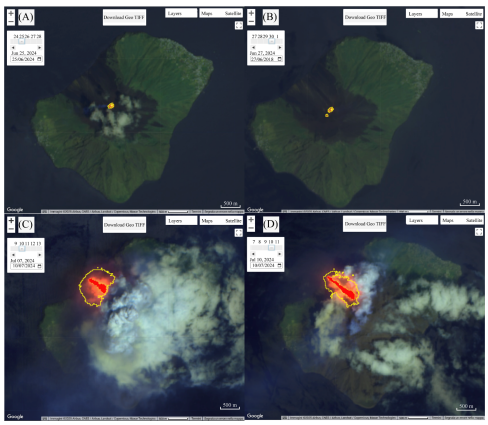

FIGURE 3.48: Stromboli Island (Aeolian Archipelago, Italy) thermal activity before the paroxysm on 11 July 2024: (A, B) increasing activity after the two overflows on 23 and 28 June 2024; (C, D) in-creased effusive and explosive activity after a sector collapses, and pyroclastic flows of 3 and 4 July 2024.

coherent, AI-driven system. This integrated approach leverages the knowledge gained from the individual tasks to optimize the design, functionality, and responsiveness of the monitoring platform, ultimately supporting more effective hazard assessment and decision-making.

The capability to acquire and process satellite data in near real-time (NRT) is a cornerstone of modern volcanic hazard monitoring, enabling the timely identification of hazardous scenarios and the implementation of appropriate mitigation measures. The efficiency of such systems relies on the temporal resolution of the available satellite sensors, which can vary from a few minutes (e.g., SEVIRI, MODIS) to several days (e.g., MSI, OLI). Recent advances in volcanology have increasingly integrated artificial intelligence (AI), leveraging both Machine Learning (ML) and Deep Learning (DL) techniques to learn from historical datasets and construct a robust knowledge base of volcanic processes. In this context, an AI-driven platform has been developed to perform end-to-end monitoring of volcanic activity, encompassing the entire operational chain from forecasting the onset of unrest, to detecting eruptive initiation, tracking and quantifying eruptive products, and finally nowcasting ongoing activity. The platform processes large volumes of multi-spectral satellite imagery, combining NRT data streams with offline analysis. ML and DL algorithms autonomously assess volcano status, extracting relevant features and triggering task-specific modules based on AI-driven decision logic. These modules are designed to address distinct monitoring objectives, such as identifying early unrest indicators, characterizing eruption dynamics, and quantifying erupted volumes using satellite thermal measurements. This approach represents an unprecedented convergence of technological innovation and scientific expertise, overcoming many of the limitations of traditional monitoring techniques by exploiting heterogeneous satellite datasets in an integrated manner. The resulting system supports informed, timely, and operationally relevant decision-making, enhancing the capability to mitigate volcanic risk and reducing the potential impacts of eruptive events on populations, infrastructure, and the environment.

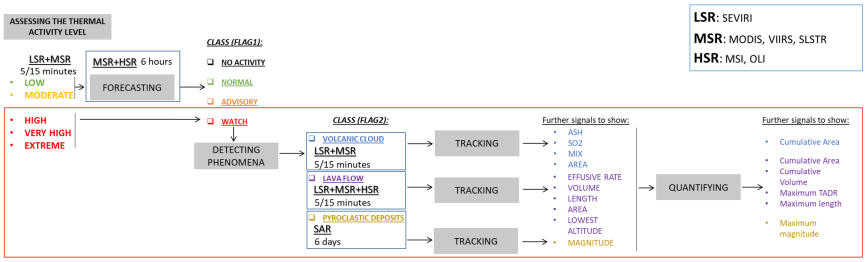

FIGURE 3.49: Workflow AI-based platform.

The use of AI, specifically ML and DL, is the cornerstone in the development of the proposed advanced volcanic monitoring platform. Its modules act both independently and collectively, representing an example of a System of Systems. The outcome, i.e., the status of the volcano, is determined by a combination of all modules. The advanced volcanic monitoring platform is structured into four key sections (Forecasting, Detecting, Tracking and Quantifying) as shown in Figure 3.49. Each section is designed to address specific needs and objectives in volcanic activity monitoring and analysis, while sharing a common unique goal, automatically assessing and describing the status of the volcano. The platform operates synergistically and recursively, with different modules activated sequentially in response to developments in volcanic activity. It provides continuous NRT monitoring of multiple volcanoes, offering a quick overview of their status with a color-coded icon indicating green, orange, or red based on the activity state determined by the outcomes of the Forecasting and Detecting modules. The primary objective of the Forecasting section is to issue warnings in the event of changes occurring during the unrest phase. Eruptions are often preceded by various indicators detectable from space, including subtle increases in surface temperature. Advanced algorithms of DL, i.e., CNN, were adopted to extract subtle to large volcanic thermal features segment the image. Forecasting module will produce either "advisory state" class if any change is detected or "normal state" class otherwise. In this case, an orange icon will be displayed together with the thermal time series showing such a trend. When the eruption officially begins, the system transitions to the crucial Detecting phase. This phase automatically switches the activity state to "watch state" based on all available satellite thermal data, using the Anomaly Detection ML algorithm, specifically Isolation Forest. During this phase, the platform is designed to accurately and promptly detect anomalous volcanic thermal activity and recognize any volcanic hazard phenomena. The Volcanic Radiative Power (VRP) time series, depicting the evolution of emitted radiance, will automatically appear. This model activates "Lava Flow", "Volcanic Cloud", and "Pyroclastic Deposits" based on the analysis of the available satellite data based on the scheme shown in Figure 3.49. As an example, the platform can detect the onset of a lava flow and then quantify its geometrical features (e.g., area, length) thanks to sophisticated detection algorithms supported by advanced ML and DL techniques, i.e., cascading models based on Random Forest and Squeezenet. Once any volcanic phenomenon is detected through the activation of the "Detecting" module, the system automatically transitions to the next tracking phase. Space-based imagery tracks activity over time, providing

information on the emplacement of volcanic deposits, the presence and characteristics of ash plumes, and potential changes in the eruption's nature. This tracking implementation focuses on the spatial-temporal evolution of eruptive products (ash clouds) using an ML approach, specifically Support Vector Machine (SVM), to discriminate between the spectral features of volcanic clouds components based on training data. Finally, after guiding users through each phase of volcanic activity, we reach the last crucial section: Quantification. This section represents the culmination of monitoring, collecting, and providing precise measurements and characterizations of volcanic activity. It offers detailed assessments of hazard magnitude, such as the amount of material emitted and the altitude reached by the lava flow. This information guides a ML model to determine if the phenomenon poses limited hazards, maintaining the activity state as "watch", or if it requires as shift to "warning". The outcomes of each module will be integrated into an ML model to assess the type of eruption, such as strombolian activity, major explosion, paroxysm, and effusive activity. Finally, an offline mode will allow users to select the volcano of interest and explore historical data. Users are introduced to an intuitive and interactive interface that offers considerable flexibility in navigating through different volcanic parameters over time.

The proposed platform demonstrates an innovative contribution, introducing a new way for thinking about volcano monitoring from space. Thanks to AI, a NRT description of volcanic behavior is automatically contextualized by merging spatiotemporal information from various satellite sources with advanced ML models. The significance of this work extends far beyond; it represents an engineering innovation that integrates AI methodologies to address complex and dynamic challenges, ranging from eruption prediction to designing solutions for risk mitigation. This fusion of volcanology and engineering, supported by AI, bridges the gap between understanding natural processes and technological innovation, opening the door for highly advanced systems capable of dynamically adjusting to changing conditions in volcanoes. This approach not only sparks scientific interest due to its contribution to predictive accuracy but also revolutionizes volcanology by enabling the development of evacuation and protection strategies driven by real-time data.

Chapter 4

Case studies in the Satellite Era: Volcanic Forcing of Climate

This chapter provides a preliminary analysis of selected components within a processing chain that spans from forecasting to nowcasting. The aim is to outline how an approach originally developed for volcanic hazard monitoring, based mainly on thermal and infrared data, can be transposed and extended to the study of atmospheric processes. This requires both a reconfiguration of observational bands and an adaptation of analytical criteria, while preserving the overall logic of the original workflow.

A preliminary step is the definition of selection criteria and data sources for identifying significant eruptions. This ensures the temporal and spatial relevance of the case studies and establishes the quality and consistency of the observational datasets, providing a robust foundation for the methodological developments that follow. The use of reanalysis products such as ERA5 is also explored to identify latent correlations between volcanic activity and atmospheric responses, signals often undetectable through conventional monitoring alone. An approach for the detection and quantification of sulfur dioxide (SO_2) using TROPOMI satellite data is presented as a further operational extension, enabling consistent measurements of volcanic gas emissions. The potential of emulator-based frameworks is investigated to bridge physical modeling and data-driven prediction, offering a pathway toward scenario, based volcanic forcing simulations. Finally, a methodology for volcanic cloud nowcasting is discussed as an operational extension of detection and analysis, supporting rapid forecasting and risk management.

Overall, the approaches presented in this chapter represent an extension of volcanic hazard monitoring practices to the assessment of the climatic impacts of eruptions. By integrating segmentation, nowcasting, and reanalysis methodologies, the chapter highlights how data-driven techniques can support the understanding of volcanic–climate interactions. Some sections present

fully developed implementations supported by results, while others introduce conceptual frameworks for future exploration.

4.1 Selection Criteria and Data Source of Significant Eruptions

As discussed in previous chapters, volcanic eruptions typically release gases such as H_2O, N_2, CO_2 and SO_2. Over the lifetime of the Earth, these emissions have represented the primary source of the planet's atmosphere and oceans, after the primitive atmosphere was lost to space. Water vapor condensed to form oceans, while CO_2 was gradually transformed by plants into O_2, with part of the carbon stored in fossil fuels. In practical terms, humans and all living beings are, in a very real sense, composed of volcanic emissions: we drink the water, breathe the oxygen, and consume the plants (and the animals that feed on them) that ultimately owe their existence to volcanic outgassing [1].
From a climatic perspective, however, the most important effect of explosive volcanic eruptions is linked to the emission of sulfur species into the stratosphere, mainly in the form of SO_2, and occasionally as H_2S. These sulfur compounds react with OH and H_2O to form H_2SO_4 within weeks, producing sulfate aerosols that dominate the radiative forcing from volcanic events. In the troposphere, SO_2 conversion to H_2SO_4 generally occurs within a few days, while in the stratosphere it can take from several weeks to months. The resulting sulfuric acid and sulfate aerosols increase the optical depth of the atmosphere and enhance its albedo.

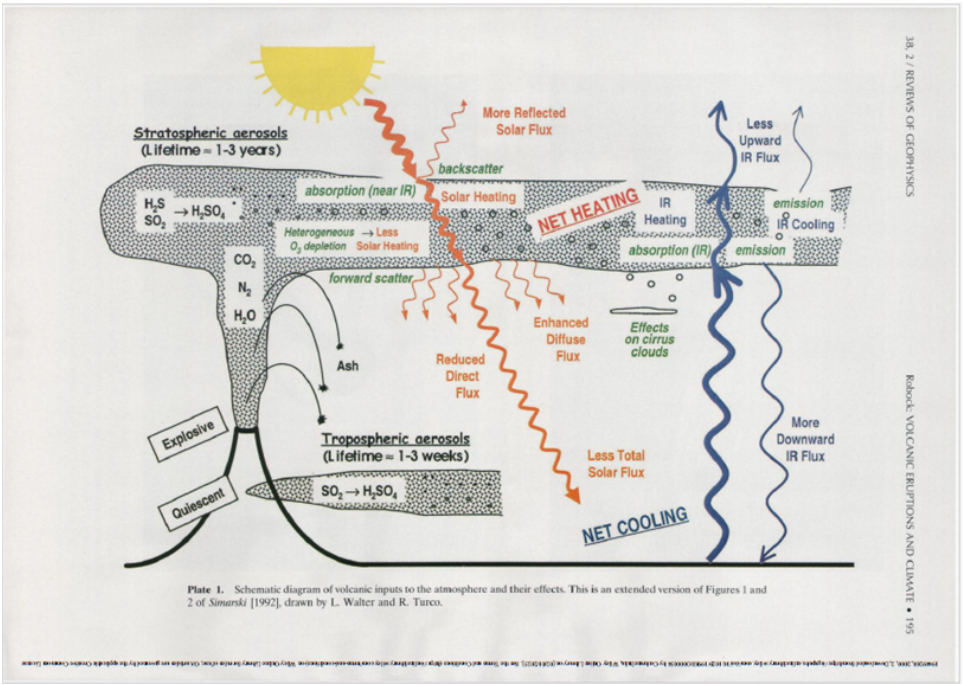

FIGURE 4.1: Indicates the major radiative processes resulting from the stratospheric aerosol cloud from a major volcanic eruption. Image courtesy of [1].

Figure 4.1 illustrates the main radiative processes triggered by the stratospheric aerosol cloud from a major volcanic eruption. Sulfate aerosols, with a typical effective radius of 0.5 µm and a single-scattering albedo close to 1, strongly scatter visible sunlight. Part of this radiation is backscattered to space, increasing planetary albedo and reducing the solar energy reaching the surface, which leads to net surface cooling. Forward scattering enhances diffuse downward radiation, partially offsetting the loss of direct sunlight.
Satellite measurements by [187] estimated that the 1982 El Chichón eruption injected about 7 Mt of SO_2 into the atmosphere, while the 1991 Pinatubo eruption injected approximately 20 Mt. These figures highlight the magnitude of emissions that can drive significant climatic impacts. Defining what constitutes a significant eruption is complex, non-trivial, and context dependent. Past case studies show that even volcanoes located at similar latitudes can produce very different climatic effects in various regions of the hemisphere. As discussed in Paragraph 1.2.1, both location and seasonality must be evaluated on a case-by-case basis. Key criteria for assessing whether an eruption can be considered significant include: Geographic location: Eruptions near the equator tend to have global impacts, while those at mid-to-high latitudes generally produce hemispheric effects; SO_2 injection: The total mass of SO_2 injected into the atmosphere, directly related to the Volcanic Explosivity Index (VEI) and the height of the eruption column; Timing within the year: Seasonal factors and pre-existing climate conditions can influence the atmospheric impact. For example, polar winter eruptions can more easily reach the stratosphere due to a lower tropopause, potentially extending aerosol residence time compared to tropical events; Interaction with climate phenomena: Events such as El Niño can modulate volcanic cooling effects, as observed for the 1982 El Chichón eruption, which coincided with a strong El Niño and displayed attenuated cooling. These criteria form the basis for selecting and analyzing significant volcanic eruptions in the context of their potential climatic influence.
Our understanding of the main mechanisms and physical processes linking volcanic eruptions to climate has not fundamentally changed since the comprehensive review by [1]. However, one significant development since then is the potential application of volcano–climate knowledge to addressing pressing environmental challenges. The relatively short residence time of volcanic aerosols in the stratosphere limits the direct climatic impact to the few years immediately following an eruption. Nevertheless, longer-term effects are possible through feedback mechanisms within the climate system. Today, the magnitude and duration of volcanic impacts can be investigated by examining volcanic and climate records, as well as by using climate models capable of simulating the complex interactions between volcanic aerosols and atmospheric, oceanic, and biospheric processes.
The first of the two observed warming phases occurred during the first half of the 20th century (Figure 4.2), a period marked by notable volcanic quiescence. Only two major eruptions, Santa Maria in 1904 and Katmai in 1912, took place early in this interval. The relative absence of significant volcanic activity may have contributed to the warming observed between 1910 and 1940. In contrast, the second half of the century saw several large eruptions; yet global temperatures continued to rise substantially (Figure 4.2). This pattern indicates that greenhouse warming, driven by the ongoing accumulation of fossil-fuel CO_2 and other anthropogenic greenhouse gases, had begun to

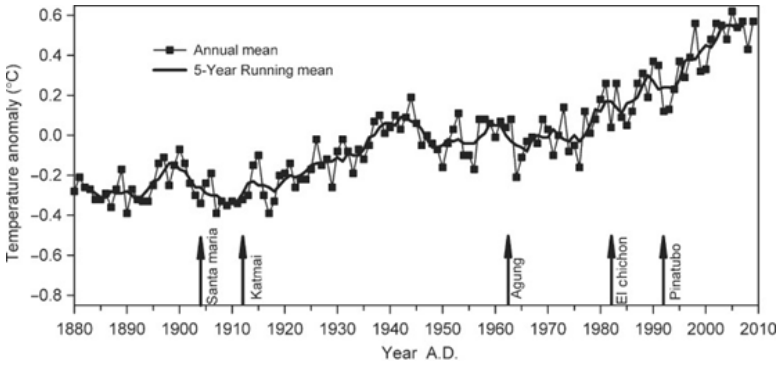

FIGURE 4.2: Global mean surface temperature anomalies (annual mean and 5-year running mean) from 1880 to 2010. Major volcanic eruptions (Santa Maria, Katmai, Agung, El Chichón, Pinatubo) are indicated with arrows along the timeline, highlighting their potential short-term cooling effects on climate variability.

outweigh climate variations from natural factors such as volcanic eruptions. Even major events, Agung (1963), El Chichón (1982), and Pinatubo (1991), produced only short-lived interruptions to the overall warming trend (Figure 4.2).

Anthropogenic greenhouse gas emissions are expected to remain the dominant driver of climate change over the coming decades and centuries. Reliable projections of future climate require a robust quantitative understanding of all forcing agents within the climate system. Among natural forcings, volcanism plays a key role in short-term climate variability, and its influence continues to be an active focus of scientific investigation.

4.2 Materials

4.2.1 Satellite Data

TROPOMI (Tropospheric Monitoring Instrument) is onboard the Sentinel-5 Precursor (Sentinel-5P) mission, developed by the European Space Agency (ESA) as part of the Copernicus program. Sentinel-5P was launched on October 13, 2017, into a sun-synchronous polar orbit at an altitude of approximately 824 km. The satellite completes one full orbit every 101 minutes, providing global daily coverage. TROPOMI is a push-broom imaging spectrometer that measures solar radiation backscattered by the Earth's atmosphere and surface across ultraviolet (UV), visible (VIS), near-infrared (NIR), and short-wave infrared (SWIR) spectral ranges. TROPOMI has a native ground resolution of approximately 3.5×7 km^2 (since August 2019) for most of its products, including sulfur dioxide, allowing detailed atmospheric trace gas mapping on a global scale. The specific data product used in this study is the Level-2 SO_2 column density retrieval, which provides estimates of sulfur dioxide concentrations in Dobson Units (DU) at 7 km. The version processed and distributed by NASA through the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) (earthdata.nasa.gov) has been used in the following methodology. The data were

generated using the DOAS (Differential Optical Absorption Spectroscopy) retrieval technique and provided in NetCDF format with full geolocation and quality assurance metadata. The analysis, focused on the SO_2 column values retrieved at an assumed plume height of 7 km, which is representative of typical volcanic injection altitudes into the lower stratosphere. These data are well-suited for detecting and tracking volcanic plumes following both explosive eruptions and continuous degassing events. Each SO_2 product includes: Latitude and longitude grids, Total vertical column density in DU, Quality flags and averaging kernels, Cloud cover fraction and surface pressure, Time and orbit metadata. The dataset offers high temporal coverage (daily) and sufficient spatial resolution to observe plume evolution and dispersion across regional and global scales.

SEVIRI (MSG) is the other satellite used in this chapter, which has already been introduced in the previous chapter, but here will explore the same sensor using different bands.

The Spinning Enhanced Visible and InfraRed Imager (SEVIRI) is the primary sensor onboard the Meteosat Second Generation (MSG) geostationary satellites, developed by EUMETSAT for operational monitoring of the atmosphere and Earth's surface. Operational since 2002, SEVIRI has been a fundamental instrument for more than two decades in monitoring volcanic emissions, particularly ash clouds, over the European and African regions. The sensor acquires multispectral imagery across 12 channels spanning the visible, near infrared, and thermal infrared spectrum, with a spatial resolution of about 3 km at nadir and a revisit frequency of 15 minutes. These features have made SEVIRI invaluable for near-real-time detection of rapidly evolving atmospheric phenomena. SEVIRI integrates critical spectral bands for distinguishing volcanic ash clouds from ice clouds, particularly the channel pairs centered around 11–12 μm and 8.7–11 μm. Radiometric differences between these bands form the basis of many threshold-based algorithms for automated ash detection and have also been a distinctive element in early applications of CNNs for the recognition of volcanic plumes.

Beyond ash monitoring, SEVIRI has also proven effective in detecting and estimating the distribution of sulfur dioxide (SO_2) emitted during volcanic eruptions, owing to the spectral sensitivity of the 8.7–11 μm band pair. The radiative differences between these channels highlight the specific absorption features of SO_2 molecules in the thermal infrared, enabling the separation of SO_2 clouds from ash and from common water or ice clouds. The use of the 8.7–11 μm difference is therefore crucial not only for ash–ice discrimination, but also for the identification and tracking of volcanic gas clouds, providing critical information for aviation safety and public health risk management in affected areas.

4.2.2 Volcano selection

Mount Etna, one of the most active volcanoes in the world, is located on the east coast of Sicily in the Metropolitan City of Catania, Italy (latitude 37.75°N, longitude 14.99°E). It is characterized by persistent degassing, particularly from its summit craters, continuously releasing sulfur dioxide (SO_2) into the atmosphere. In the next section, the effusive eruptive phase between February and March 2025, when SO_2 emissions experienced a notable increase, will be analyzed. According to Sentinel-5P TROPOMI data, SO_2

emissions began to rise significantly starting 11 February 2025, reaching a peak of approximately 7.0kt/day on 25 February 2025 ()link to the website. This specific time window provides an excellent opportunity to analyze how well SAM 2 can segment and quantify continuous SO_2 emissions during sustained eruptive activity, offering insights into plume behavior and model performance under active degassing conditions.

The **Hunga Tonga Hunga Haapai (HTHH)** submarine volcano, located at approximately 20.55°S, 175.38°W, lies about 65km north of Tongatapu (Tonga's main island) within the active Tonga–Kermadec volcanic arc. The main explosive phase occurred on 15 January 2022, producing one of the most powerful eruptions of the modern era (VEI 5-6). The eruption sent a towering plume of ash, vapor water, and gas into the mesosphere, reaching altitudes of up to 55km [222]. Satellite retrievals estimate that the eruption injected approximately 0.5Tg of sulfur dioxide (SO_2) into the stratosphere. This amount is substantially lower than that released by historic eruptions like Pinatubo in 1991 (15Tg), and hence the climatic impact from SO_2 is expected to be minimal. Regarding dispersion, the SO_2 plume initially spread in a southeast–northwest direction, eventually covering distances up to 6,24km within approximately 30 hours after the eruption.

4.3 Methodology

4.3.1 Zero-Shot Segment Anything Model (SAM2)

Most zero-shot learning (ZSL) approaches in remote sensing have been developed and evaluated primarily on visible-range (RGB) imagery [223]. These methods commonly rely on pretrained computer vision architectures or vision-language models such as CLIP, which are inherently optimized for natural scenes with well-defined objects and semantic richness. Although efforts to extend ZSL to hyperspectral imaging (HSI) are increasing, most of them adopt a few-shot learning (FSL) configuration, where a small number of labeled examples for novel classes are provided for model adaptation [224]–[226]. These methods typically leverage metric-based learning, prototype networks, or transductive inference to overcome the high dimensionality and low interclass variance of HSI data [227]. Pure zero-shot segmentation in HSI remains relatively underexplored. More recently, researchers have begun applying ZSL to other remote sensing tasks such as change detection, land cover classification, and object segmentation. Among them, foundation models such as the Segment Anything Model (SAM) and SAM 2 in zero-shot mode have been used to segment objects or regions of interest across various domains, including urban village boundary extraction [228], medical imaging [174], land use and land cover (LULC) classification [229], and synthetic aperture radar (SAR) imagery [230]. While these applications confirm the generalization ability of foundation models across diverse visual domains, their extension to non-RGB, physically derived datasets, such as ultraviolet spectral imagery used for atmospheric gas retrieval, remains virtually unexplored.

However, a growing body of work has started to address atmospheric monitoring with AI. For example, [69] proposed a deep learning framework for quantifying volcanic SO_2 emissions from Sentinel-5P TROPOMI over Mt. Etna, demonstrating the potential of artificial intelligence in interpreting UV-based gas retrieval products. This work addresses that gap by evaluating the

capabilities of SAM 2 in segmenting SO_2 plumes in Sentinel-5P TROPOMI data under a zero-shot setting. The input preparation stage represents one of the most challenging aspects of this work, especially when dealing with satellite imagery. The TROPOMI *SO2__column_number_density* product originally featured a ground resolution of approximately 3.5 × 7 km, which since 2019 has been improved to 3.5 × 5.5 km, allowing for a finer spatial resolution. When working with satellite data, particularly with the aim of detection and quantification, it is crucial to pay close attention to the resolution grid to which the product must be mapped, both to ensure correct visualization and, above all, to maintain data consistency. In this case, the Level-2 COBRA(CBR) product made available by Copernicus has been used. This detail is essential when the goal is to quantify atmospheric products, as in the present case. Since the downloaded product is global, the study area corresponding to each selected volcano is extracted from the full dataset.
From the original array, a sub-image is then generated that is perfectly consistent with the source product in terms of dimensions, spatial resolution, and pixel values. This sub-image is subsequently passed as input to the SAM 2 model for segmentation. When working with SAM 2, it is also necessary to adapt to the type of input the model expects. The literature suggests that, for optimal results, the model should receive images of at least 1024 × 1024 pixels, or at most 512 × 512 in constrained scenarios. To place the model in the best possible conditions for segmentation, the images created and passed to SAM 2 in this study have an average resolution of about 1000 × 2000 pixels. This choice depends on the specific study area: for example, in cases like the Hunga Tonga and Raikoke eruptions, where the SO_2 cloud later drifted away from the source point, the area of interest became larger, requiring correspondingly larger input images.
The SAM2 architecture was already discussed in the previous chapter. Here, in the zero-shot setting, to guide the model in recognizing the object within an image, a selection logic that maximizes the useful information from the input has been adopted. Specifically, identifyng the point with the highest value in the input image; this point corresponds to the highest likelihood of object membership. Since neighboring pixels tend to have similar values, selecting the maximum point helps create spatial continuity in the target object, thereby improving the consistency and completeness of the resulting segmentation.
Once the prompt point is provided to the model, SAM 2 predicts the segmentation by generating three masks associated with the input image. Among these three masks, the one with the highest score is always selected, ensuring that the final segmentation corresponds to the most confident prediction. Furthermore, the zero-shot mode was also tested both with and without prompting in more challenging cases, for example, when the SO_2 region to be segmented did not exhibit spatial continuity. In such scenarios, indicating a single point within the input often resulted in identifying only a portion of the SO_2 cloud, rather than capturing its entire extent.

4.3.2 Nowcasting volcanic ash clouds using Deep Learning

The proposed model integrates three-dimensional U-Net and ConvLSTM modules, aiming to capture spatial, spectral, and temporal dependencies simultaneously. The primary goal is to predict the spatial and compositional evolution of volcanic clouds, distinguishing ash, SO_2, and mixed plumes, up to 60 minutes ahead, with inference times compatible with operational use. By combining the strengths of geostationary satellite observations with advanced neural architectures, this work seeks to provide a scalable and near real-time framework to support aviation safety, civil protection, and environmental monitoring.

The algorithms presented in the following sections share a common initial segmentation pipeline. Multispectral SEVIRI images are first converted into binary masks that isolate the volcanic cloud from the surrounding atmospheric background. These masks represent the foundation of the forecasting workflow, upon which all subsequent models are built. The 3D models proved to be well suited for capturing both the spatio-temporal dynamics of the volcanic cloud and the spectral information related to its chemical composition, showing qualitative improvements in the separation between SO_2 and ash compared with binary models.

However, the increased complexity of the data and the limited availability of eruptive sequences resulted in a higher risk of overfitting, highlighting the need for additional regularization strategies and data augmentation techniques to fully exploit the potential of multispectral datasets. After analyzing in detail, the strengths and limitations of the previous architectures, the following section introduces a hybrid U-Net3D–ConvLSTM3D model that synergistically integrates spatial and spectral feature extraction with explicit temporal dynamics modeling. The goal is to overcome the critical issues identified in the earlier sections and to provide a more accurate and consistent forecast of the compositional temporal evolution of volcanic clouds.

Dataset and Input preparation

The dataset used come from [126] has already developed the segmentation procedure for SEVIRI images. It includes 41 eruption sequences acquired between January 2021 and December 2024, each recorded at a temporal resolution of 15 minutes by the SEVIRI instrument aboard the Meteosat satellites. For each frame, on [126], the algorithm produces a binary mask that reliably separates the volcanic cloud from the background by exploiting SEVIRI's thermal channels. These binary masks serve as the starting point for all predictive models described in this work.

Data preparation was structured in three main steps. First, each segmented mask was resized to 100×100 pixels (single channel) and temporally ordered, producing tensors of the form (T, 100, 100, 1). During the exploratory phase, variable-length temporal windows were tested, extracting samples consisting of 2–5 input frames followed by 4 target frames (corresponding to a 60-minute forecast). However, this heterogeneity introduced instability in training, as mini-batches became non-homogeneous. To ensure stability, a fixed windowing strategy was adopted: 4 input frames and 4 output frames, extracted with stride 1. The resulting samples were aggregated into a single dataset, randomly shuffled at the beginning of each epoch, and provided to the models in batches of size 4.

Training Setup - The training pipeline was implemented in PyTorch and designed for models outputting three-dimensional volumes of shape (B, 1, T, H, W). This formulation reduces memory usage and accelerates computation on tensor cores.
Loss Function and Metrics The optimization metric is defined as a composite loss that balances a mean squared error (MSE) term with a perceptual loss (LPIPS). Given the prediction tensor $\hat{Y}$ and the target Y, the loss is formulated as:

$$L_{mix} = \alpha\, MSE(\hat{Y}, Y) + (1-\alpha)\,\frac{1}{T}\sum_{t=1}^{T} LPIPS(\hat{Y}_t, Y_t), \quad \alpha = 0.7$$

The LPIPS contribution is computed frame-wise and subsequently averaged over the temporal dimension, thereby preserving both photometric and perceptual consistency. The adoption of this hybrid loss is motivated by the need to combine the numerical precision provided by MSE with the perceptual quality ensured by LPIPS, thus producing predictions that are not only close to the targets in a Euclidean sense but also visually realistic and free from typical L2-related artifacts, such as loss of fine details and the tendency to generate overly smoothed images.
Optimizer. Network weights were updated using the Adam optimizer with an initial learning rate of 1e-3. No learning rate scheduler was applied in this preliminary phase, in order to preserve a controlled and interpretable optimization dynamic.
Training and Validation Procedure. Training proceeded for 120 epochs. At each epoch, all mini-batches (size 4) underwent forward propagation, loss computation, gradient backpropagation, and parameter updates. The dataset was split into training (70%), validation (15%), and test (15%) subsets.
Validation was carried out at the end of each epoch without weight updates, and the model corresponding to the lowest validation loss was retained as the final checkpoint. Early stopping was not employed.
Monitoring. Training and validation performance were tracked by plotting Dice Loss and balanced accuracy across epochs, enabling the detection of underfitting or overfitting trends and providing guidance for future improvements.

Fase	**Dettagli**
Forward pass	Encoder-Decoder 3D $\rightarrow$ $\text{Conv}_{1\times1\times1}$
Loss	$\mathcal{L}_{\text{mix}} = 0.7\,\text{MSE} + 0.3\,\text{LPIPS}$
Back-prop	Gradiente su tutti i parametri
Optimizer	Adam; LR adattato da `ReduceLROnPlateau`
Stop	Early stopping se Δval_loss $< 10^{-4}$ per 80 epoche
Checkpoints	`best`, `last` e uno ogni 10 epoche

FIGURE 4.3: Summary of the training configuration adopted in the proposed model. The table details the main steps of the training pipeline, including forward pass, loss formulation, optimization strategy, stopping criterion, and checkpointing procedure.

The final dataset comprises 91 eruptions recorded between 2007 and 2024. Each folder contains the complete ashRGB frame sequence corresponding to a single event. The events were manually labeled into four operational classes

Classe	Significato	Eruzioni	%
1	Strong Plume	34	37.4
2	Weak Plume	10	11.0
3	Plume + Meteo	38	41.8
4	Mixed types	9	9.8

FIGURE 4.4: Distribution of the events for each class.

based on the intensity and morphology of the cloud (Figure 4.4). The adopted taxonomy reflects the variety of phenomena observed at Mount Etna and is motivated by both physical and operational considerations.
Class 1 – Strong Plume: includes plumes generated by high-energy explosive episodes capable of producing well-developed convective columns that rise above the summit and expand into a large cloud of ash and gas.
Class 2 – Weak Plume: groups together short-lived or low-intensity emissions, often composed almost exclusively of fine ash or sulfur dioxide. These clouds are thinner, intermittent, and rapidly dispersed by winds, with less-defined textures and lower spectral contrast. To harmonize the duration of the clips with the nowcasting model specifications, each raw sequence is subdivided into partially overlapping temporal windows of eight frames, where the first four serve as input and the subsequent four constitute the target.
Class 3 – Plume + Meteo: identifies events in which the volcanic plume coexists or interacts with meteorological clouds, complicating the radiometric separation between the volcanic component and the cloud cover.
Class 4 – Mixed types: gathers situations that do not fall into the previous categories, such as sequences with two consecutive explosive emissions or mixed transitional phases. This class serves as an open category to preserve dataset variability without forcing arbitrary assignments.
The class subdivision is used to ensure that the training, validation, and test sets each receive a proportional share of every eruption type, thereby reducing the risk of sampling bias. The chosen sliding step corresponds to a single frame, maximizing the number of samples without introducing excessive redundancy. All generated windows have a uniform size of (4, 100, 100, 3) for the input portion and (4, 100, 100, 3) for the target portion, ensuring compatibility with sequential forecasting models. Each sample retains the class label associated with its eruption of origin, preserving the ability to balance these classes across splits in subsequent stages.
Overall, the subdivision process yields a total of 964 clips across the dataset, a value dependent on the original duration of each sequence. This structured dataset forms the basis for the comparative experiments described in the following sections. Once the clips were generated, the dataset was divided into three subsets (train, validation, and test) according to the class distribution defined in Figure 4.4. The split follows a 70% / 15% / 15% proportion for each class, ensuring balanced representation even in the less populated categories. Each split is associated with a DataLoader configured with a batch size of four. To enhance batch diversity and reduce the risk of overfitting, the

training set is shuffled at each epoch (shuffle=True), while the validation and test sets maintain a fixed order (shuffle=False) to ensure repeatability and consistency of measurements. The entire process thus guarantees a balanced, normalized, and fixed-length clip dataset, which represents an essential condition for the comparative evaluations presented in the subsequent sections.

Architectures

The *Deep 3D U-Net* RGB, is an extension of the standard U-Net to multichannel volumetric sequences while introducing greater depth. The encoder is organized into four levels of spatial compression through three-dimensional convolutions with stride applied only along the spatial dimensions, combined with repeated blocks of convolution, normalization, and activation. The central bottleneck processes the most compact representation before the decoding stage. The decoder progressively restores the original resolution using three-dimensional transposed convolutions and convolutional blocks, recovering the initial spatial dimensions. A key feature of this model is the pre without skip-connections that concatenate the feature maps from the encoding path with those from the decoding path at each level, thereby improving reconstruction quality and preserving both spatial and spectral details. Furthermore, the network employs targeted padding to ensure dimensional compatibility of the feature maps before concatenation. The fully volumetric and multichannel design has proven effective in maintaining the coherence of the chemical components and in separating SO from Ash throughout the temporal sequence; however, the temporal evolution of the cloud remains partially inaccurate, suggesting that the network has a limited ability to fully model complex temporal dynamics.

Subsequently, an Encoder–Decoder *3D-ConvLSTM* was implemented, a three-dimensional architecture capable of explicitly modeling spatio-temporal and spectral dependencies. The network is composed of an encoder built on multiple hierarchical layers of 3D-ConvLSTM cells, each equipped with normalization and volumetric dropout to mitigate overfitting. The encoder processes the volumetric sequence (B, 3, T, H, W) and stores the cloud dynamics in its internal three-dimensional states. The symmetric decoder uses the encoder's final states to generate the predicted sequence, through an additional stack of 3D-ConvLSTM cells followed by a final convolution that outputs a three-channel volumetric sequence consistent with the expected chemical composition. Compared to previous versions, the network explicitly incorporates batch normalization and dropout at the LSTM gate level, improving both stability and generalization capacity. The absence of constraints on the final activation allows the model to produce unconstrained outputs, which can later be binarized or subjected to further transformations. This architecture combines the explicit memory of LSTMs with the ability of three-dimensional convolutions to capture volumetric and multichannel context, making it particularly well-suited for scenarios with complex temporal dynamics. From a qualitative comparison, the model successfully reproduces the structure and direction of the volcanic cloud, correctly placing SO and Ash. However, the predictions appear blurrier: contours are smoothed, peaks attenuated, and fine details vanish after a few steps. This indicates a tendency toward over-smoothing of high-frequency patterns and an underestimation of rapid local variations. The Figure 4.5 shows the workflow proposed. In particular, the 3D U-Net employs residual blocks to improve training stability and facilitate effective gradient propagation. The encoder is organized into five hierarchical

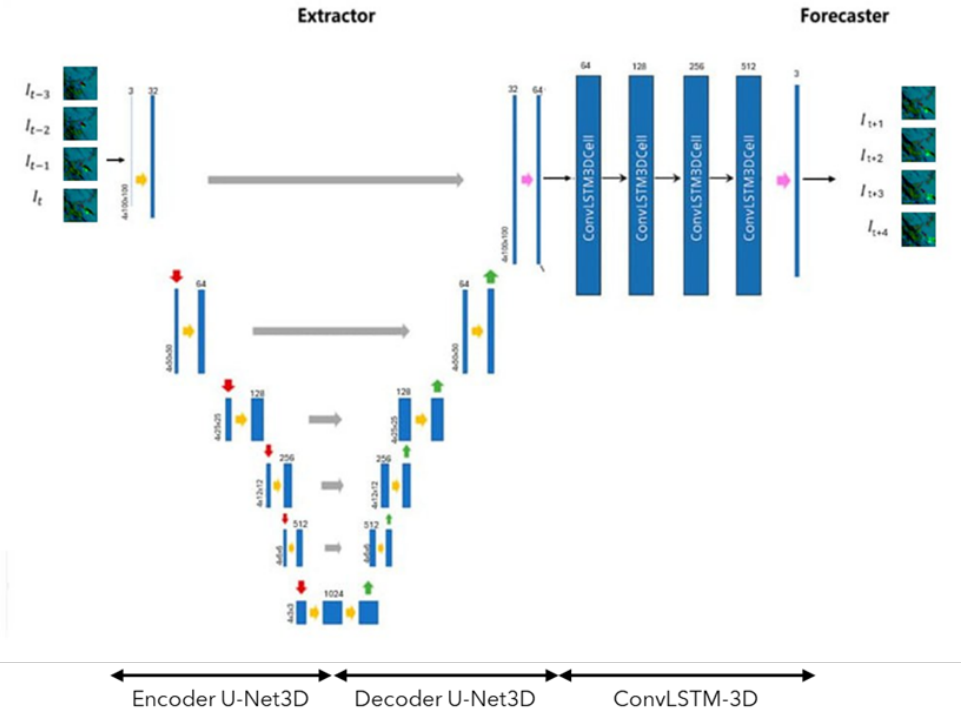

FIGURE 4.5: Workflow proposed

levels with three-dimensional convolutions, alternated with spatial downsampling operations via max pooling. Each level uses a ResidualBlock3D that processes spatial and temporal information while progressively increasing the number of channels. Each residual block consists of two 3D convolutions, each followed by batch normalization and ReLU activation, as well as skip connections that preserve the original signal representations, thereby enhancing the model's ability to learn complex structures. The central bottleneck effectively synthesizes spatio-temporal and spectral information. The decoder restores the original resolution using three-dimensional transposed convolutions, integrated with skip connections from the corresponding encoder levels.
The final convolution of the decoder produces a rich intermediate representation, which is then used as input for the subsequent stage of the model. Indeed, the temporal component of the architecture is handled by a 3D ConvLSTM module, composed of a hierarchical sequence of ConvLSTM-3D cells. Each cell employs three-dimensional convolutions within the update gates of the hidden states and internal memory, enabling effective modeling of temporal dynamics. The 3D ConvLSTM module receives as input the volumetric features extracted by the 3D U-Net (B, 64, T, 100, 100) and progressively modulates the representation across four levels.
The combined use of 3D U-Net and 3D ConvLSTM, rather than their individual application, is motivated by the need to simultaneously manage the spatial, spectral, and temporal complexity of ashRGB sequences. The 3D U-Net can identify and preserve fine spatial and spectral details, but it cannot independently capture the temporal dependencies that govern the evolution of volcanic clouds. Conversely, the 3D ConvLSTM is specifically designed to

model complex temporal dynamics through three-dimensional recurrent memories, yet it exhibits limitations in extracting and retaining high-resolution spatial and spectral details. The synergistic integration of the two models overcomes the intrinsic limitations of each approach: the 3D U-Net provides a detailed representation of spatial and spectral features, while the 3D ConvLSTM effectively models their temporal evolution. The optimal combination of these modules thus achieves superior predictive performance compared to the individual use of either architecture, making it particularly well-suited for accurate forecasting of multispectral volcanic cloud sequences.

4.4 Forecasting

Unlike the nowcasting approaches presented in other sections, the forecasting component discussed here does not yet rely on a consolidated methodology or quantitative results. Instead, it is conceived as a research perspective, outlining potential directions for future investigation.
A particularly interesting avenue concerns the role of volcanic degassing, both pre-eruptive and persistent background emissions, as a driver of atmospheric and climatic impacts. While it is well established that major explosive eruptions represent strong climate forcings, injecting vast amounts of SO_2 and aerosols into the stratosphere, much less attention has been devoted to the cumulative effect of continuous degassing from multiple volcanoes worldwide. This includes both emissions that occur prior to eruptions, which may precede explosive phases, and emissions associated with non-eruptive or quiescent activity. Such contributions, although smaller on an event-by-event scale, are far from negligible when considered in aggregate over time. Large persistently degassing systems, such as Kīlauea in Hawaii, which emits several kilotons of SO_2 per day, or Mount Etna in Italy, represent important test cases. Incorporating their emissions, together with those from other active volcanoes, into climate impact assessments could provide new insights into how the background flux of volcanic SO_2 contributes to atmospheric composition and potentially modulates radiative forcing. Future work in this direction could aim to build datasets that quantify these continuous and pre-eruptive emissions, evaluate their cumulative effect across different regions, and investigate how they interact with larger eruptive injections. Such efforts would significantly expand the current framework of volcanic climate impact studies, which has traditionally focused on singular major events, by integrating the role of persistent degassing as a continuous volcanic forcing on the atmosphere.

4.5 Detecting

In this section, tha applicability of the zero-shot methodology will be presented. After about three months of quiescence, during which Etna volcano exhibited ordinary passive degassing from its summit craters, a new effusive eruption began on 8 February 2025. The event was marked by the opening of an eruptive fissure at approximately 3,050 meters above sea level, located between the Bocca Nuova (BN) and Southeast (SEC) craters. Unlike other volcanoes analyzed in this study, such as HTHH and Raikoke, where the SO_2 cloud dispersed over thousands of kilometers from the main source, during this effusive activity Etna produced a continuous and relatively localized SO_2

emission, confined mainly to the area surrounding Sicily. This detail is particularly relevant, as it allows us to focus on a single satellite orbit and use the SO_2 column density from the TROPOMI product directly as input to our model. As in some of the other case studies presented later, certain situations

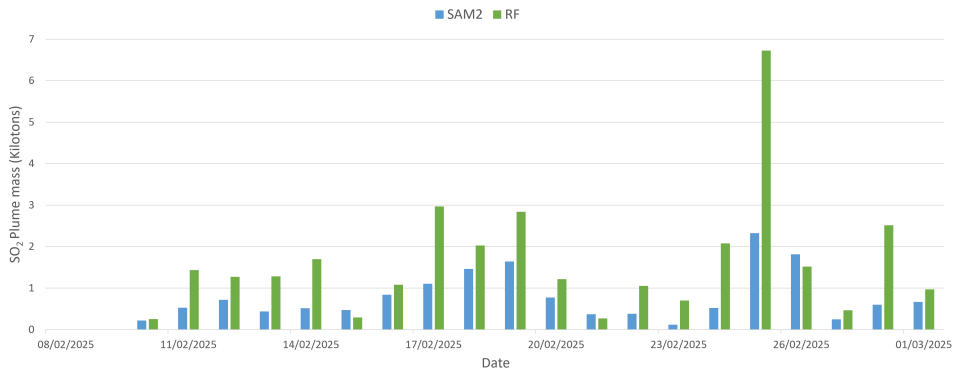

FIGURE 4.6: Time series of SO_2 plume mass (in kilotons) retrieved by the SAM2 (blue) and Random Forest – RF (green) algorithms between 8 February and 1 March 2025.

here also allowed us to test the model in both prompt and no-prompt modes, ensuring that the detection and subsequent segmentation of the SO_2 cloud in the atmosphere were robust under different input conditions.

To evaluate our model's performance, we compared its results with those obtained using the RF model of [69]. Figure 4.6 illustrates the temporal evolution of SO_2 emissions during the February - March 2025 eruptive sequence. Both algorithms show a coherent pattern, characterized by a gradual increase in plume mass starting around 11 February and culminating in a clear peak on 25 February 2025, followed by a sharp decrease in the subsequent days. Overall, the SO_2 masses estimated by the Random Forest model are systematically higher than those retrieved by SAM2, particularly during the most active eruptive phases. This discrepancy primarily arises from the methodological differences between the two approaches. SAM2 relies on the total SO_2 column density computed assuming a fixed plume height of 7 km, whereas the RF algorithm, following [69], uses the Cloud Top Height (CTH) to infer the actual plume altitude and then interpolates linearly between the SO_2 masses calculated at 1 km and 15 km.

The tendency of the RF retrievals to yield higher SO_2 masses reflects the fact that, in many cases, the CTH-derived plume altitudes are lower than the fixed 7 km assumed in SAM2. A lower plume altitude corresponds to a longer optical path and thus to a higher apparent SO_2 column, explaining the systematic positive bias in the RF results.

Despite these absolute differences, both methods exhibit strong temporal consistency, capturing synchronous variations in the SO_2 signal and reproducing similar trends in plume evolution. The 25 February 2025 episode represents the main emission peak, with approximately 7 kilotons of SO_2 estimated by RF and about 2.5 kilotons retrieved by SAM2.

Two representative cases were selected, 17 February and 25 February 2025, which in the previous figures exhibited the largest discrepancies in SO_2 quantification. The purpose of this analysis is to closely examine the input images

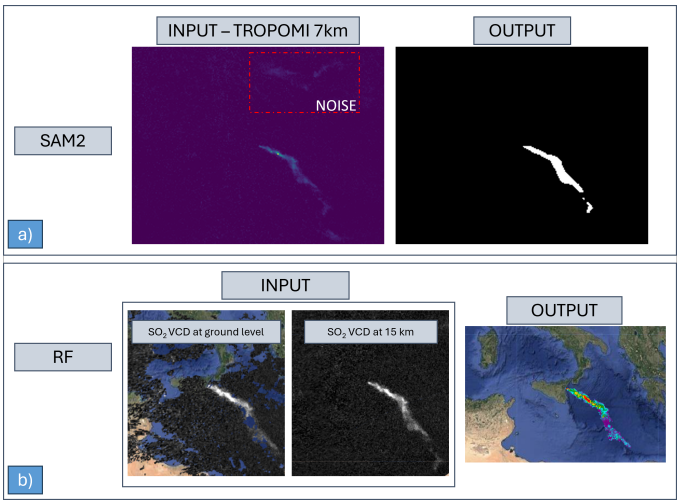

FIGURE 4.7: Comparison between the SAM2 a) and RF b) approaches for SO_2 plume detection on 17 February 2025

provided to the model and to assess whether the algorithm is able to accurately identify the plume area while avoiding the inclusion of noise or spurious signals in the surrounding regions. Specifically, our SAM 2 implementation uses the TROPOMI product at 7 km as input, whereas the RF method of [69] uses GEE-processed and interpolated products at 1 km and 15 km, combined with six GLCM texture bands, and performs interpolation across these datasets. Such differences in spatial resolution and data processing can lead to significant variations in the quantified SO_2 mass. In addition, in this case the average plume altitude, derived from cloud_top_height, is around 3 km, which further contributes to the underestimation by SAM 2, given its fixed assumption of a 7 km plume height.

The first case, observed on 17 February 2025, corresponds to an SO_2 mass of 1.106 kton estimated by the SAM2 model and 2.97 kton produced by the RF algorithm. In terms of SO_2 detection, the SAM2 product successfully discriminates the plume even in the presence of noise, clearly visible in the input image of Figure 4.7a (top right). The model is able to avoid misclassifying noise or acquisition artifacts as SO_2, correctly focusing instead on the plume originating from the volcanic source. Although the plume in this case appears relatively weak, hence represented with a moderate color intensity, the SAM2 algorithm effectively captures its full spatial extent, resulting in an accurate and noise-free output, as shown in the right panel of the figure.
In Figure 4.7b, the corresponding input and output used by the RF model are displayed, which form the basis for the SO_2 quantification discussed previously. This visual comparison highlights the different input structure of the RF approach and illustrates how the model reconstructs the plume distribution using the interpolated information between the 1 km and 15 km SO_2 vertical column densities. Furthermore, an additional example of segmentation with a fixed threshold is presented for the same date, in order to highlight the intrinsic noise within the image and to show how the segmented region varies as a function of the chosen threshold value Figure 4.8. The same mechanism explains the even larger discrepancy observed on 25 February 2025, when SAM 2 estimates 2.324 kton compared to 6.725 kton reported by RF.

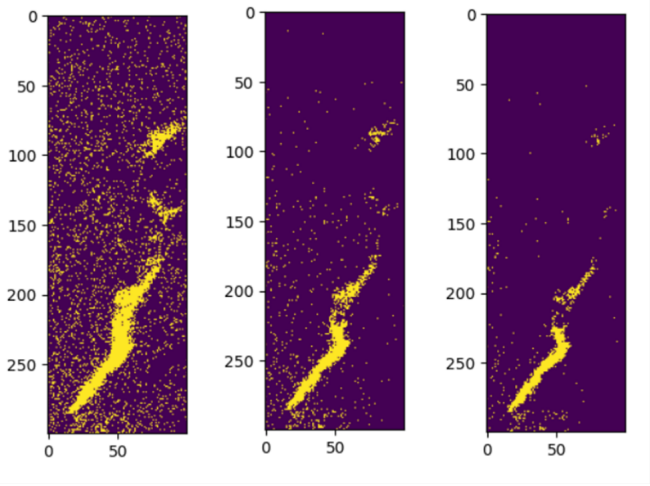

FIGURE 4.8: Fixed threshold – 17 February 2025: 1 DU, 2.25 DU, and 3 DU, corresponding to 1.61 kton, 0.98 kton, and 0.76 kton, respectively.

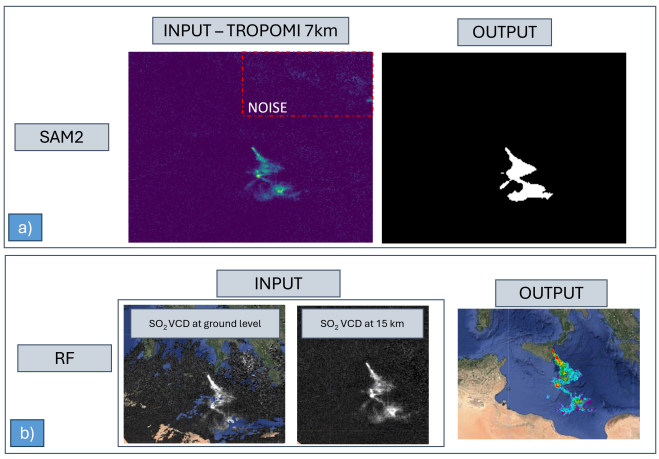

FIGURE 4.9: Comparison between the SAM2 a) and Random Forest b) approaches for SO_2 plume detection on 25 February 2025

The second case, as shown in Figure 4.9, observed on 25 February 2025, corresponds to the day of maximum SO_2 emission during the analyzed period, with 2.49 kton estimated by SAM2 and 6.99 kton retrieved by the RF model. As shown in Figure 4.9a, the SAM2 product, which assumes a fixed plume altitude of 7 km, effectively captures the main SO_2 structure near the volcanic source while avoiding the influence of noise present in the northern portion of the input image. Despite the presence of multiple high-intensity regions in the input field, the SAM2 segmentation (right panel) correctly isolates the volcanic plume, producing a clean and spatially compact output. Figure 4.9b illustrates the inputs and output of the RF model. The resulting output (right panel) shows a broader and more continuous plume that extends southeastward, reflecting the transport and dispersion of the volcanic emission at multiple altitudes. An additional case study is shown in Figure 4.10, covering the period from 2 to 30 June 2025. During this interval, several eruptive episodes occurred: on 2 June a lava fountain accompanied by a pyroclastic flow from the Southeast Crater, and on 19–20 June intense

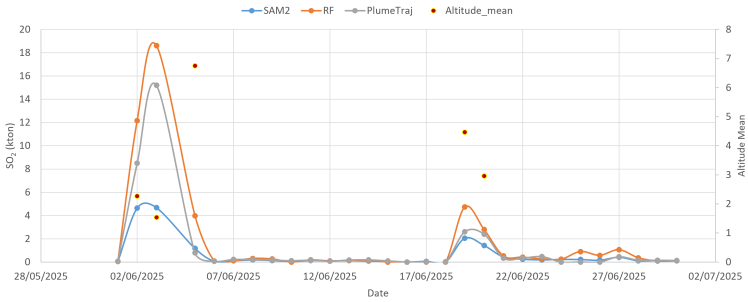

FIGURE 4.10: Time series of SO_2 plume mass (in kilotons, left axis) and mean plume altitude (right axis) from 2 June to 30 June 2025. The plot includes SO_2 mass estimates retrieved from SAM2 (blue), RF (orange), and PlumeTraj (gray), together with the mean plume altitude provided from CTH (yellow markers).

Strombolian activity at the same crater with ash emission, along with variable degassing activity at Bocca Nuova, Voragine, and Northeast Craters. The comparison was carried out between the outputs of SAM 2, Random Forest, and PlumeTraj [231], [232], while also including information on the average plume altitude (from CTH) in the figure. The analysis shows that the plume altitude generally remains confined to values close to 1 km, and it is precisely in these cases that the largest discrepancies between the different methods are observed. Conversely, when the altitude approaches 7 km, the SAM 2 estimates are consistent with those from the other approaches. This confirms that the eruptive episodes considered were characterized by very low plume heights, and that the assumption of a fixed 7 km height systematically leads to an underestimation of the SO_2 mass quantified by SAM 2. The figure highlights the clear dependence of retrieved SO_2 mass on the altitude of the volcanic plume. When the plume altitude is significantly lower than 7 km, the SO_2 mass retrieved by the product fixed at 7 km (SAM2) is systematically underestimated compared to the other two methods that account for the actual plume height. This behavior is particularly evident during the first eruptive phase (2–3 June 2025), when the mean plume altitude ranges between 4 and 6 km. In this period, the SAM2 values remain distinctly lower, while the other datasets capture a larger SO_2 mass, reflecting the enhanced optical path and higher apparent column at lower altitudes. Conversely, when the plume height approaches around 7 km, as observed on 5, 19 June 2025, the differences among the three retrievals become minimal. In these cases, the SO_2 mass segmented and subsequently quantified by the fixed-7 km product is fully consistent with the estimates obtained from the other two methods. Overall, the comparison demonstrates that the accuracy of SO_2 quantification strongly depends on the correct representation of plume height. Underestimation by fixed-altitude retrievals occurs primarily when the volcanic cloud is confined to lower atmospheric layers, while consistency across methods is achieved when the actual plume altitude aligns with the assumed value of 7 km. This result reinforces the importance of incorporating realistic altitude information when interpreting satellite-based SO_2 measurements, particularly during variable eruptive activity.

4.6 Tracking

Although the SO_2 emission from the Hunga Tonga–Hunga Haapai eruption was negligible compared to major historic events such as Pinatubo, Tambora, in the modern era it stands out as one of the largest eruptions, with a Volcanic Explosivity Index (VEI) of approximately 5–6. Nevertheless, the amount of SO_2 injected was relatively modest, with estimates of about 0.5 Tg. This eruption remains of great scientific interest due to the unprecedented quantity of H_2O injected into the stratosphere, which is thought to have acted as a form of natural volcanic climate forcing, potentially influencing weather and climate on both short and long timescales, locally and globally. These values highlight how the largest fraction of SO_2 was released in a short, intense phase, followed by a gradual dispersion of the cloud. Analogously to the Etna case, where the retrieved values were analyzed in relation to plume altitude, it is even more critical to make this consideration for the Hunga Tonga–Hunga Ha'apai (HTHH) eruption. The plume height reached during this event far exceeded the range captured by cloud_top_height products: the eruption column of 15 January 2022 rose well into the stratosphere and beyond.

Most operational CTH algorithms are tuned for the troposphere and tend to saturate (or become invalid) above 16–18 km; moreover, stratospheric temperature inversions challenge IR brightness temperature–based methods. As a result, a significant discrepancy between estimated and actual plume heights is expected. Indeed, plume altitudes were reported to be 15–20 km on 14 January, 25–30 km on 15 January, and 25 km on 16 January. [222] provide a detailed table of SO_2 emission values for these three days derived from OMPS and TROPOMI, adopting for convenience a fixed plume height of 15 km for TROPOMI. Although, as discussed, the use of data at 7 km inevitably leads to an underestimation of the amount of SO_2 emitted, the focus here is on tracking, following the movement and displacement of the SO_2 cloud from 14 to 17 January 2022. As already noted, the cloud propagated for more than 6000 km westward from the HTHH volcano. Figure 4.11 shows the sequence of input images provided to the SAM2 model and the corresponding output segmentation masks. A first difference that can be highlighted, when comparing the HTHH case to that of Etna, is the spatial consistency and continuity of SO_2 emissions, which in the case of HTHH is markedly different. The SO_2 cloud generated by HTHH represents a higher level of complexity that SAM2, when applied in zero-shot mode, is not always able to fully capture. On 14 January 2022, the SO_2 cloud appears localized and compact, and the model is able to segment it correctly. However, in the following days, as the SO_2 plume disperses into the atmosphere, the cloud loses continuity and becomes increasingly fragmented. In these conditions, SAM2 has greater difficulty in capturing the full spatial extent of the emissions. For example, on 17 January, part of the SO_2 cloud at the outer edge is not segmented, since the model does not recognize it as belonging to the same structure relative to the prompt point provided.

From a tracking perspective, however, the overall displacement of the SO_2 cloud is still captured: the model allows us to follow its westward movement from the eruption site, thereby reconstructing its large-scale spatial evolution. It is important to note that SAM2 is a model originally trained on RGB natural images, and it has never been exposed to data of this type. This strongly

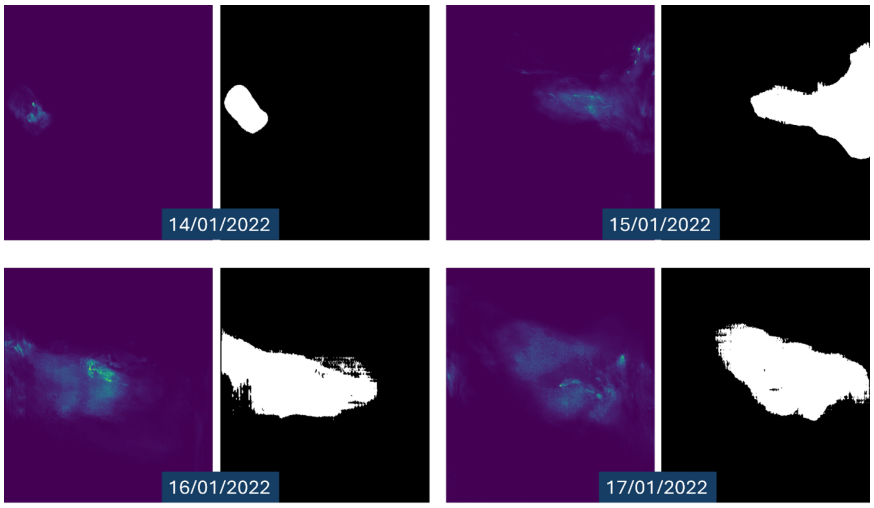

FIGURE 4.11: Example of SO_2 plume detection from the Hunga Tonga–Hunga Ha'apai eruption between 14 and 17 January 2022, applying SAM2 zero-shot modality.

influences its behavior: for large-scale SO_2 emissions that disperse spatially and temporally in the atmosphere, using a single-point prompt corresponding to the maximum intensity often does not provide enough information to guide SAM2 toward a complete and coherent segmentation. In the future, SAM2 could be further tested in zero-shot mode by exploring different types of prompts, such as bounding boxes, which may help to capture portions of the SO_2 cloud that are otherwise missed due to the lack of spatial continuity.

4.7 Quantifying

In the context of quantification, an important research direction concerns the development of advanced methodologies to better evaluate the climatic impact of volcanic eruptions by integrating multi-source satellite data with artificial intelligence techniques. While the focus of many existing studies has been on the radiative forcing associated with single major explosive events, a future line of investigation should also account for the cumulative role of continuous volcanic degassing and the detailed quantification of gas injections during both eruptive and pre-eruptive phases.
Emerging approaches, such as the application of super-resolution (SR) models to satellite datasets (e.g., Sentinel-5P TROPOMI, MTG FCI), could significantly enhance the spatial and temporal fidelity of SO_2 and other gas observations, allowing for more precise quantification of emission sources and their atmospheric dispersion. By reconstructing high-resolution emission fields, even under partial or low-quality observational conditions, SR methods would improve the derivation of climate-relevant parameters such as Stratospheric Aerosol Optical Depth (SAOD), plume injection altitude, and radiative forcing proxies. In addition to single-event analysis, quantification should also consider persistent emitters, such as Etna or Kīlauea, which release large amounts of SO_2 on a daily basis. Integrating these emissions into long-term observational frameworks would make it possible to assess how continuous

volcanic degassing interacts with major eruptions to shape atmospheric composition and climate forcing. This line of research, although still conceptual, represents a natural extension of current hazard-oriented monitoring frameworks toward a more comprehensive evaluation of volcanic impacts on climate.

4.8 Nowcasting

This section contains some results related to nowcasting of volcanic ash clouds. The analysis of the combined loss trend during training represents a crucial indicator of both the quality of the optimization procedure and the model's ability to generalize on the test data. Figure 4.12 illustrates the evolution of

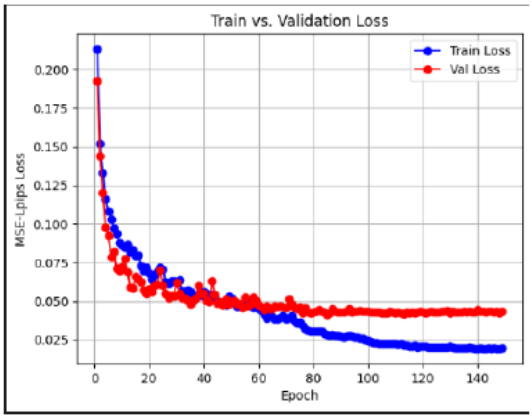

FIGURE 4.12: Training and validation loss curves over 150 epochs. The plot shows the convergence behavior of the model using the combined MSE + LPIPS loss function. Both training and validation losses decrease rapidly in the initial epochs and then stabilize, with validation loss closely following the training curve, indicating good generalization and absence of overfitting.

the loss on the training and validation sets, showing an initial rapid decrease followed by a gradual flattening, a behavior typical of well-regularized deep models. Notably, the validation curve never exhibits a sharp inversion, indicating that overfitting remains limited and that the balance between MSE and LPIPS, together with the chosen learning-rate scheduling, resulted in a stable optimization dynamic. In particular, the curve profile does not follow a purely linear decay but rather displays three distinct regimes, corresponding to the learning-rate reductions and the different aspects of the signal that the model progressively learns to reproduce.

Phase 1 – Rapid descent (0–20 epochs). The model learns the macroscopic geometry of the cloud, with the aggregated loss dropping from 0.21 to 0.06, driven primarily by the MSE component, which dominates over the perceptual term.

Phase 2 – Slow refinement (20–70 epochs). After the first plateau, the learning rate is halved, the descent slows down, and the two curves converge, reflecting the stabilization of the main structures and the onset of textural learning, particularly at the plume–meteorological cloud boundaries.

Phase 3 – Fine-tuning ($\geq 70 epochs$). With a learning rate $\leq 5\times10^{-4}$ the training loss decreases to 0.019, while the validation loss stabilizes at 0.042–0.043,

with an average spread ≈ 0.022. This indicates mild but controlled overfitting. This persistent gap between the curves, though non-negligible, is typical of scenarios in which the validation set distribution is more heterogeneous than that of the training set. The stability of the validation curve confirms that the model does not degenerate and retains good generalization capability, even as optimization continues beyond the validation minimum.
For each case, the four input frames, the four ground-truth targets, and the four model predictions for the subsequent time steps are shown. The analysis allows for a visual assessment of the physical plausibility of the forecasts, particularly regarding the network's ability to preserve the spatial coherence of the volcanic cloud, its expansion direction, and the radiometric distribution, even though a slight attenuation of fine textures can be observed in the more distant frames.

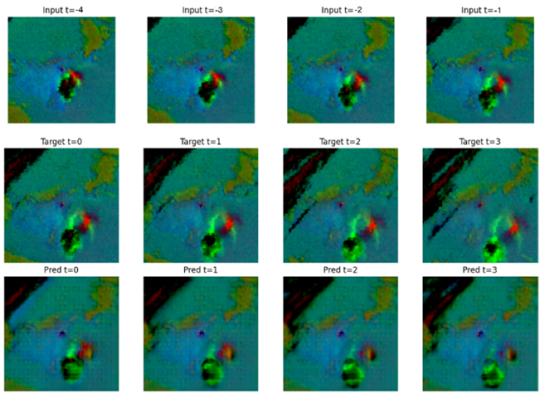

FIGURE 4.13: Cloud dispersion prediction example. The first row shows the input sequence at times t=-4 to t=-1. The second row illustrates the ground truth target frames from t=0 t=3. The third row presents the corresponding model predictions, highlighting its ability to reproduce both spatial distribution and temporal evolution of cloud dynamics.

In Sequence 1 (Figure 4.13), the volcanic cloud moves toward the southwest, an unusual direction for this dataset since most paroxysmal eruptions of Mount Etna typically propagate eastward. Moreover, meteorological clouds partially interfere with the plume. Despite these atypical conditions and the overall complexity of the scene, the model's prediction remains consistent with the morphology and radiometric distribution observed in the targets.
In Sequence 2 (Figure 4.14), the summit craters of Etna are partially covered by meteorological clouds, making it more challenging to separate the volcanic plume from the atmospheric background. Even under these complex conditions, the model successfully tracks the plume's evolution, maintaining spatial coherence and a reasonable separation between the different components. In Sequence 3 (Figure 4.15), the scene occurs under optimal meteorological conditions, without cloud cover interference. In this case, the prediction maintains strong spatial coherence, with stable plume contours. Furthermore, the model correctly predicts the gradual fading of the ash component within the cloud in the later frames, in agreement with the target. Overall, the model demonstrates a solid predictive capability even under unfavorable meteorological conditions or in anomalous configurations relative to the dataset. Spatial coherence is preserved in all cases, and the loss of detail

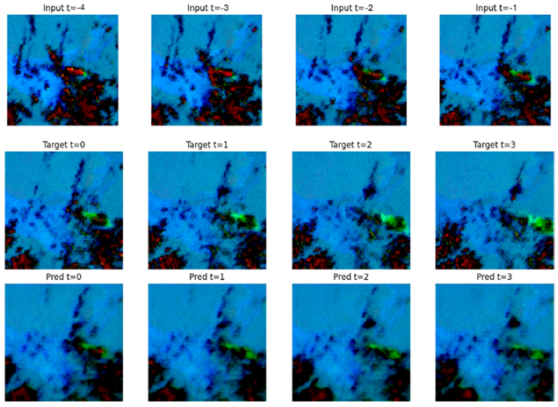

FIGURE 4.14: Meteorological clouds overlapping the summit craters.

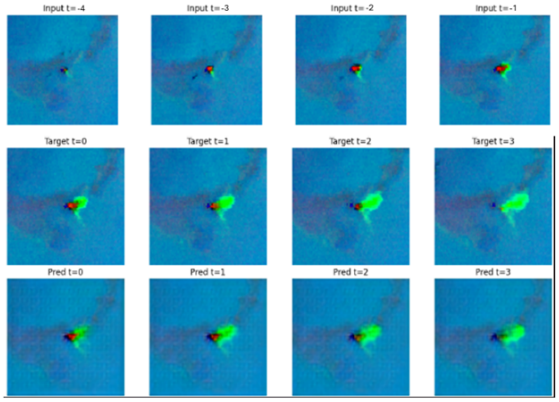

Figura 4.4: Sequenza 3: condizioni meteorologiche ottimali e plume compatto.

FIGURE 4.15: Optimal meteorological conditions and a compact plume.

remains limited and consistent with the nature of the problem. However, a slight qualitative degradation is observed in the final frames (t = 2–3), with increased blurring and reduced sharpness of the plume boundaries.

This research line has pursued the goal of addressing some of the open challenges in volcanic cloud nowcasting, with a particular focus on Etna plumes observed by the SEVIRI sensor. Starting from the observation that physical, numerical models suffer from operational latency and that threshold-based algorithms do not provide temporal forecasts, a hybrid architecture, U-Net3D + ConvLSTM-3D, was designed to simultaneously process the spatial, spectral, and temporal information contained in Ash RGB representations. The results demonstrate that a deep data-driven approach, when trained on carefully curated datasets, can reliably anticipate the evolution of volcanic clouds within an operational horizon of one hour, providing spatial resolution and response times superior to traditional methods. However, the proposed solution also presents structural limitations.

These include reduced generalization due to training being focused exclusively on Etna, loss of detail for lead times beyond one hour, absence of uncertainty quantification, spectral sensitivity to illumination conditions, and the lack of physical or dynamical constraints, which may reduce accuracy under extreme

scenarios. To address these limitations, several directions for future research can be outlined. Among them are transfer learning and domain adaptation, to adapt the model to volcanoes with different morphology and atmospheric conditions; the integration of attention-based architectures, such as spatiotemporal Transformers, to extend the forecast horizon beyond one hour; multi-source fusion and data assimilation, incorporating wind fields from meteorological models and polar-orbiting satellite observations into physics-informed frameworks.
In conclusion, the findings of this work represent a concrete step toward rapid and reliable alert systems for volcanic eruption monitoring. While further developments are needed to ensure physical robustness, global generalization, and integration into international operational frameworks, the proposed approach clearly demonstrates the potential of deep learning to bridge the gaps of traditional methods and to open new perspectives for the observation and forecasting of volcanic clouds.

4.9 Analysing the Climatic Response to Volcanic Forcing

This section explores how the AI-based volcanic monitoring framework can be extended to investigate the climatic impacts of eruptions. While previous sections focused on the detection, tracking, and quantification of volcanic phenomena, here the emphasis shifts to the analysis of how such events influence the atmosphere. Two complementary methodologies are presented: one based on reanalysis data (ERA5) and the other on statistical emulation techniques. Both aim to characterize the same process, the atmospheric and climatic response to volcanic forcing, but differ in terms of data sources and computational strategy. The first approach, based on the ERA5 reanalysis, provides an observation, driven reconstruction of the atmospheric state, combining satellite and ground-based data through data assimilation. It enables the examination of large-scale temperature, pressure, and radiative anomalies linked to major eruptions, serving as a diagnostic tool to validate also the AI-based detection and quantification results presented earlier. The second approach employs emulator models, which reproduce the behavior of complex climate simulations at a fraction of the computational cost. By learning statistical relationships between volcanic forcing parameters and climate model outputs, emulators offer a rapid and flexible means of exploring volcanic–climate interactions. This methodology complements the reanalysis approach by providing predictive capabilities and allowing fast scenario testing, especially when running full-scale climate models would be computationally prohibitive.
Together, these two approaches extend the forecasting-to-nowcasting workflow into the climate domain. Their combined use supports a deeper understanding of how volcanic eruptions perturb the atmosphere, contributing both to rapid assessment of past events and to forward-looking predictive analyses.

4.9.1 Reanalysis-Based Assessment: The 2022 Hunga Tonga Eruption

Reanalysis products represent a cornerstone in modern atmospheric and climate research, combining observational datasets with numerical weather prediction (NWP) models through sophisticated data assimilation techniques [233], [234]. As discussed in Chapter 2, data assimilation integrates heterogeneous observations, originating from satellites, ground-based stations, radiosondes, and aircraft measurements, into a consistent four-dimensional representation of the atmosphere. This process ensures that the reanalysis output is not merely an interpolation of observations, but rather the best physically consistent estimate of the atmospheric state, constrained by both dynamical laws and real-world measurements. The analysis of global tem-

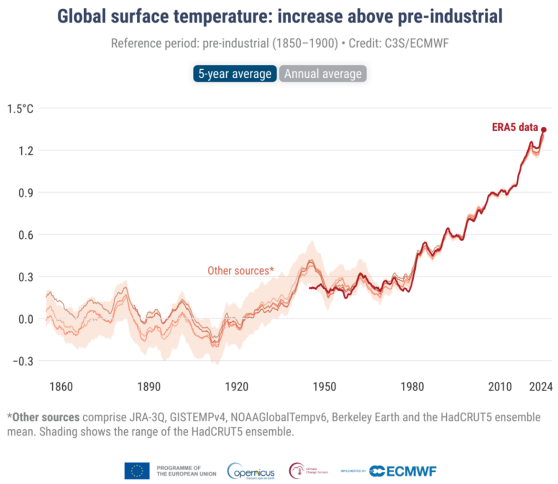

FIGURE 4.16: The evolution of global surface temperature relative to the pre-industrial period (1850–1900), shown here using the ERA5 dataset and compared with other independent historical series (JRA-3Q, GISTEMPv4, NOAAGlobalTempv6, Berkeley Earth, and HadCRUT5), reveals a steady and significant increase starting in the second half of the 20th century. The graph highlights how, in recent decades, the rate of increase has accelerated, leading to a thermal anomaly exceeding +1.4 °C in 2024 compared to pre-industrial levels.

perature trends over the past two centuries (Figure 4.16) shows a marked increase in surface temperatures relative to the pre-industrial period, providing the climatic background against which the local or global effects of natural events such as volcanic eruptions can be assessed.

In the context of volcanic eruption studies, reanalysis products offer several advantages: Long-term temporal coverage enables the study of historical events, even prior to the modern satellite era; Global spatial consistency ensures the ability to compare eruptions across different geographic regions; multi-variable data availability allows for the concurrent analysis of dynamical, thermodynamical, and compositional atmospheric parameters relevant to volcanic forcing of climate. These characteristics make reanalysis indispensable for detecting and quantifying the atmospheric perturbations induced by volcanic events, from immediate stratospheric aerosol injections to subsequent climatic feedback [160]. Multiple institutions produce reanalysis datasets,

each optimized for different temporal spans, spatial resolutions, and variable suites. ERA5 (ECMWF Reanalysis v5) – Produced by the European Centre for Medium-Range Weather Forecasts [121], ERA5 provides hourly data at 31 km horizontal resolution, with comprehensive coverage from 1940 to the present. It assimilates a wide range of satellite radiances and in-situ measurements using the 4D-Var assimilation system within the IFS (Integrated Forecasting System). MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, Version 2), Developed by NASA's Global Modeling and Assimilation Office (GMAO) [235], MERRA-2 covers 1980 onwards, with strengths in aerosol assimilation and atmospheric chemistry. NOAA 20th Century Reanalysis (20CRv3), Produced by NOAA and the Cooperative Institute for Research in Environmental Sciences (CIRES), it extends back to 1836, assimilating only surface pressure observations and using prescribed sea surface temperature and sea ice boundary conditions [236].
Here, ERA5 has been used because is considered one of the most advanced and reliable reanalysis products currently available, owing to its: High spatio-temporal resolution (hourly data at 31 km); Comprehensive observation assimilation, including modern hyperspectral satellite radiances, GNSS radio occultation data, and advanced ground-based networks; Consistent long-term coverage back to 1940, enabling both near-real-time monitoring and historical event analysis; Robust representation of the stratosphere, critical for volcanic aerosol and SO_2 transport studies. Several comparative studies have shown that ERA5 often outperforms earlier reanalysis generations (e.g., ERA-Interim, MERRA-2) in accurately representing upper-tropospheric and lower-stratospheric dynamics [121], making it particularly suitable for studying volcanic plume dispersion and associated climatic impacts. Reanalysis datasets have been widely applied to investigate the atmospheric and climatic impacts of major volcanic eruptions. [237] used ERA-Interim to link a major stratospheric aerosol load from the Nabro eruption to monsoon-driven transport. [238] combined ERA-Interim data with climate model simulations to assess radiative effects of the Sarychev eruption. Eyjafjallajökull (2010): ERA5 data facilitated the modeling of ash dispersion and its implications for aviation safety.
In the present work, ERA5 is leveraged to conduct a signal analysis aimed at detecting subtle atmospheric perturbations linked to volcanic activity. This involves isolating anomalies in meteorological variables, such as temperature, wind, geopotential height, and humidity, from their climatological background, with the goal of revealing possible links between volcanic emissions and short-to-medium-term climate variability.

4.9.2 Data source: ERA5

ERA5 represents the fifth generation of global reanalysis products developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) and is the natural evolution of ERA-Interim, with significant improvements in spatial and temporal resolution, the completeness of available variables, and overall data quality [121]. It covers a time span from 1940 to the present and provides atmospheric, surface, and oceanic fields with a spatial resolution of approximately 31 km and hourly temporal resolution, thus allowing for the

investigation of both long-term climatic phenomena and short-lived atmospheric events characterized by high variability. At the core of ERA5's production system is the Integrated Forecasting System (IFS), cycle 41r2, which employs a four-dimensional variational data assimilation approach (4D-Var) [234]. This methodology enables the consistent integration of an extremely wide range of heterogeneous observations, originating from both satellite platforms and ground-based or upper-air measurement networks. These include radiance from hyperspectral and microwave instruments, atmospheric profiles derived from GNSS radio occultation, observations from surface meteorological stations, ocean buoys, radiosondes, ships, and aircraft, as well as a set of reprocessed historical satellite datasets that extend the temporal consistency of the reanalysis back several decades.

This balanced integration of multiple sources ensures an atmospheric representation that is both physically consistent and closely aligned with observations, an essential characteristic for both climatological studies and the analysis of extreme events. One of ERA5's most remarkable features is the extensive range of variables it provides. The reanalysis includes a wide array of meteorological and climate parameters, from surface variables such as two-metre temperature and humidity, surface pressure, ten-metre wind, radiative fluxes, and precipitation, to upper-air atmospheric fields on 37 pressure levels, including temperature, geopotential height, wind, humidity, and ozone, thereby covering the entire troposphere and stratosphere.

This richness of information makes it possible not only to describe the state of the atmosphere at a given time, but also to reconstruct in detail the three-dimensional dynamics and the interactions between physical and chemical variables. From an operational standpoint, ERA5 offers the possibility to begin with a global analysis, investigating anomalies in key parameters on a planetary scale, and then progressively narrow the focus to more restricted geographical areas, such as a single hemisphere or a specific sector within it. This scalable approach is particularly useful in studies of the atmospheric impact of volcanic eruptions: starting from the identification of global signals, for example, changes in stratospheric temperatures or zonal wind fields, it is possible to progressively zoom in on the regions actually influenced by the volcanic cloud, track its spatial and temporal evolution, and quantify its local climatic consequences.

The combination of global coverage, temporal detail, and variable richness makes ERA5 a particularly powerful tool for seamlessly moving from a synoptic global perspective to a targeted analysis of regional atmospheric phenomena, such as those induced by major explosive eruptions. In this work, the use of ERA5 data is made possible through the Copernicus Climate Data Store (CDS) API, which provides programmatic access to the entire ERA5 archive, allowing the automated retrieval of tailored subsets of data for specific time ranges, geographical domains, and variables of interest. This approach not only ensures reproducibility and efficiency but also facilitates large-scale data processing within advanced analytical workflows. More information on the Copernicus CDS API is available at: Copernicus Link.

4.9.3 The 2022 Hunga Tonga-Hunga Ha'apai Eruption

On 15 January 2022, the submarine volcano Hunga Tonga–Hunga Ha'apai, located in the South Pacific approximately 65 km north of Tongatapu, produced one of the most powerful eruptions of the satellite era. The event was characterized by an exceptionally violent explosive phase, generating an umbrella cloud that penetrated the mesosphere and triggering a global pressure wave that was detected worldwide. The eruption's magnitude in terms of plume height and energy release was notable, yet its atmospheric composition differed markedly from that of other large-magnitude volcanic events. While many major eruptions are defined by their substantial emissions of sulfur dioxide (SO_2), a precursor to stratospheric sulfate aerosols capable of exerting a cooling effect on the global climate, the Hunga Tonga event released only a relatively modest quantity of SO_2.
As will be discussed in subsequent sections, this limited sulfur injection suggests that the eruption's potential for inducing a measurable global radiative cooling was minimal. In contrast, one of the most remarkable aspects of this eruption was the unprecedented injection of water vapor into the stratosphere. From a radiative perspective, the introduction of large quantities of water vapor into the stratosphere enhances the greenhouse effect, as H_2O is a potent infrared absorber. This could be led to localized warming in the atmospheric regions directly influenced by the injected plume, particularly in the vicinity of the eruption's source region and along the principal transport pathways of the moist air mass.
It is important to underline that, despite the extraordinary vertical development of the eruption column and the record-breaking water vapor injection, the event's overall climatic impact on a global scale is expected to have been negligible. The anticipated radiative effects are predominantly regional in nature, manifesting as short- to medium-term temperature anomalies over limited spatial domains in the southern hemisphere, closely linked to the dispersal patterns of the plume. As such, the Hunga Tonga eruption represents an exceptional opportunity to study localized radiative and dynamical responses in the atmosphere, without the complicating influence of a large-scale aerosol-induced cooling signal.

4.9.4 Results

The primary objective is to eventually identify any detectable change in the atmospheric temperature signal that can be directly associated with, and proportional to, the eruption event. To achieve this, the analysis focuses on the 2 m temperature product from ERA5, which represents the ambient air temperature at two meters above the surface. This parameter is derived from a combination of surface observations and model outputs and is widely used in climate and meteorological studies as a reliable indicator of near-surface thermal conditions.
Given the expectation that the Hunga Tonga eruption would not produce a substantial shift in global mean temperature, the analysis adopts a more targeted spatial approach. In addition to the global dataset, particular attention is devoted to temperature fields over the Southern Hemisphere, the region in which both the SO_2 plume and the injected water vapour were confirmed to have been present and actively dispersing for up to one week following the eruption. By narrowing the analysis to this hemisphere, the study aims

to maximize the sensitivity to localized warming signals potentially induced by the stratospheric water vapor injection, while minimizing the influence of unrelated temperature variability in unaffected regions. This methodological focus allows the detection of short-term, spatially confined thermal anomalies, thereby offering a more nuanced understanding of the atmospheric response to the January 2022 Hunga Tonga–Hunga Ha'apai eruption than would be possible through a purely global-scale analysis. The figure 4.17 shows the evo-

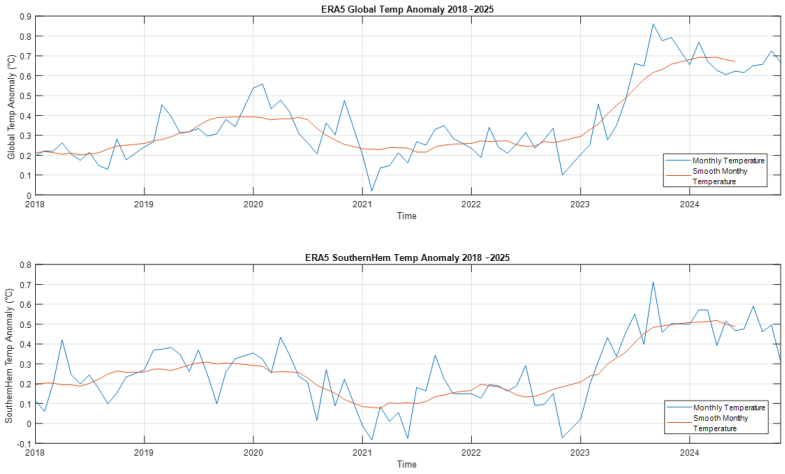

FIGURE 4.17: The evolution of mean 2 m temperature anomalies in global and Southern Hemisphere.

lution of mean 2 m temperature anomalies, derived from the ERA5 product, for the period 2018–2024. The analysis period, 2018–2025, was deliberately chosen to include four years prior to and four years following the January 2022 eruption. This temporal window is considered adequate to capture any potential shift in the temperature signal attributable to the event, while also providing sufficient pre-eruption baseline data for comparison. The four-year span on each side allows for the detection of short- to medium-term anomalies without being overly influenced by unrelated long-term climate trends.
The upper panel represents global temperature anomalies, while the lower panel shows anomalies for the Southern Hemisphere only. In both cases, the blue line indicates monthly anomalies, whereas the red line represents a smoothed moving average, which highlights longer-term trends by filtering out high-frequency variability. In the global panel, there is no marked increase in temperature immediately following the Hunga Tonga–Hunga Ha'apai eruption (January 2022), consistent with the expectation that the event did not produce a detectable global radiative impact. The variability observed in the months thereafter falls within typical interannual oscillations, and temperature trends become more pronounced only from 2023 onwards, likely in connection with broader-scale climatic forcings not directly related to the eruption. In the Southern Hemisphere panel, however, the signal is somewhat more intriguing. In the months immediately following the eruption, a slight positive anomaly is observed in Smooth Monthly Temperature signal, which could be compatible with a localized radiative effect due to the massive

stratospheric water vapor injection into the Southern Hemisphere. Nevertheless, this signal remains modest in magnitude and is quickly absorbed into the natural variability of the region. Thus, no evidence emerges of a persistent anomaly large enough to confirm a lasting thermal impact; rather, the data suggest a potential short-lived, spatially confined effect. For both the global dataset and the Southern Hemisphere, the analysis was conducted over a symmetrical time window encompassing 48 months prior to and 48 months following the January 2022 Hunga Tonga–Hunga Ha'apai eruption. This approach ensures a balanced comparison, providing a robust pre-eruption baseline against which to evaluate any post-event changes, while also allowing sufficient time for the potential signal to emerge and persist if present. For the global mean 2 m temperature, the post-eruption period shows an average increase of 0.077 °C compared to the pre-eruption baseline. The associated p-value (0.0081) indicates that this difference is statistically significant. In statistical terms, the p-value expresses the probability of obtaining a difference equal to or greater than that observed if there were no effect at all, in this case, if the eruption had no influence on global temperature. A value of 0.0081 means there is less than a 1% chance that the observed warming is due purely to random variability, which strongly suggests that the shift in temperature signal is real within the studied period.

In the Southern Hemisphere, the mean temperature increased by 0.051 °C in the post-eruption period relative to the baseline. The corresponding p-value (0.0612) is slightly above the conventional 0.05 threshold for statistical significance. This implies that, while the warming pattern is consistent with expectations, especially given that the SO_2 plume and the injected water vapor were primarily confined to the Southern Hemisphere, we cannot conclusively rule out the possibility that this increase is due to natural variability. Overall, these results suggest a clearer and more statistically robust signal at the global scale than in the Southern Hemisphere alone. While the global outcome shows a change unlikely to be explained by chance, the Southern Hemisphere data indicate a weaker signal that aligns with the physical expectations of localized warming but does not reach conventional significance levels. This distinction underscores the importance of combining statistical evidence with physical interpretation when assessing the climatic implications of volcanic events.

The statistical analysis of temperature anomalies was carried out by dividing the time series into two distinct intervals: a pre-eruption period (14 January 2017 – 14 January 2022) and a post-eruption period (14 January 2022 – 14 January 2025). For each interval, both mean values and variances were computed in order to assess potential changes attributable to the eruptive event. The comparison of variances was performed using a two-sample F-test (vartest2 function in MATLAB), while the comparison of means was carried out through a two-sample t-test without assuming equal variances.

Figure 4.18 reports the boxplot comparing temperature anomalies before and after the January 2022 eruption. Each box represents the distribution of values in the two periods: the central red line indicates the median, the blue box boundaries represent the first (Q1) and third quartiles (Q3), and thus the interquartile range (IQR), while the whiskers extend to the minimum and maximum values that fall within 1.5 × IQR from the quartiles. Any values beyond this range would appear as outliers (none are evident in this case). The pre-eruption anomalies show a lower median (0.05) and a wider spread,

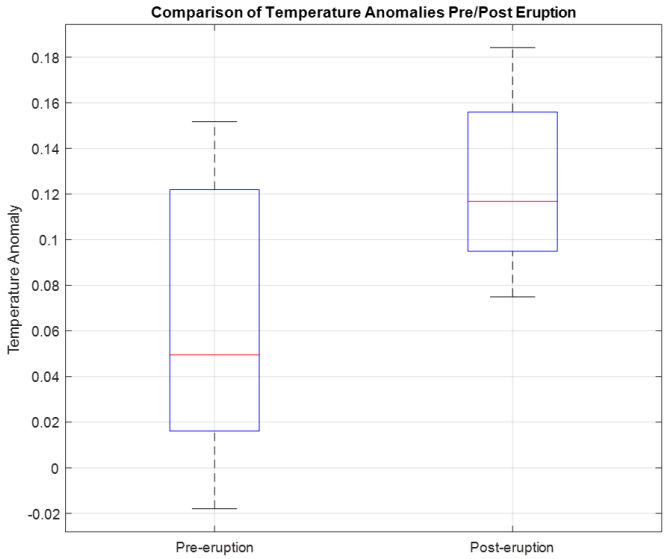

FIGURE 4.18: Boxplot of temperature anomalies before and after the eruption. The data show an increase in the median and a reduction in variability in the post-eruption period compared to the pre-eruption period.

indicating greater variability in temperature anomalies before the event. The post-eruption anomalies, instead, display a higher median (0.12) and a narrower box, reflecting that the anomalies not only increased in magnitude but also became more stable over time. The results highlight a significant increase in the mean during the post-eruption period ($p < 0.001$), indicating a persistent rise in thermal anomalies. Moreover, the test on variances ($p \approx 0.017$) revealed a statistically significant difference in data dispersion, suggesting that not only the average level of anomalies has changed, but also their variability. These findings confirm that the January 2022 eruption had a statistically relevant impact on both the intensity and the stability of the thermal anomaly signal, providing quantitative evidence of a detectable climatic effect on a hemispheric scale.

In addition, a Cross-correlation is applied, that is a mathematical technique used to measure the degree of similarity between two signals as a function of a temporal shift (lag) applied to one of them. In other words, it quantifies how well one signal matches another when displaced in time. From a practical perspective, cross-correlation is applied to: Identify temporal delays between two signals, for instance to determine whether a phenomenon observed in one signal precedes or follows that in another; Evaluate similarity and periodicity, by assessing whether two signals exhibit common patterns even if they are not perfectly aligned; Extract dependency information, revealing relationships between physical processes that are reflected in correlated signals. In MATLAB, the function xcorr is commonly used to compute the cross-correlation between two discrete signals. It returns a vector representing the correlation for different lag values, where the maximum correlation coefficient indicates the temporal shift at which the two signals are most strongly aligned.

The cross-correlation computed between the 2 m mean temperature anomalies (both for the Southern Hemisphere) for the two selected time intervals, 12

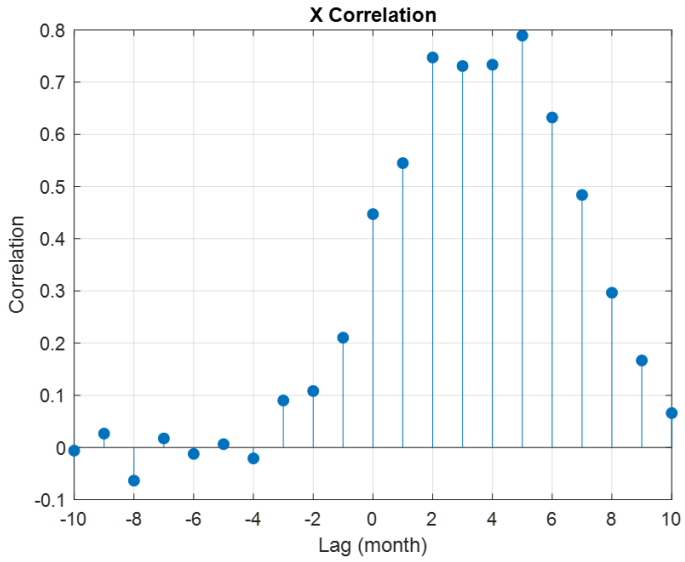

FIGURE 4.19: The trend shows that correlation values are low or close to zero for negative lags, indicating little similarity between the two series when the post-eruption period is shifted forward relative to the pre-eruption period. Starting from a lag of 0 months, the correlation increases rapidly, reaching values above 0.7 for lags between 2 and 5 months. The absolute maximum is observed at around a lag of +5 months, with a correlation coefficient of approximately 0.79.

months prior (14 January 2021 – 14 January 2022) and 12 months after (14 January 2022 – 14 January 2023) the Hunga Tonga–Hunga Ha'apai eruption, shows a maximum peak at a lag of +5 months, with a correlation coefficient of 0.789 (Figure 4.19. This indicates that the thermal signal of the year following the eruption is most like that of the year preceding it when shifted forward by approximately five months.

From a physical perspective, a positive lag of this magnitude may suggest a delayed thermal response in the region under consideration, potentially linked to the persistence of atmospheric and radiative effects resulting from the stratospheric water vapor injection into the Southern Hemisphere. At the same time, part of this correlation may reflect seasonal cyclicity, where the thermal pattern of one year tends to repeat in the following year, with possible shifts in phase due to temporary forcings. The high value of the maximum correlation suggests that, despite the volcanic event and associated anomalies, the overall structure of the thermal signal between the two years remains coherent, with differences mainly in timing rather than in the overall shape. This supports the view that the direct thermal effects of the eruption were localized and transient, while the underlying behavior continued to be dominated by seasonal variability and broader-scale climatic interconnections.

4.9.5 Discussion

The analyses carried out in this study indicate that the January 2022 Hunga Tonga–Hunga Ha'apai eruption did not produce a strong or persistent global climatic signal, but rather short-lived and spatially limited effects. ERA5 2 m temperature data show a modest yet statistically significant global warming signal (+0.077 °C, p = 0.0081) when comparing the 48 months before

and after the event, while the Southern Hemisphere, directly affected by the plume, displayed a smaller increase (+0.051 °C) that did not reach conventional significance levels ($p = 0.0612$). Cross-correlation analysis between pre- and post-eruption temperature anomalies revealed a high degree of similarity in the overall thermal pattern, with a maximum correlation of 0.789 at a lag of +5 months, suggesting a temporal shift rather than a complete change in the temperature regime. An additional point of interest that emerged during the study concerns the analysis of Aerosol Optical Depth (AOD) around the time of the Hunga Tonga–Hunga Ha'apai eruption. For this purpose, the MERRA-2 Aerosol Optical Depth Analysis dataset was selected, as its nature as a monthly reanalysis product reduces the high-frequency variability typical of daily observational datasets such as those derived from MODIS or OMI. This characteristic allows the signal associated with the eruption to be isolated and highlighted more clearly, removing much of the seasonal or meteorological noise. In the present case, the post-eruption AOD peak is clearly visible and readily interpretable, making it easier to assess the atmospheric impact of the event.

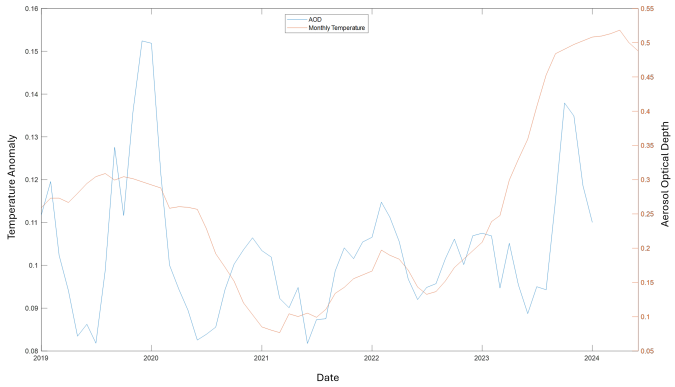

FIGURE 4.20: Time series of Aerosol Optical Depth (AOD, left axis) and monthly temperature anomaly (right axis) from 2019 to 2024. The plot highlights the temporal relationship between aerosol concentration and surface temperature variability.

Observing and monitoring AOD is particularly important in the study of volcanic eruptions, as this parameter represents an integrated measure of aerosol concentration in the atmosphere and their interaction with solar radiation. An increase in AOD following an eruption is often linked to the presence of volcanic particulates and sulfate aerosols, which can influence the Earth's radiative balance and, in some cases, have significant climatic effects. Although the Hunga Tonga eruption did not produce a large amount of SO_2 compared to other historical events, monitoring AOD still provides key information for assessing the extent and duration of the atmospheric disturbance induced by the eruption. Figure 4.20 shows the temporal evolution of Aerosol Optical Depth (AOD) and monthly temperature anomalies between 2019 and 2024. Future developments of this work could include the integration of additional atmospheric variables, such as precipitation, to investigate whether the eruption influenced hydrological patterns. Analyzing precipitation anomalies and potential shifts in annual trends could help to identify more subtle impacts

of the event and provide a more complete understanding of the volcanic perturbation's influence on the Earth system.

4.9.6 Role of Emulators in Scenario Analysis and Prediction

Regional Climate Model (RCM) emulators are tools designed to rapidly and at low computational cost reproduce the downscaling function that links the coarse-resolution variables of Global Climate Models (GCMs) to high-resolution local climate fields. This capability makes them ideal for exploring a wide range of future scenarios and quantifying the uncertainty of projections without the need to run a full, dynamically resolved RCM for every hypothesis.

In essence, emulators work as "distilled" versions of complex climate models: they retain the fundamental relationships learned from high-resolution simulations but can deliver results in computation times that are orders of magnitude shorter. In [239] two different deep learning training approaches for emulators are compared, both aimed at downscaling the Surface Mass Balance (SMB) over Antarctica. In the so-called perfect model framework, the emulator is trained on RCM data that have been artificially degraded to the resolution of a GCM (upscaled RCM data). This approach performs well in this controlled context but shows reduced accuracy when applied to real GCM data, due to structural discrepancies between global and regional models. In the so-called imperfect model framework, training is carried out directly on real GCM inputs, allowing the emulator to learn and compensate for these differences and to produce more realistic predictions when used in practice.

The study by [240] proposes instead a hybrid approach that combines statistical downscaling techniques with the explicit physics of regional climate models. In this case, a neural network architecture learns the downscaling function of an RCM and uses it to produce daily high-resolution climate fields, on the order of 12 kilometers, while maintaining good fidelity in reproducing the spatial structure and daily variability simulated by the original model. All of this is achieved with drastically reduced training and prediction times: around two hours of training on a GPU and less than a minute for producing a forecast, compared to the days or weeks required by a dynamic RCM.

Beyond these general advantages, emulators can play a decisive role in studying the interaction between volcanic activity and climate. Large explosive eruptions inject enormous amounts of sulfate aerosols into the stratosphere, increasing the reflection of solar radiation back into space (shortwave scattering) and causing a temporary cooling of the global climate system. The magnitude and distribution of this cooling depend in a complex way on the latitude and season of the eruption, the injection height of the aerosols, and the total sulfur dioxide (SO_2) load emitted. Studying these effects in detail with full dynamical climate models demands enormous computational resources, especially when the goal is to explore a broad range of possible eruptive configurations.

In this context, emulators act as high-speed virtual laboratories, capable of rapidly generating maps of temperature, precipitation, and atmospheric circulation anomalies for hundreds of different scenarios, each characterized by specific combinations of eruption intensity, timing, and physical parameters. They can also be used to systematically vary the properties of an eruption

and observe how these variations influence the global and regional climate response, identifying the most important factors for cooling or for changes in precipitation regimes. Thanks to their efficiency, emulators also make it possible to incorporate probabilistic approaches, estimating the range of possible climate responses and their associated uncertainties, and to investigate large-scale interactions such as the influence on atmospheric interconnections, for example ENSO, or on monsoonal circulations.

A particularly relevant aspect is that, with the imperfect model approach, the emulator can absorb the discrepancies between GCMs and RCMs in the background climate conditions that modulate the eruption's response, thereby improving the fidelity of the simulations. With the hybrid approach, on the other hand, the physical coherence of the RCM is preserved in the learned transformation mapping, combining speed with physical rigour. These characteristics make them ideal tools for coupling climate models with socio-economic or environmental impact models, quickly propagating the effects of a simulated volcanic event through to potential consequences on agricultural production, water availability, and public health. In this way, emulators not only accelerate the study of volcano–climate interactions but also provide an integrated and quantitatively robust framework for assessing and managing the risks associated with such events.

Chapter 5

Conclusions

5.1 Summary of key findings

This thesis has demonstrated the significant potential of satellite-based approaches, when combined with artificial intelligence (AI), to advance both volcanic monitoring and the study of climatic impacts associated with major eruptions. By integrating remote sensing observations with advanced data-driven models, the research highlights how these methodologies can complement traditional techniques, offering new perspectives for near real-time surveillance and long-term climate analysis. On the hazard-monitoring side (Chapter 3), the cascading approach integrating SqueezeNet and Random Forest models proved highly effective in detecting eruptive activity and classifying eruptive styles, achieving accuracies consistently above 0.90 across multiple case studies.

This multi-step framework not only improved the robustness of detection but also reduced false positives, thereby increasing confidence in automated monitoring pipelines. The introduction of few-shot segmentation models (SAM2) further extended the capacity to map eruptive features with minimal training data, a particularly valuable asset in data-scarce volcanic regions. Moreover, feature engineering strategies were shown to be critical for enhancing both classification and nowcasting performance, allowing for the identification of subtle patterns in satellite-derived thermal anomalies. This was especially evident in the case of Mount Etna, where the models demonstrated the ability to anticipate paroxysmal episodes with operationally significant lead times.

Another key element was the adoption of data fusion techniques, which enabled the integration of multispectral observations, temporal features, and ancillary information into a unified analytical framework. This integration proved fundamental for capturing the complex and multi-scale nature of volcanic activity, ultimately strengthening the interpretability of results and the robustness of predictions.

All these analyses and methodological advances converge toward the development of an automatic platform for volcanic monitoring, conceived as a data-driven system in which information derived from diverse satellite sources is automatically processed, integrated, and transformed into actionable knowledge. Such a platform, exemplified by the V-STAR prototype, demonstrates how combining artificial intelligence with cloud-based geospatial infrastructures can bridge the gap between raw data and operational decision-making. By turning large volumes of heterogeneous Earth Observation data into knowledge, this intelligent and modular framework lays the foundation for a new paradigm in volcanic monitoring: adaptive, scalable, and directly supportive of civil protection and scientific investigation.
On the climate side (Chapter 4), the thesis explored how volcanic activity can act as a forcing of atmospheric variability. The analysis of the 2022 Hunga Tonga-Hunga Ha'apai eruption using ERA5 reanalysis data revealed clear and statistically significant changes in atmospheric temperature anomalies before and after the event, providing empirical evidence of the eruption's influence on the global climate system. This reinforces the role of extreme eruptions in modulating short-term climate variability, complementing existing studies of volcanic forcing. Beyond temperature, innovative methodologies were applied to track volcanic gas emissions and aerosols. In particular, the zero-shot segmentation of SO_2 plumes demonstrated the potential of foundation models to generalize across different eruption contexts without explicit retraining, while the deep-learning-based nowcasting of volcanic ash clouds showed how AI can capture spatio-temporal dependencies in multispectral satellite data, supporting aviation safety and atmospheric modeling. Taken together, these applications highlight the versatility and scalability of AI-driven Earth Observation in capturing multi-scale volcanic impacts, from localized thermal anomalies to global atmospheric perturbations.

5.2 Limitations and Unresolved issues

Despite the promising results, several limitations remain. The accuracy of volcanic monitoring is still strongly constrained by external and technical factors. Cloud cover continues to represent a critical obstacle, as it can obscure thermal anomalies or plume signatures, limiting the continuity of observations even in regions with frequent volcanic unrest. Similarly, sensor spatial resolution poses an intrinsic trade-off: while high-resolution instruments such as Sentinel-2 allow for detailed mapping of lava flows or small-scale anomalies, their revisit times are relatively long and coverage is limited. Conversely, geostationary sensors such as SEVIRI provide high temporal resolution (up to minutes), which is crucial for capturing rapidly evolving eruptive events, but their coarse spatial resolution prevents fine-scale discrimination of features such as narrow lava flows or small vents. This spatio-temporal imbalance highlights the importance of data fusion strategies, integrating complementary datasets to compensate for individual sensor limitations. From a methodological perspective, machine learning models, while powerful, also face challenges. They can be sensitive to class imbalance, where eruptive scenes are underrepresented relative to quiescent periods, potentially biasing detection performance.

In addition, the presence of noise in satellite time series, whether from atmospheric effects, sensor artifacts, or thermal halos, complicates the extraction of reliable features. The availability of labeled training data is another limiting factor, since ground truth is often scarce or difficult to obtain in remote or hazardous volcanic environments. These aspects collectively constrain the scalability and generalization of the models across different volcanic settings.
On the climate side, significant limitations persist in quantifying the full extent of volcanic forcing. The attribution of volcanic signals remains particularly challenging, as distinguishing them from background variability and anthropogenic influences requires both long, continuous observational records and multi-model inter comparisons. Moreover, the representation of stratospheric processes in models is still uncertain: aerosol microphysics, particle size evolution, transport dynamics, and water vapor feedback all contribute to uncertainties in estimating the radiative impact of large eruptions. The limited vertical and temporal resolution of reanalysis products further complicates the detection of short-lived but intense perturbations, such as those observed after the 2022 Hunga Tonga-Hunga Ha'apai eruption.
In summary, the current limitations are not solely algorithmic but reflect a complex interplay of observational constraints, methodological challenges, and intrinsic climate variability, underlining the need for future improvements in sensor design, data integration, and hybrid modeling approaches.

5.3 Future perspective: Toward integrated Monitoring framework and next generation AI

Looking ahead, the integration of advanced AI methodologies with increasingly sophisticated Earth observation systems paves the way for a new generation of volcanic and climate monitoring frameworks. One of the key directions will be the development of fully integrated platforms capable of harmonizing multisensor datasets, thereby overcoming the inherent limitations of spatial and temporal resolution. For example, data fusion techniques combining high-resolution polar-orbiting sensors with high-frequency geostationary platforms will enable both the fine-scale detection of small anomalies and the continuous monitoring of rapidly evolving events.
Another critical perspective lies in the adoption of adaptive machine learning models, capable of mitigating issues such as class imbalance, noisy time series, and limited labeled training data. Semi-supervised and self-supervised approaches, along with the application of Foundation Models trained on massive heterogeneous datasets, represent promising avenues for building more generalizable and robust classifiers.
In addition, uncertainty quantification frameworks will become increasingly important, ensuring that predictions can be used operationally for decision-making in risk management. From the climate perspective, addressing the limitations identified in Chapter 4 requires both technological and methodological innovations. Future reanalysis products, enriched by higher vertical and temporal resolution and improved assimilation of volcanic aerosols and trace gases, will be essential to refine the attribution of volcanic forcing. Furthermore, the integration of multi-model ensembles with AI-based emulators will help disentangle volcanic signals from natural variability and

anthropogenic trends, thereby reducing uncertainties in climate impact assessments.
Looking further ahead, the rise of quantum machine learning (QML) offers a transformative opportunity. Quantum-enhanced AI models, leveraging the principles of quantum computing, could address the scalability issues of current deep learning architectures, enabling faster processing of large-scale satellite datasets and more efficient handling of complex, high-dimensional problems such as time series forecasting and spatio-temporal anomaly detection. Although still at a conceptual and experimental stage, QML has the potential to open entirely new paradigms in Earth observation, offering both computational speedups and enhanced learning capabilities.
In summary, future monitoring frameworks will likely evolve toward intelligent, data-driven platforms that integrate multi-sensor observations, advanced AI, and emerging quantum technologies. Such frameworks will not only enhance near-real-time volcanic hazard monitoring but also improve our capacity to assess and predict the broader climatic impacts of eruptions, providing a more resilient and adaptive infrastructure for disaster risk reduction and climate policy support.

5.4 Implications for volcanic risk and climate policy

The findings of this thesis have direct implications for both disaster risk reduction and climate governance, underscoring the dual relevance of satellite- and AI-driven approaches. On the hazard management side, enhanced forecasting, detection, tracking, quantifying, and nowcasting capabilities provide civil protection agencies with timely and actionable information, allowing for rapid evacuation strategies, resource allocation, and risk communication to exposed communities. This operational advantage is particularly critical for volcanoes with high population density in their surroundings, where minutes or hours of lead time can significantly reduce casualties and economic losses. By offering automated, near-real-time assessments, the proposed methodologies reduce the dependency on manual interpretation and increase the efficiency of emergency response systems. On the climate side, the improved quantification of volcanic forcing contributes to refining climate projections and attribution studies, strengthening the capacity to disentangle volcanic signals from anthropogenic drivers of climate variability. This has a direct impact on international policy frameworks, such as the IPCC Assessment Reports and the Paris Agreement, where accurate estimates of radiative forcing are necessary to evaluate mitigation pathways and climate resilience strategies. By providing robust evidence of the role of major eruptions in perturbing the climate system, these methods can inform both scientific consensus and policy negotiations at the global level.
Moreover, the integration of AI-driven monitoring into global observing systems represents not only a scientific advancement but also a strategic tool for societal resilience. These systems can serve as prototypes for broader multi-hazard monitoring platforms, capable of integrating volcanic, seismic, and climatic data into a unified decision-support framework. Finally, the broader implication of this work is to highlight how data-driven intelligence can bridge the gap between Earth system science and societal needs. By transforming

massive volumes of satellite observations into operational knowledge, the approaches developed in this thesis contribute to building a more adaptive, anticipatory, and resilient society in the face of both acute volcanic crises and the chronic challenge of climate change.

Appendix A

Review of satellite remote sensing

Remote sensing is the observation of an object without being in direct physical contact with it. There are different ways to collect data, and different sensors are used depending on the application. Sensors are carried by remote sensing platforms, which are mainly three: ground-level platforms (like cranes and towers), aerial platform (like helicopter and aircraft) and spaceborne platforms (like space shuttles, geostationary and polar-orbiting satellites). Remote sensing is generally referred to as a technology to survey the Earth's surface observing and analyzing the electromagnetic radiation that is either reflected or emitted from the surface (Shimizu, s.d.). The sensors onboard the remote sensing platforms receive electromagnetic radiation from different objects and surfaces and convert them into a signal that can be recorded and exhibited as numerical data or as images.

A.1 Physical basics

Electromagnetic radiation is the energy propagated through free space or through a material medium in the form of electromagnetic waves. It can be described by two principal theories, the theory of the waves and the theory of the particles (Mårtensson, s.d.). According to the theory of waves, an electromagnetic radiation is a sinusoidal wave which travels at the speed of light c (constant around 3 x 108 m/s). The wavelength and the frequency are respectively the distance from one wave peak and the next and the number of peaks passing a point per unit time. The wavelength and the frequency are related inversely, as is evident from the equation (1).

$$c = v \cdot \lambda$$

The characteristics of the electromagnetic radiation vary in the wavelength

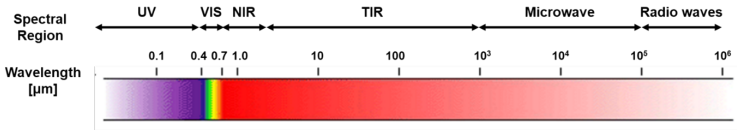

FIGURE A.1: Ranges of electromagnetic spectrum.

and the while range of wavelengths is called electromagnetic spectrum (Figure A.1). The electromagnetic spectrum is simply the full range of wave frequencies that characterizes solar radiation, and it extends from shorter wavelengths, such as gamma and X-rays, to longer wavelengths, including microwaves and broadcast radio waves. In remote sensing applications, the micrometer (µm) is the most common measuring units used when referring to the electromagnetic spectrum. Most of the electromagnetic spectrum cannot be detected by the human eye. Even satellite detectors only capture a small portion of the entire electromagnetic spectrum. The electromagnetic spectrum is generally divided into seven regions (in order of decreasing wavelength and increasing frequency and energy): radio waves, microwaves, infrared (IR), visible (VIS) light, ultraviolet (UV) light, X-rays, and gamma-rays.
There are several regions of the electromagnetic spectrum which are useful for remote sensing, the most frequently used are the bands in the visible and near visible spectrum. A band, or channel, is an interval of electromagnetic spectrum with similar characteristics. On the shorter and on the longer wavelength side of the visible spectrum are respectively UV wavelengths and the Near InfraRed (NIR) wavelengths, followed by the Thermal InfraRed (TIR) wavelengths. NIR-band consists of radiation reflected from the sun, while TIR of emitted radiation from various heat sources, and the boundary between these two regions is at approximately 3 μm. According to the theory of particles, all objects above 0°K (-273°C) emits electromagnetic radiation (Blackett, 2017). The total energy increases rapidly as the temperature increases, in all the wavelengths of emission, following the Stefan-Boltzmann's Law (J, 1879):

$$M = \sigma T^4$$

Where M is the total amount of radiated energy from the surface of an object (W/m2), is the Stefan-Boltzmann constant ($5.669710^(-8)$ W/m2/°K) and T is the absolute temperature (°K). Therefore, the radiant flux density (i.e., rate of energy flow per unit area) emitted by an object is directly proportional to the fourth power of the object's surface temperature, and then the hotter an object, the grater the total energy it will radiate (Blackett, 2017). Hot objects emit energy at higher rate than cold objects. The Stefan-Boltzamann's Law is valid for a blackbody, which is ad idealized body that absorbs and re-emits all the incident electromagnetic radiation. The blackbody temperature has an inverse relationship with the wavelength with the highest emissive power, according to the Wien's Displacement Law (Wien, 1896):

$$\lambda_{\max} = \frac{A}{T}$$

Where λm is the wavelength with the maximum radiation, A is a constant (2898 μm K) and T is the absolute temperature (K) of the object. So, the

wavelength at which maximum emission occurs shifts to shorter wavelengths as an object's temperature increases.

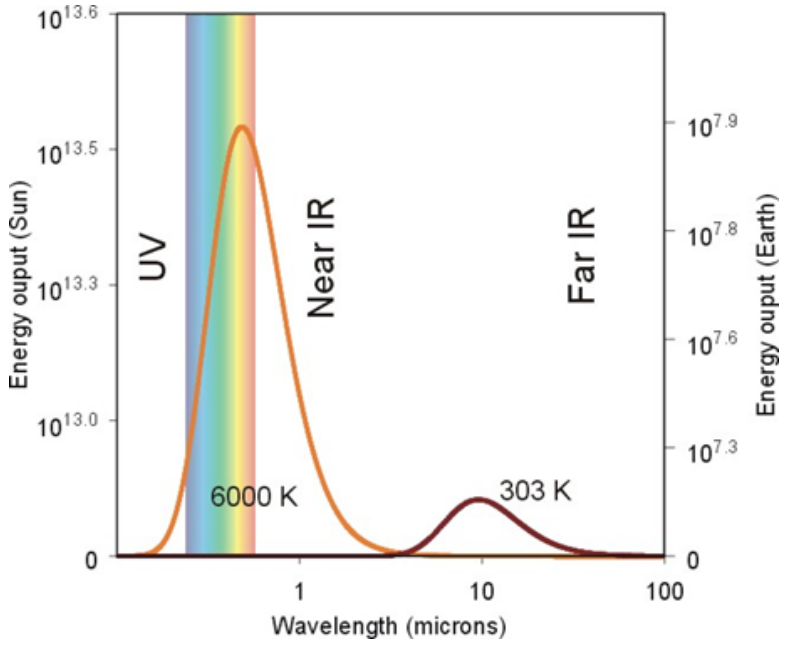

FIGURE A.2: Emission as a function of wavelength of two objects with different temperatures, the sun and the earth.

Figure A.2 shows the emission as a function of the wavelengths of two objects with two different temperatures, the earth and the sun. The earth has an ambient temperature of 300 K and emits radiation predominantly in the infrared spectrum, with a maximum around 10 m. On the contrary, the sun has a high surface temperature, of about 6000 K, and emits radiation at shorter wavelengths, more specifically it emits visible light having a maximum at around 0.5 μm. The intersection of the two curves occurs slightly below $5\mu m$, indicating that observations of the earth's surface at wavelengths less than approximately 4 μm are primarily influenced by reflected sunlight. Conversely, wavelengths exceeding 6 μm are dominated by the Earth's natural emissions. To accurately measure the natural energy emitted by the earth, it is necessary to choose a sensor with wavebands in the 10-12 μm range, close to the maximum radiation of the Earth. The energy radiated from an object is defined radiance, quantified in terms of radiance per unit of wavelength. The spectral radiance Lλ(T), in units of W $\cdot$ m^2 $\cdot$ $sr^{-}1$ $\cdot$ μm-1, emitted by a blackbody at a temperature T and wavelength , can be calculated using the Planck function (Planck, 1901):

$$L_\lambda(T) = \frac{C_1}{\lambda^5 \cdot \left(e^{\frac{C_2}{\lambda T}} - 1\right)}$$

Where C_1 and C_2 are two constants, respectively 1.19 $\cdot$ 10-16 and 1.44 $\cdot$ 10-2 mK (Wooster and Rothery, s.d.).
The Stefan-Boltzmann's Law, the Wien's Displacement Law and the Planck function are represented in Figure A.3.
Natural objects are not blackbodies: not all the incident energy upon them

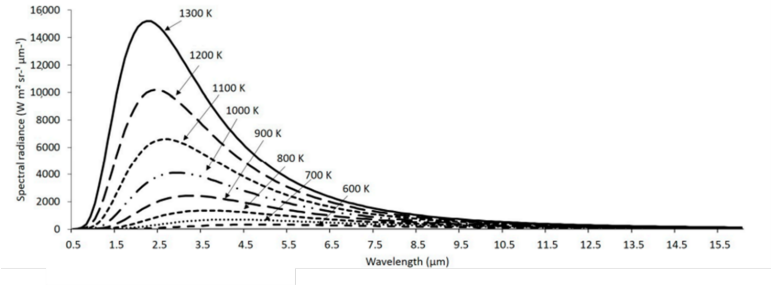

FIGURE A.3: Relationship between the spectral radiant emissions (in terms of quantity and wavelength) and temperature of the emitting surface, thus representing the Planck function and both the Stefan–Boltzmann and Wien displacement laws (sourced from Blackett et al. 2017)

is absorbed and they emit only a fraction of the energy emitted from a blackbody. Therefore, they can have different emissivity in different parts of the spectrum. The emissivity is the emitting ability of a surface defined as the ratio of the energy radiated from an object's surface to the energy radiated from a blackbody at the same temperature:

$$\varepsilon = \frac{L(\lambda, T)}{L_{BB}(\lambda, T)}$$

Where L(λ,T) is the real blackbody spectral radiance while LBB(λ,T) is the blackbody spectral radiance. All radiation is influenced by the atmosphere in different way. The atmosphere influences the incoming radiation by atmospheric scattering and atmospheric absorption. Scattering is a process of diffusion, caused by the interaction of radiation with matter present in the atmosphere, resulting in a redirection of part of the energy to other directions. Absorption occurs when part of radiation is absorbed into the particles and some of it will eventually be re-emitted from the atmosphere ash heat. Both processes depend on the size of the present atmospheric particles and on the wavelength of the radiation. The atmosphere's ability to permit the passage of radiation is known as its transmissivity (λ), and it depends on the wavelength or on the type of the radiation (Figure A.4). The gases that influ-

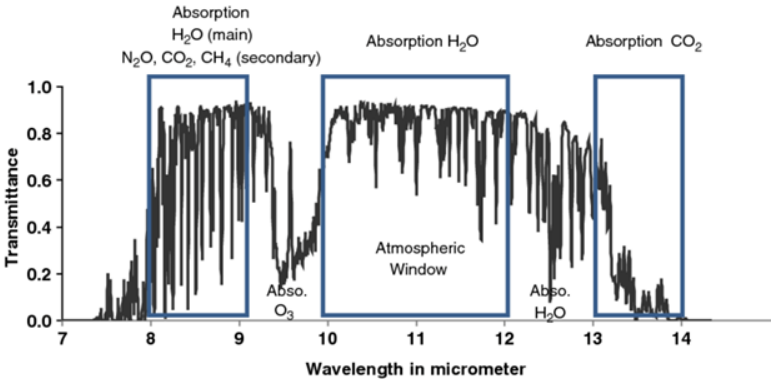

FIGURE A.4: Atmospheric transmittance in infrared region.

ence the atmosphere absorb radiation at specific wavelengths while allowing

radiation with different wavelengths to pass through. The most important absorbers of the incoming solar radiation in the atmosphere are water vapour, carbon dioxide and ozone. In contrast to the absorption bands, there exist regions in the electromagnetic spectrum where the atmosphere shows transparency, meaning there is minimal or no absorption of radiation at specific wavelengths. These wavelength ranges are termed "atmospheric windows" because they facilitate the passage of radiation through the atmosphere to the Earth's surface. The atmosphere is essentially opaque to electromagnetic radiation in part of the mid-infrared (mid-IR) and the entire far-infrared (far-IR) regions. On the other hand, in the microwave region, a significant portion of this radiation travels through unobstructed, allowing radar waves to reach the Earth's surface. Electromagnetic radiation incident on the surface of a body is partly reflected, partly absorbed or transmitted depending on radiation wavelength, material and surface conditions of the body. Remote sensing is based on the measurement of reflected or emitted radiation from different bodies. Radiance is directly measured by remote sensing instruments.

$$L(\lambda) = L_s(\lambda)\,\tau(\lambda) + L_u(\lambda)$$

Where $L_U(\lambda)$ is the upwelling radiance, i.e. the radiance emitted by the atmosphere, and $L_G(\lambda)$ is the ground radiance, defined as the sum of the emitted radiance $L_e(\lambda, T_G)$ and the reflected radiance $L_R(\lambda)$, respectively Eqs. (A.1) and (A.2).

$$L_e(\lambda, T_G) = \varepsilon(\lambda)\, L_{BB}(\lambda, T_G) \tag{A.1}$$

$$L_R(\lambda) = (1 - \varepsilon(\lambda))\, L_D(\lambda) \tag{A.2}$$

Where LD(λ) is the downwelling radiance, i.e. the radiance downwelling from the atmosphere that reaches the ground. Part of this is reflected back.

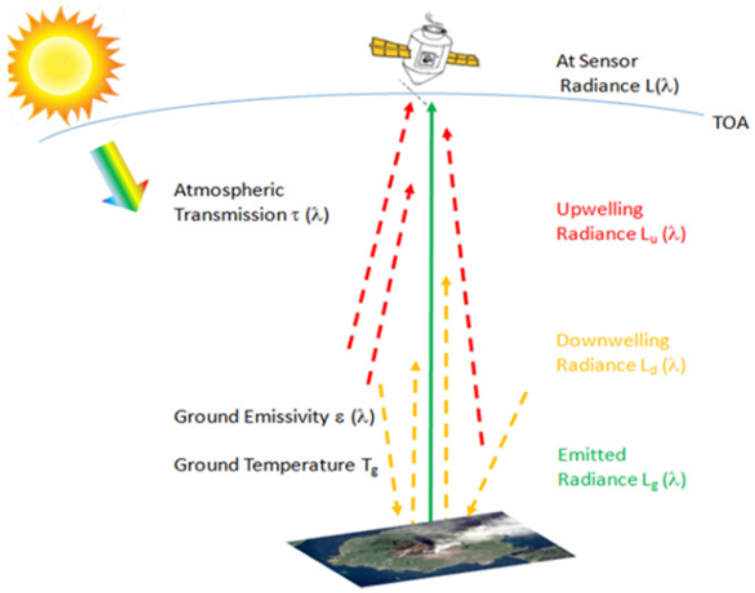

FIGURE A.5: Atmospheric transmittance

A.2 Satellite systems

The path followed by a satellite is referred to as its orbit. Generally, there are two major orbits, which are called geostationary and polar orbits. A geostationary orbit is located about 36.000 km above the Earth's equator. A geostationary satellite orbits around the earth at the same rotational speed as the Earth itself, so it needs exactly 24 hours to complete an orbit. As a result, it appears to remain stationary relative to a fixed point on the Earth's surface. This stationary position allows geostationary satellites to provide continuous coverage of a specific area, making them ideal for applications like communication, weather monitoring, and Earth observation. Due to the high altitude of the satellite orbit, the geometric resolution is very low, but this limitation is offset by the advantage of a more frequent revisiting capability. The METEOSAT series of satellites, operated by EUMETSAT, are examples of geostationary satellites.
A polar orbit is a type of orbit in which a satellite passes over the Earth's poles as it travels from north to south and vice versa with an inclination near 90°. This orbit is also called sun-synchronous because the section monitored by the satellites is always radiated by the sun in the same way. Satellites in polar orbits typically have lower altitudes, often ranging from a few hundred kilometers to a few thousand kilometers above the Earth's surface. The polar orbit is not stationary, and the satellite's path covers the entire Earth's surface over time. Days or weeks later, the satellite orbits again above the same section. This allows for global coverage and is suitable for applications like Earth observation, remote sensing, and environmental monitoring. Hence, the temporal resolution of these satellites is limited compared to geostationary satellites, but they offer superior spatial resolution. The LANDSAT series, managed by NASA and the U.S. Geological Survey, is an example of polar orbiting satellites.

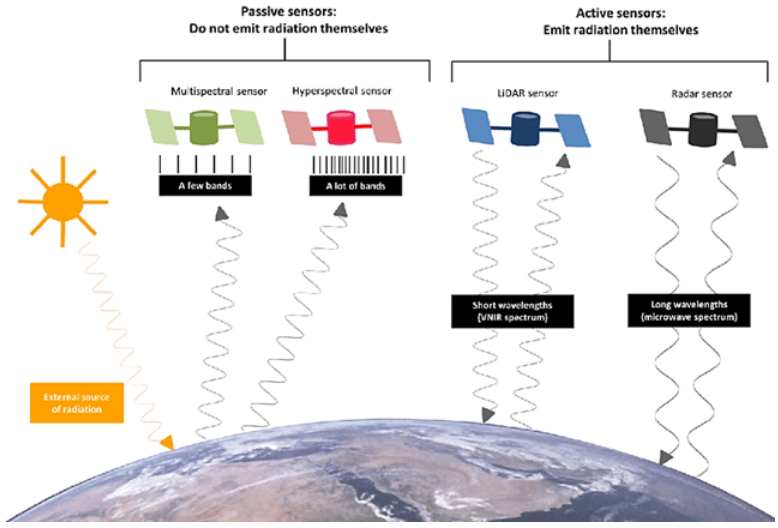

FIGURE A.6: Active and passive sensors.

Another way to distinguish between Earth observation satellites is by examining the types of sensors they employ. Typically, these sensors fall into two categories: passive sensors that detect reflected sunlight or thermal radiation,

and active sensors that generate their own radiation source (Figure A.6). Passive sensors detect and measure natural radiation emitted or reflected by the Earth's surface and atmosphere. They do not emit their own energy source. Common examples of passive sensors include optical sensors that capture visible and infrared light, as well as thermal sensors that detect heat radiation. Most passive sensors include scanners for imaging. Equipped with spectrometers, these sensors are capable of simultaneously measuring signals across multiple spectral bands and then creating multispectral images. Active sensors, in contrast, emit artificial radiation to monitor the earth surface or atmospheric features, and then measure the return signals that interact with the target. Radar (Radio Detection and Ranging) and lidar (Light Detection and Ranging) systems are typical examples of active sensors. Radar uses microwave signals, while lidar uses laser pulses. Active sensors do not depend on daylight and are hardly affected by bad weather conditions. Satellite systems measure electromagnetic radiation in different bands of the electromagnetic spectrum. Each band of a multispectral image can be displayed individually as a grayscale image or in a combination of three bands at time as a color composite image. The three primary colors of light are red, green, and blue (RGB). A natural or true color composite is an image displaying a combination of visible red, green and blue bands to the corresponding RGB channels. The resulting composite closely resembles what the human eye would naturally perceive: vegetation appears green, water appears dark blue to black, and bare ground and impervious surfaces appear as light grey and brown.
However, a true color composite may not always be ideal for distinguishing subtle differences in features, since it can sometimes have lower contrast and appear somewhat hazy due to the scattering of blue light by the Earth's atmosphere. False color images use different combinations of bands and enable us to visualize wavelengths that the human eye cannot see, such as near-infrared. Therefore, a false color image is used to enhance spectral separation and often makes the data more interpretable.
Satellite sensors store information about objects as a grid. Digital data is gathered from the covered area as individual image points, which are referred to as pixels. A pixel is the smallest unit within a digital image, representing a discrete portion of the image's content. Resolution plays a crucial role in how data from a sensor can be used, and it can vary based on factors such as the satellite's orbit and sensor design. There are generally four types of resolution to consider for any dataset:

- **Spatial resolution** (also known as ground resolution) is the ground surface area that forms one pixel in the satellite image. Higher resolution, which means a finer grid with smaller pixels, enables recognition of detailed features on the Earth's surface.

- **Temporal resolution** is the measure of the repeat cycle with which a sensor revisits the same portion of the Earth's surface. The temporal resolution is determined by altitude and orbit of the satellites as well as its sensor characteristics.

- **Spectral resolution** is the ability of a sensor to discern finer wavelengths. Sensors with 3–10 bands are called multispectral, while sensors with hundreds to even thousands of bands are called hyperspectral.

- **Radiometric resolution** is the amount of information in each pixel, that is, the number of bits representing the recorded energy. It specifies how well the differences in brightness in an image can be perceived. The higher the radiometric resolution is, the better small differences in reflected or emitted radiation can be measured.

Appendix B

Review of machine learning techniques

Machine learning (ML) is a branch of artificial intelligence (AI) that enables computers to "self-learn" from training data and improve over time, without being explicitly programmed. The term machine learning was first introduced in the 1950s by Artur Samuel. He developed a self-learning computer program for playing checkers and he observed that the more the system played, the better it performed (Agile Software Development, 2023). Machine learning algorithms are automated processes that are able to identify patterns within data and learn from them, so they solve problems and make predictions according to past observations, unlike traditional programming in which a computer uses instructions based on if-then structure to transform input data. Machine learning has experienced significant evolution and growth over the past three decades for several key reasons:

- **Big data**: the digital age has generated immense volumes of data. Machine learning thrives on data, and the availability of big data sources has provided the raw material needed for training and refining models.
- **Algorithms and techniques**: the field of machine learning has witnessed remarkable progress in the development of new algorithms and techniques, which have enabled more sophisticated and accurate models.
- **Advancements in computing power**: the exponential increase in computing power, coupled with the development of specialized hardware like GPUs (Graphics Processing Units), has allowed machine learning algorithms to process vast amounts of data and perform complex calculations more efficiently.
- **Open Source and Collaboration**: open-source frameworks and tools, like TensorFlow, PyTorch, and scikit-learn, have fostered collaboration and knowledge sharing within the machine learning community. This

has accelerated the development and deployment of machine learning applications.

- **Applications and Success Stories**: as machine learning has demonstrated its effectiveness in various applications, from recommendation systems to healthcare diagnostics or monitoring of natural phenomena, more industries and sectors have embraced it, spurring further research and innovation.

Machine learning techniques can be divided into two main domains, namely unsupervised and supervised learning. In the case of unsupervised learning, the model identifies unknown patterns in data without pre-existing knowledge (like classes or label), while in supervised learning models are trained on labeled data to make predictions. In machine learning, the dataset is composed of different data points, also called samples, where each data point represents an entity we aim to analyze (Badillo et al., 2020).
A feature is an individual, measurable property or characteristic of a data point, and it is used as input of a model to make predictions. Feature engineering is a crucial aspect of machine learning, as selecting, transforming, and encoding the right features can significantly impact the model's performance. The process of feature selection and extraction is often guided by domain knowledge and data analysis to ensure the most relevant and informative features are used. ML algorithms need a lot of domain knowledge and human intervention. This is the reason why in some context deep learning (DL) holds a bit more promise. DL is a subset of ML which uses advanced computing structure (neural networks) to understand complex patterns, with little to no human intervention.

B.1 Unsupervised Learning

Unsupervised learning is a type of machine learning technique that identifies patterns in dataset that are neither classified nor labeled. Then, it can classify, group and label unstructured data, without any training or human interface. Unsupervised learning models are widely used for three main tasks (Delua, 2021):

- **Clustering**: it is a technique for grouping unlabeled data based on their similarities or differences, without having a predefined hypothesis on the properties that subgroups might have. One simple approach to clustering is the *K-means* algorithm.

- **Association**: it is a method which uses different rules to find relationships between variables in a given dataset.

- **Dimensionality reduction**: it is a technique used when the number of features in a given dataset is too high and therefore reduces the number of data inputs to a manageable size, while preserving data integrity. An example of an algorithm used to reduce dimensionality is the *Principal Component Analysis (PCA).*

B.1.1 K-means

K-means clustering is one of the most popular unsupervised machine learning algorithms. A cluster is a grouping of data points aggregated together

according to their similarities. The aim of the K-means is to partition a set of unlabeled data point into k clusters in such a way the sum of squared distances between data points and the mean of the cluster to which they are assigned is minimized. The steps of the algorithm are the following:

1. **Choosing the number of clusters**: the first step is to define the number of clusters k in which the data will be grouped. This is the unique parameter fixed *a priori* by the user.

2. **Initializing centroids**: a centroid is the center of a cluster, and initially random data points are used as centroids for each cluster.

3. **Assign data points to the nearest cluster**: once the centroids are initialized, each data point X_n is assigned to its closest cluster centroid C_k. During this step, for every data point the Euclidean distance (Eq. B.1) is measured from the centroids, and whichever distance is smaller, that point is assigned to that centroid.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{B.1}$$

4. **Re-initialize centroids**: the centroids are re-initialized by taking the average of all data points in that cluster.

5. **Repeat steps 3 and 4**: the algorithm repeats the calculation of centroids and assignment of points until convergence. Convergence occurs when the cluster assignments no longer change significantly (the sum of distances between the data points and their corresponding centroid is minimized), or when a predetermined number of iterations is reached.

In practice, selecting the optimal number of clusters k can be challenging, and various techniques like the Elbow Method and Silhouette Score are used to determine the most appropriate number of clusters for a given dataset. K-means clustering offers the advantages of being a straightforward and easily interpretable algorithm, efficient with large datasets, and versatile for various data types. However, it comes with the challenges of sensitivity to initial centroid placement, the need to determine an optimal number of clusters, and assumptions about spherical, equally sized clusters, which may not hold in all scenarios. K-means clustering is widely used in various fields, including image segmentation, customer segmentation, anomaly detection, and document classification, among others.

B.1.2 Principal component analysis (PCA)

Principal Component Analysis (PCA) is a widely used dimensionality reduction technique and a statistical method in machine learning and data analysis. Its primary purpose is to simplify complex datasets while preserving as much of the important information as possible. Principal Component Analysis (PCA) is a widely used dimensionality reduction technique and a statistical method in machine learning and data analysis. Its primary purpose is to simplify complex datasets while preserving as much of the important information as possible. PCA aims to capture the maximum amount of variability in the data. It achieves this by finding linear combinations of the original features,

called principal components, that maximize the variance in the data. The first principal component explains the most variability, the second explains the second most, and so on.
PCA performs an orthogonal linear transformation on the data, ensuring that the principal components (new variables) are uncorrelated with each other. This makes interpreting and analyzing the transformed data more straightforward. It involves computing the eigenvalues and eigenvectors of the covariance matrix of the data. The eigenvalues represent the variance explained by each principal component, while the eigenvectors define the directions of these components. After dimensionality reduction, PCA allows for the reconstruction of data points using a subset of the principal components. This reconstruction helps in understanding how much information is retained when reducing the dimensionality.
PCA is widely used in various fields, including image processing, feature selection, data compression, finance, and genetics. It is particularly useful when dealing with high-dimensional data, such as in computer vision or gene expression analysis. While PCA is a powerful technique, it may not be suitable for all datasets. It assumes linearity and may not perform well when relationships in the data are highly nonlinear. Additionally, it doesn't inherently provide feature importance or interpretability for the reduced components. In summary, Principal Component Analysis is a valuable technique for reducing the dimensionality of data while retaining essential information. It is widely used in data preprocessing and feature extraction to improve the efficiency of machine learning models and enhance the understanding of complex datasets.

B.2 Supervised Learning

Supervised machine learning is a type of machine learning used to describe prediction tasks because the goal is to forecast/classify a specific outcome of interest (Jiang et al., 2020). It learns the relationship between input and output. The inputs are known as features while the desired output is generally referred to as target. The learning process is divided into two steps: training and testing. In training process, labeled data which contains both the features and the target are used as input to train and build the model. The training set is used to teach model to yield the desired output. The algorithm measures its accuracy by utilizing the loss function, iteratively adjusting its parameters until the error has been sufficiently minimized. In the testing phase, the trained model is applied to a dataset that it has not seen during the training process to assess the performance. Supervised learning can be separated into two types of problems:

- **Classification**: the algorithm learns from the data to predict results in a discrete output. An example of a classification problem is when, given the medical records of a patient with a tumor, we have to predict whether the tumor is malignant or benign. When the outcome can take two possible values (True or False, Yes or No), it is known as *binary classification*. If the possible values as output are more than two, it is referred to as *multiclass classification*. Common classification algorithms are linear classifiers, Support Vector Machines (SVM), decision trees, k-nearest neighbor, and Random Forest (RF).

- **Regression**: the algorithm learns from the data to predict continuous values. In contrast to classification, where the aim is to assign input data points to discrete categories, regression involves predicting a quantity that can take any real value within a specific range. An example of a regression problem is when, given data about the size of houses on the real market, we have to predict their prices. Linear regression, logistic regression, and polynomial regression are popular regression algorithms.

In supervised learning approaches, the goal is minimizing the error between the predictions made by the trained ML model and the corresponding target. A cost function is a measure of the error between the predicted values and the actual values in the training data. The purpose of a cost function is to quantify how well or poorly a model is performing, providing a way to guide the learning algorithm during the training process. Gradient descent is an algorithm used to optimize the cost function and it is used to find the minimum value of error possible in the model. This algorithm uses the gradients of the cost function with respect to the model parameters to update the parameters in the direction that minimizes the cost. This iterative process continues until the model reaches a state where the cost is sufficiently minimized, indicating that the model has learned the underlying patterns in the data.

B.2.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the most popular ML algorithms, used for both regression and classification tasks. The goal is to define the best decision boundary that can separate n-dimensional spaces into classes, in order to easily classify new data points in the correct category.

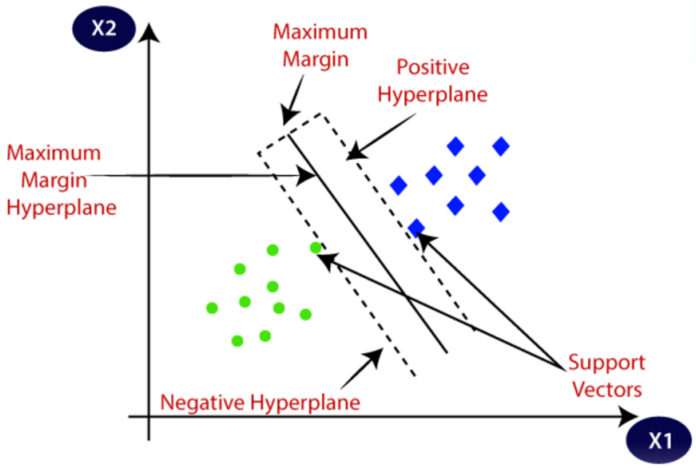

FIGURE B.1: The Support Vector Machine (SVM) optimizes margin between support vectors (Source: www.analyticsvidhya.com).

In two-dimensional space, the boundary is a line dividing a plane in two parts, while in a n-dimensional space it is called hyperplane. The support vectors are the points closest to the hyperplane, while the margin is the distance between the hyperplane and the support vectors (Figure B.1). The best hyperplane is the plane which maximizes the distance between the classes. This is done by identifying various hyperplanes which classify the data points in the best

way and then selecting the one that is the farthest from the support vectors and has the maximum margin. There are two types of SVM algorithms:

- **Linear SVM**: when the data is perfectly linearly separable, the data points can be classified into two classes using a straight line.
- **Non-Linear SVM**: when the data is not perfectly separable, more advanced techniques (such as kernel functions) can be used to classify the data points.

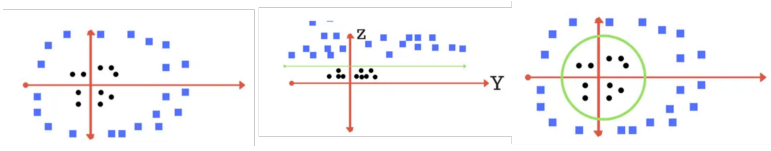

FIGURE B.2: A kernel function is used in the case of Non-Linear SVM. (a) The data point can not be separated in two classes by a straight line. (b) A new dimension (z-axis) is added thanks to a transformation and a line separate the data point. (c) The line transforms to circle in the x-y plane.

In the case of a Non-Linear SVM, the idea is to use a kernel function to transform the input data into a higher-dimensional feature space. For example, if we consider the data point in Figure B.2a, they can not be separate in two classes by a straight line. We can apply a transformation (w = x2 + y2) and add one more dimension, called z-axis. If we plot in z-axis (Figure B.2b), a clear separation is visible and a straight line can be drawn. When we transform back to x-y plane, the line transforms to circle (Figure B.2c). These transformations are called kernels. Therefore, a kernel converts non separable problems to separable problems.
We have a training dataset with input feature vectors X and the corresponding class labels Y. The linear kernel is defined as:

$$K(w, b) = w^T x + b$$

Where w is the normal vector to the hyperplane and b the offset. One of the most popular kernel functions is the Radial Basis Function (RBF) defined as

$$K(w, b) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right) \tag{B.2}$$

Where $\|x_i - x_j\|^2$ is the Euclidean distance between two points x_i and x_j. γ is a hyperparameter that controls the width of the kernel. The bigger the γ, the more "linear" the decision boundary will be. If γ approaches 0, it results in a greater number of support vectors and a more non-linear boundary.

Another important hyperparameter is the cost c, which controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights. When c is larger, the algorithm is more tolerant of a smaller margin if it improves the accuracy of classifying all training points. On the contrary, a lower c promotes a larger margin, leading to a simpler decision function at the cost of training accuracy.

B.2.2 Random Forest (RF)

Random Forest (RF) is a ML algorithm which combines the output of multiple decision trees to reach a single result. A decision tree is a method that predicts the value of a target variable by learning simple decision rules inferred from the data features. Multiple decision trees form an ensemble in the random forest algorithm, which predict more accurate results, especially when the individual trees exhibit low correlation with one another (Parmar et al., 2019). Ensemble learning methods consist of a collection of classifiers, such as decision trees, whose predictions are combined to determine the most prevalent outcome. An example of boosting method is the bagging method: several data samples are generated selecting random sample of data with replacement in a training set (data points ca be repeated) and then these models are then trained independently (Gaikwad & Thool, 2015). The RF algorithm extends the bagging method by incorporating both bagging and feature randomness, in order to create an uncorrelated forest of decision trees. In contrast to decision trees, which evaluate all potential feature splits, random forests take a different approach by only considering a subset of those features. The three main hyperparameters to be set before training are node size, the number of trees and the number of features sampled.
The random forest algorithm is composed of a group of decision trees, each built on a data sample drawn from a training set with replacement, known as the bootstrap sample. Within this training sample, a portion (one-third) is reserved as test data, referred to as the out-of-bag (oob) sample. Another layer of randomness is introduced through feature bagging, enhancing dataset diversity and reducing correlation among the decision trees. The prediction strategy varies depending on the task: for regression, individual decision trees are averaged, while for classification, a majority vote determines the predicted class. Finally, the oob sample is utilized for cross-validation, concluding the prediction process (What Is Random Forest?, s.d.).

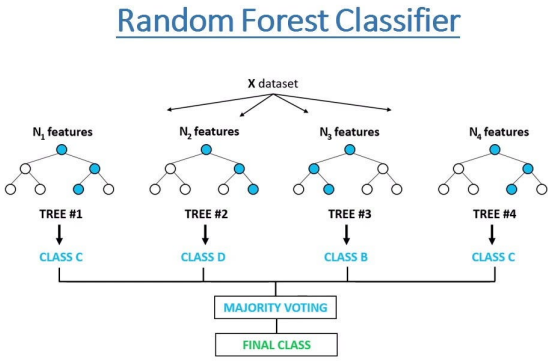

FIGURE B.3: Random forest model.

B.3 Deep Learning (DL)

Deep learning is a subfield of machine learning which involves the use of artificial neural networks (ANN) to model and solve complex problems. An ANN

is a ML algorithm inspired by biological neural network (James et al., 2021). It is composed by nodes that communicate with each other via connections. Perceptron is a single layer neural network (NN), which takes multiple input features (each representing a characteristic of the input data) and attempts to find a line, plane, or hyperplane that separates the classes in a two-, three-, or hyper-dimensional space, respectively (Rumelhart et al., 1986). This algorithm uses an activation function (like the sigmoid function) to transform multiple input into a single output.
Multiple perceptrons connected to each other create an ANN, which contains a layer of input nodes, a number of "hidden layers" and a layer of output nodes. Hidden layers contribute to the ability of neural networks to learn hierarchical representations, making them powerful tools for tasks such as image recognition, natural language processing, and more. Information of each node in the previous layer passes to each node in the next layers, undergoes transformation and is subsequently forwarded to each node in the subsequent layer. This is the reason why it is known as feedforward neural network (Choi et al., 2020). Each node in an ANN contains an activation function.
For image recognition tasks, each input into a feedforward ANN corresponds to a pixel in the image. This imply that there are no connections between nodes in a layer and the spatial context of features in the image are lost (Krizhevsky et al., 2012). A convolutional neural network (CNN) is a particular case of ANN which overcomes these limits: instead of using single pixels as input, a CNN feeds patches of an image to specific nodes in the next layer of nodes (rather than all nodes), preserving the spatial relationship between pixels in an image. These patches of nodes learn to extract specific features and are known as convolutional filters (Choi et al., 2020).

B.3.1 U-Net

The U-Net was first designed and applied in 2015 to process biomedical images (Ronneberger et al., 2015). In biomedical cases, the focus is not only to distinguish whether there is a disease, but also to localize the area of abnormality. For this reason, U-Net is popular for fast and precise segmentation of images. It is a U-shaped encoder-decoder network architecture, which consists of four encoder blocks and four decoder blocks that are connected via a bridge: input images are fed into an encoder architecture to capture high-level contextual information, which is subsequently transmitted to a decoder architecture to recover spatial details and generate pixel classification results (Tomar, 2021; Weng et al., 2019).
The encoder network (or contracting path) acts as a feature extractor, acquiring an abstract representation of the input image through a series of encoder blocks. Each encoder block is composed of two 3x3 convolutions, with each convolution followed by a Rectified Linear Unit (ReLU) activation function. The ReLU activation function introduces non-linearity into the network, enhancing its ability to generalize from the training data. The output from the ReLU serves as a skip connection for the corresponding decoder block.
The skip connections facilitate gradient flow during backpropagation, which in turn helps the network to learn better representation. Then a 2x2 max pooling operation is applied to reduce by half the spatial dimensions (height and width), the number of trainable parameters decreases and consequently

the computational costs are reduced (Shin et al., 2016). Thus, at each downsampling step, the number of feature channels doubles while the spatial dimensions become half. The bridge connects the encoder and the decoder network and consists of two 3x3 convolutions, where each convolution is followed by a ReLU activation function. The decoder network (or expansive path) transforms the abstract representation into a semantic segmentation mask.
It consists of a 2x2 transpose convolution, which halves the number of feature channels, followed by the concatenation with the corresponding skip connection feature map from the encoder block. Then, two 3x3 convolutions are applied, with each convolution followed by a ReLU activation function. At the final layer, a 1x1 convolution is used to map the channels to the desired number of classes.

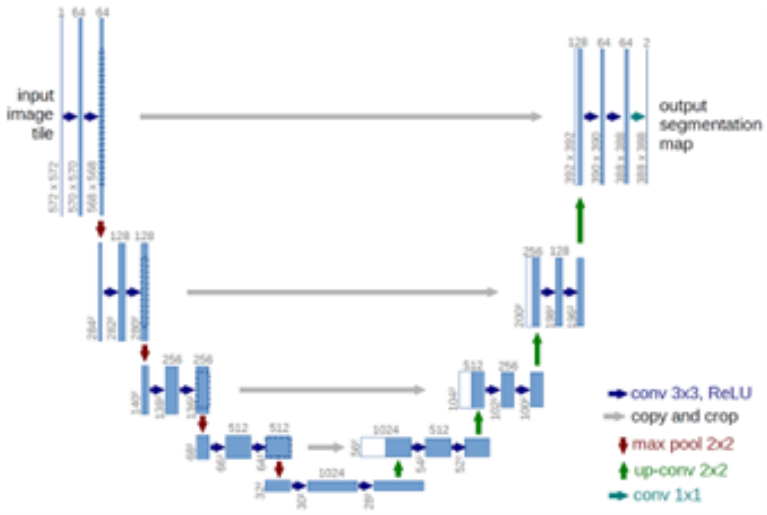

FIGURE B.4: The U-Net architecture (example for 32x32 pixels, sourced from (Tomar, 2021))

Appendix C

Review of Foundation models

A foundation model in the context of Earth Observation (EO) is an artificial intelligence model pre trained on massive, heterogeneous datasets, often unlabelled, and designed to be adaptable to a wide range of geospatial tasks. These models stand out for their ability to learn general representations from multispectral, multiscale and multimodal data, which makes them usable in multiple applications with only minimal fine-tuning. Recent studies have highlighted the advantages of such an approach. For example, Prithvi is a transformer-based model pre-trained on more than one terabyte of Harmonized Landsat–Sentinel-2 multispectral data. It has been shown that fine-tuning Prithvi allows tasks such as cloud removal, flood mapping, burn scar segmentation, and multi-temporal crop segmentation to be performed with significant improvements compared to generic models, even when only a small number of labelled samples are available.

To facilitate the practical adoption of foundation models, the TerraTorch toolkit has been developed. Built on PyTorch Lightning, TerraTorch provides a modular environment that supports fine-tuning and benchmarking of geospatial models in a largely no-code framework. It integrates pre-configured modules for satellite data handling, EO-specific tasks, and automated hyperparameter optimization. By connecting directly to GEO-Bench, TerraTorch ensures systematic comparisons across tasks and models, encouraging reproducibility and standardization within the EO community.

In 2025, the ESA ϕ-lab in collaboration with IBM Research Europe introduced TerraMind, a groundbreaking generative "any-to-any" multimodal foundation model for EO. TerraMind was trained on TerraMesh, a dataset comprising more than nine million multimodal geotemporal samples. It incorporates optical and radar satellite imagery, topography, land use and land cover data, terrain geometry, descriptive textual information, and precise geographic location. Its architecture is a symmetric transformer-based encoder-decoder that can operate simultaneously at pixel and token level, enabling a deep contextual and spatial understanding of Earth processes. One of TerraMind's most innovative features is the "Thinking-in-Modalities" (TiM) tuning paradigm.

During fine-tuning or inference, the model is able to auto-generate intermediate data from other modalities, for example reconstructing land cover maps from multi-temporal observations. This ability allows the model to fill data gaps or mitigate distortions, ensuring more robust outputs in real-world monitoring scenarios. Benchmarks such as PANGAEA have demonstrated that TerraMind outperforms previous EO models by at least eight percent in tasks like land-cover classification, change detection, and multispectral analysis, all while operating at a computational cost of less than one-tenth compared to using separate models for each modality.

The use of foundation models in Earth Observation brings several concrete advantages. Because they are pre-trained through self-supervised learning on large unlabelled datasets, they drastically reduce the dependence on manually labelled samples, which are expensive and time-consuming to obtain. Their ability to generalize across new tasks, sensors and geographical areas makes them highly adaptable and scalable. At the same time, models such as TerraMind have been designed with efficient architectures that lower the computational burden, making large-scale deployments more feasible. The capacity to combine multimodal data, ranging from radar and optical imagery to textual descriptors, further enhances robustness and interpretability, while their zero-shot and few-shot capabilities allow them to be applied immediately to critical events, such as natural disasters, without the need for costly training campaigns. Additionally, their generative functions make it possible to complete missing or degraded datasets by reasoning across modalities and contexts.

Nonetheless, several challenges remain. The training and deployment of foundation models demand high-performance computing infrastructures, which limits their accessibility. Biases may emerge because some regions or phenomena are underrepresented in the training data, while the opacity of the models can limit interpretability in scientific contexts where transparency is critical. Current models also face difficulties in deep temporal reasoning, such as the representation of seasonal climatic cycles or long-term volcanic dynamics, and in quantifying uncertainty in predictions.

Future developments are expected to address many of these issues. Hybrid approaches that combine physical constraints with data-driven models could help ensure greater scientific validity. The multimodal nature of foundation models may be further extended to incorporate text, climate simulations, physical models and historical geospatial records, providing a richer contextual basis for prediction.

Appendix D

Cloud Computing Platform

Cloud computing is the on-demand availability of computing resources, including storage and infrastructure, delivered as services over the internet. Cloud computing has the potential to significantly enhance the availability of IT resources and offers numerous advantages compared to other computing techniques. For instance, it facilitates self-help services without requiring manual interactions with service providers. Additionally, all cloud resources are made accessible to users, allowing for dynamic leasing of physical or virtual resources without the need to be aware of their specific locations. Moreover, on a cloud computing platform, all resources can be rapidly and elastically deployed (Peng et al., 2009). Google provides cloud computing services that run on the same infrastructure that Google uses internally for their own consumer products. Google Earth Engine (GEE) and Google Colaboratory (Colab) are two powerful tools provided by Google that can be used together for various geospatial data analysis tasks.

D.1 Google Earth Engine (GEE)

Google Earth Engine (GEE) is a cloud-based computing platform for scientific analysis and visualization of geospatial datasets. It simplifies the access to high-performance computing resources for the processing of extensive geospatial datasets, alleviating the current challenges associated with IT infrastructure management (Gorelick et al., 2017). GEE store satellite imagery in a public archive containing historical Earth images spanning over forty years. Moreover, it includes a set of library functions, optimized for the parallel processing of geospatial data. GEE has an intuitive interface, like it is shown in Figure 42. The most important features are:

1. **Search bar**: the user can search datasets and places.
2. **Left panel**: it includes the *Scripts* tab, the *Docs* tab (with a categorical list of available functions), and the *Assets* tab (which contains any files uploaded to GEE).

3. **Script editor**: it is the environment to write code (in JavaScript) and execute it by clicking the *Run* button.

4. **Right panel**: it includes the *Inspector* (to change the interaction with the Map viewer), the *Console* (which contains information related to data being accessed in the script editor), and *Tasks* (through which the user can track the progress of any data in the queue for export/download).

5. **Map viewer**: this is the panel to visualize images, draw shapes, and interact with the map.

GEE has become a powerful tool for remote sensing, thanks to its multiple applications in various fields such as vegetation monitoring (Tsai et al., 2018), land use and cover (Velastegui-Montoya et al., 2022), surface temperature (Ermida et al., 2020), forest fires (Parks et al., 2018) and volcano monitoring (Amato et al., 2021; Genzano et al., 2020). GEE store huge amounts

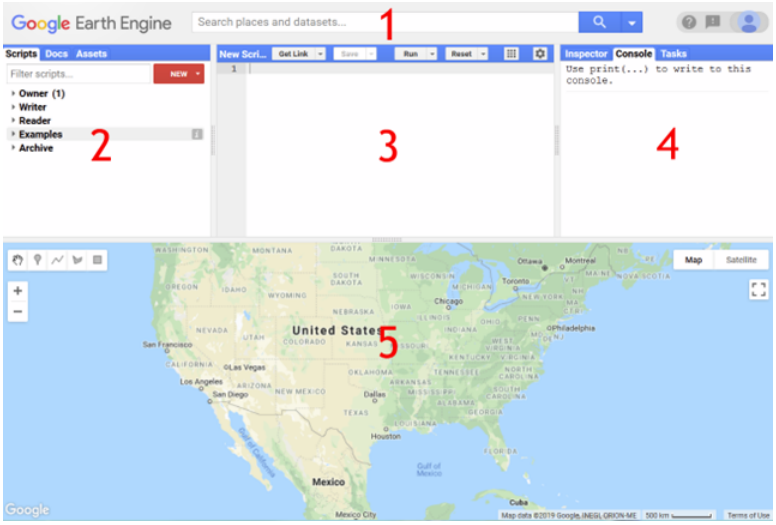

FIGURE D.1: Google Earth Engine interface. Source: odata.nrel.colostate.edu.

of satellite data useful to monitor volcanic activity, like Sentinel2, MODIS, ASTER, and it allows for the easy gathering and analysis of these data. Moreover, GEE has built-in capabilities to allow users to build and use ML models. This fortuitous combination of plentiful satellite data and powerful data analysis has significantly increased the capability to produce accurate forecasting of volcanic hazards during an ongoing eruption.

D.2 Google Colab

Google Colaboratory (Colab) is a Google research project created to help disseminate machine learning education and research. A Jupyter Notebook is an interactive laboratory notebook, which includes notes, data and code to manipulate data. Google Colab hosts Jupyter notebooks, without requiring any local software installation, and enables to write and execute Python code entirely on the cloud. Colab enables you to utilize the power of Python libraries for data analysis and visualization, but mostly it is particularly suitable to

develop ML and DL models. It overcomes the computational limitations related to local machines, because, running the code on the cloud, it is possible to train large scale ML models, even without a powerful machine or a high-speed internet connection.
Colab can be used to interact with GEE through the Earth Engine Python API. This library facilitates the integration of GEE functionalities with python programming, enabling users to access Earth Engine data, perform geospatial analysis and visualize results. To use Earth Engine, users need to authenticate their Colab notebook, in order to link the Google account with Earth Engine and to access the geospatial data and services.
In GEE, there are certain limitations and drawbacks related to the availability of datasets. While GEE provides access to a wide range of geospatial data, it may not cover all datasets that users might need. The process of retrieving data from these space agency portals is quite slow and time-consuming, because the user has to download the products manually, selecting them one by one. By using Google Colab, it is possible to speed up and automate the data downloading processes, using Python API.

Bibliography

[1] A. Robock, "Volcanic eruptions and climate," *Reviews of geophysics*, vol. 38, no. 2, pp. 191–219, 2000.

[2] M. M. Chim, T. J. Aubry, C. Smith, and A. Schmidt, "Neglecting future sporadic volcanic eruptions underestimates climate uncertainty," *Communications Earth & Environment*, vol. 6, no. 1, p. 236, 2025.

[3] L. M. Polvani, A. Banerjee, and A. Schmidt, "Northern hemisphere continental winter warming following the 1991 mt. pinatubo eruption: Reconciling models and observations," *Atmospheric Chemistry and Physics*, vol. 19, no. 9, pp. 6351–6366, 2019.

[4] P. J. Wallace, "From mantle to atmosphere: Magma degassing, explosive eruptions, and volcanic volatile budgets," in *Developments in volcanology*, vol. 5, Elsevier, 2003, pp. 105–127.

[5] M. Edmonds, E. Liu, and K. Cashman, "Open-vent volcanoes fuelled by depth-integrated magma degassing," *Bulletin of Volcanology*, vol. 84, no. 3, p. 28, 2022.

[6] A. Harris, *Thermal remote sensing of active volcanoes: a user's manual*. Cambridge university press, 2013.

[7] Y. Li, C. Pedersen, J. Dykema, *et al.*, "In situ measurements of perturbations to stratospheric aerosol and modeled ozone and radiative impacts following the 2021 la soufrière eruption," *Atmospheric Chemistry and Physics*, vol. 23, no. 24, pp. 15 351–15 364, 2023.

[8] R. Benoit, H. Vernier, J.-P. Vernier, *et al.*, "Organics in the stratosphere: New insights from weather balloon flights conducted in france," *npj climate and atmospheric science*, vol. 8, no. 1, p. 221, 2025.

[9] M. McCormick, P. Hamill, T. Pepin, W. Chu, T. Swissler, and L. McMaster, "Satellite studies of the stratospheric aerosol," *Bulletin of the American Meteorological Society*, vol. 60, no. 9, pp. 1038–1047, 1979.

[10] N. Anantrasirichai, J. Biggs, F. Albino, P. Hill, and D. Bull, "Application of machine learning to classification of volcanic deformation in routinely generated insar data," *Journal of Geophysical Research: Solid Earth*, vol. 123, no. 8, pp. 6592–6606, 2018.

[11] E. Amato, C. Corradino, F. Torrisi, and C. Del Negro, "A deep convolutional neural network for detecting volcanic thermal anomalies from satellite images," *Remote Sensing*, vol. 15, no. 15, p. 3718, 2023.

[12] C. Corradino, E. Amato, F. Torrisi, and C. Del Negro, "Data-driven random forest models for detecting volcanic hot spots in sentinel-2 msi images," *Remote Sensing*, vol. 14, no. 17, p. 4370, 2022.

[13] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[14] G. Bonaccorso, *Machine Learning Algorithms: Popular algorithms for data science and machine learning*. Packt Publishing Ltd, 2018.

[15] N. Anantrasirichai and D. Bull, "Artificial intelligence in the creative industries: A review," *Artificial intelligence review*, vol. 55, no. 1, pp. 589–656, 2022.

[16] W. F. Wiggins and A. S. Tejani, "On the opportunities and risks of foundation models for natural language processing in radiology," *Radiology: Artificial Intelligence*, vol. 4, no. 4, e220119, 2022.

[17] A. Kolides, A. Nawaz, A. Rathor, *et al.*, "Artificial intelligence foundation and pre-trained models: Fundamentals, applications, opportunities, and social impacts," *Simulation Modelling Practice and Theory*, vol. 126, p. 102 754, 2023.

[18] M. Schwabe, L. Pastori, I. de Vega, *et al.*, "Opportunities and challenges of quantum computing for climate modelling," *arXiv preprint arXiv:2502.10488*, 2025.

[19] A. Sebastianelli, D. A. Zaidenberg, D. Spiller, B. Le Saux, and S. L. Ullo, "On circuit-based hybrid quantum neural networks for remote sensing imagery classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 565–580, 2021.

[20] A. Sebastianelli, F. Mauro, G. Ciabatti, *et al.*, "Quanv4eo: Empowering earth observation by means of quanvolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

[21] P. Baker, "Geological aspects of volcano prediction," *Journal of the Geological Society*, vol. 136, no. 3, pp. 341–345, 1979.

[22] J. P. Lockwood, R. W. Hazlett, and S. de la Cruz-Reyna, *Volcanoes: global perspectives*. John Wiley & Sons, 2022.

[23] M. Ort and N. Riggs, *Volcanology: Processes, deposits, geology and resources, volcanology: Processes, deposits, geology and resources, by raf cas, g. giordano, and jv wright, springer textbooks in earth sciences, geography, and environment, 2024, 1833 pp., https://doi. org/10.1007/978-3-319-66613-6*, 2025.

[24] A. J. Sutton, T. Elias, *et al.*, "One hundred volatile years of volcanic gas studies at the hawaiian volcano observatory," *Characteristics of Hawaiian volcanoes*, vol. 1801, pp. 295–320, 2014.

[25] H. Shinohara, "Excess degassing from volcanoes and its role on eruptive and intrusive activity," *Reviews of Geophysics*, vol. 46, no. 4, 2008.

[26] T. Girona, V. Realmuto, and P. Lundgren, "Large-scale thermal unrest of volcanoes for years prior to eruption," *Nature Geoscience*, vol. 14, no. 4, pp. 238–241, 2021.

[27] M. S. Ramsey, C. Corradino, J. O. Thompson, and T. N. Leggett, "Statistical retrieval of volcanic activity in long time series orbital data: Implications for forecasting future activity," *Remote Sensing of Environment*, vol. 295, p. 113 704, 2023.

[28] K. Reath, M. Ramsey, J Dehn, and P. Webley, "Predicting eruptions from precursory activity using remote sensing data hybridization," *Journal of Volcanology and Geothermal Research*, vol. 321, pp. 18–30, 2016.

[29] C. Kern, S. Arellano, R. Campion, S. Hidalgo, and R. Kazahaya, *Remote sensing of volcanic gas emissions from the ground, air, and space*, 2023.

[30] A. La Spina, M. Burton, and G. G. Salerno, "Unravelling the processes controlling gas emissions from the central and northeast craters of mt. etna," *Journal of Volcanology and Geothermal Research*, vol. 198, no. 3-4, pp. 368–376, 2010.

[31] Y. Zhan, H. Le Mével, D. C. Roman, T. Girona, and P. M. Gregg, "Modeling deformation, seismicity, and thermal anomalies driven by degassing during the 2005-2006 pre-eruptive unrest of augustine volcano, alaska," *Earth and Planetary Science Letters*, vol. 585, p. 117 524, 2022.

[32] E. Biagioli, G. La Spina, M. Polacci, B. Bonechi, J. E. Romero, and M. Burton, "Numerical modeling integrated with field observations and analytical data of the 2021 cumbre vieja eruption improves understanding of eruption dynamics at mafic volcanoes," *Journal of Geophysical Research: Solid Earth*, vol. 130, no. 6, e2024JB029692, 2025.

[33] M. Toohey, Y. Jia, S. Khanal, and S. Tegtmeier, "Stratospheric residence time and the lifetime of volcanic stratospheric aerosols," *EGUsphere*, vol. 2024, pp. 1–28, 2024.

[34] A. Robock, "Climatic impact of volcanic emissions," *State of the Planet (in prep.)*, 2004.

[35] R. J. Carey, D. A. Butterfield, and M. R. Clark, "Submarine volcanic eruptions and their impacts on hydrothermal systems and biological communities," *Annual Review of Earth and Planetary Sciences*, vol. 53, no. 1, pp. 397–423, 2025.

[36] H. Vömel, S. Evan, and M. Tully, "Water vapor injection into the stratosphere by hunga tonga-hunga ha'apai," *Science*, vol. 377, no. 6613, pp. 1444–1447, 2022.

[37] M. A. Pfeffer, B. Bergsson, S. Barsotti, *et al.*, "Ground-based measurements of the 2014–2015 holuhraun volcanic cloud (iceland)," *Geosciences*, vol. 8, no. 1, p. 29, 2018.

[38] E. Carboni, T. A. Mather, A. Schmidt, *et al.*, "Satellite-derived sulfur dioxide (so 2) emissions from the 2014–2015 holuhraun eruption (iceland)," *Atmospheric Chemistry and Physics*, vol. 19, no. 7, pp. 4851–4862, 2019.

[39] E. Tam, R. Miike, S. Labrenz, *et al.*, "Volcanic air pollution over the island of hawai'i: Emissions, dispersal, and composition. association with respiratory symptoms and lung function in hawai'i island school children," *Environment international*, vol. 92, pp. 543–552, 2016.

[40] R. Del Moral and S. Y. Grishin, "Volcanic disturbances and ecosystem recovery," *Ecosystems of the world*, pp. 137–160, 1999.

[41] D. Sparice, V. Amoretti, F. Galadini, *et al.*, "A novel view of the destruction of pompeii during the 79 ce eruption of vesuvius (italy): Syn-eruptive earthquakes as an additional cause of building collapse and deaths," *Frontiers in Earth Science*, vol. 12, p. 1 386 960, 2024.

[42] G. Luongo, A. Perrotta, C. Scarpati, E. De Carolis, G. Patricelli, and A. Ciarallo, "Impact of the ad 79 explosive eruption on pompeii, ii. causes of death of the inhabitants inferred by stratigraphic analysis and areal distribution of the human casualties," *Journal of Volcanology and Geothermal Research*, vol. 126, no. 3-4, pp. 169–200, 2003.

[43] R. D'Arrigo, R. Seager, J. E. Smerdon, A. N. LeGrande, and E. R. Cook, "The anomalous winter of 1783–1784: Was the laki eruption or an analog of the 2009–2010 winter to blame?" *Geophysical Research Letters*, vol. 38, no. 5, 2011.

[44] M. Mazzocchi, F. Hansstein, and M. Ragona, "The 2010 volcanic ash cloud and its financial impact on the european airline industry," in *CESifo Forum*, München: ifo Institut für Wirtschaftsforschung an der Universität München, vol. 11, 2010, pp. 92–100.

[45] B. Zambri, A. Robock, M. J. Mills, and A. Schmidt, "Modeling the 1783–1784 laki eruption in iceland: 1. aerosol evolution and global stratospheric circulation impacts," *Journal of Geophysical Research: Atmospheres*, vol. 124, no. 13, pp. 6750–6769, 2019.

[46] B. J. Soden, R. T. Wetherald, G. L. Stenchikov, and A. Robock, "Global cooling after the eruption of mount pinatubo: A test of climate feedback by water vapor," *science*, vol. 296, no. 5568, pp. 727–730, 2002.

[47] J.-P. Guo, X.-Y. Zhang, Y.-R. Wu, *et al.*, "Spatio-temporal variation trends of satellite-based aerosol optical depth in china during 1980–2008," *Atmospheric Environment*, vol. 45, no. 37, pp. 6802–6811, 2011.

[48] M. Chin, T Diehl, Q Tan, *et al.*, "Multi-decadal aerosol variations from 1980 to 2009: A perspective from observations and a global model," *Atmospheric Chemistry and Physics*, vol. 14, no. 7, pp. 3657–3690, 2014.

[49] N. Butchart, "The brewer-dobson circulation," *Reviews of geophysics*, vol. 52, no. 2, pp. 157–184, 2014.

[50] S. M. Andersson, B. G. Martinsson, J.-P. Vernier, *et al.*, "Significant radiative impact of volcanic aerosol in the lowermost stratosphere," *Nature communications*, vol. 6, no. 1, p. 7692, 2015.

[51] Q. Zhao, L. Yu, Z. Du, *et al.*, "An overview of the applications of earth observation satellite data: Impacts and future trends," *Remote Sensing*, vol. 14, no. 8, p. 1863, 2022.

[52] S. Conti, "Advancing earth observation with a multi-modal remote sensing foundation model: Remote sensing," *Nature Reviews Electrical Engineering*, pp. 1–1, 2025.

[53] P. Kansakar and F. Hossain, "A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth," *Space Policy*, vol. 36, pp. 46–54, 2016.

[54] R. Mota, J. M. Pacheco, A. Pimentel, and A. Gil, "Monitoring volcanic plumes and clouds using remote sensing: A systematic review," *Remote Sensing*, vol. 16, no. 10, p. 1789, 2024.

[55] D. M. Pyle, T. A. Mather, and J. Biggs, "Remote sensing of volcanoes and volcanic processes: Integrating observation and modelling–introduction," 2013.

[56] K. Reath, M. Pritchard, M Poland, *et al.*, "Thermal, deformation, and degassing remote sensing time series (ce 2000–2017) at the 47 most active volcanoes in latin america: Implications for volcanic systems," *Journal of Geophysical Research: Solid Earth*, vol. 124, no. 1, pp. 195–218, 2019.

[57] G. S. Di Bella, C. Corradino, S. Cariello, F. Torrisi, and C. Del Negro, "Advancing volcanic activity monitoring: A near-real-time approach with remote sensing data fusion for radiative power estimation," *Remote Sensing*, vol. 16, no. 16, p. 2879, 2024.

[58] D Pieri and M Abrams, "Aster watches the world's volcanoes: A new paradigm for volcanological observations from orbit," *Journal of Volcanology and Geothermal Research*, vol. 135, no. 1-2, pp. 13–28, 2004.

[59] F Albino, J Biggs, C Yu, and Z Li, "Automated methods for detecting volcanic deformation using sentinel-1 insar time series illustrated by the 2017–2018 unrest at agung, indonesia," *Journal of Geophysical Research: Solid Earth*, vol. 125, no. 2, e2019JB017908, 2020.

[60] S. Ebmeier, B. Andrews, M. Araya, *et al.*, "Synthesis of global satellite observations of magmatic and volcanic deformation: Implications for volcano monitoring & the lateral extent of magmatic domains," *Journal of Applied Volcanology*, vol. 7, no. 1, p. 2, 2018.

[61] C. Corradino, M. S. Ramsey, S. Pailot-Bonnétat, A. J. Harris, and C. Del Negro, "Detection of subtle thermal anomalies: Deep learning applied to the aster global volcano dataset," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.

[62] R. Wright, L. Flynn, H. Garbeil, A. Harris, and E. Pilger, "Automated volcanic eruption detection using modis," *Remote sensing of environment*, vol. 82, no. 1, pp. 135–155, 2002.

[63] D. Coppola, M. Laiolo, D Delle Donne, M Ripepe, C. Cigolini, *et al.*, "Mirova (middle infrared observation of volcanic activity)," 2013.

[64] C. Corradino, A. B. Malaguti, M. S. Ramsey, and C. Del Negro, "Quantitative assessment of volcanic thermal activity from space using an isolation forest machine learning algorithm," *Remote Sensing*, vol. 16, no. 11, p. 2001, 2024.

[65] D. Coppola, M. Laiolo, C. Cigolini, *et al.*, "Thermal remote sensing for global volcano monitoring: Experiences from the mirova system," *Frontiers in Earth Science*, vol. 7, p. 362, 2020.

[66] C. Spinetti, F. Mazzarini, R. Casacchia, *et al.*, "Spectral properties of volcanic materials from hyperspectral field and satellite data compared with lidar data at mt. etna," *International journal of applied earth observation and geoinformation*, vol. 11, no. 2, pp. 142–155, 2009.

[67] N. Theys, P. Hedelt, I. De Smedt, *et al.*, "Global monitoring of volcanic so2 degassing with unprecedented resolution from tropomi onboard sentinel-5 precursor," *Scientific reports*, vol. 9, no. 1, p. 2643, 2019.

[68] V. J. Flower, T. Oommen, and S. A. Carn, "Improving global detection of volcanic eruptions using the ozone monitoring instrument (omi)," *Atmospheric Measurement Techniques*, vol. 9, no. 11, pp. 5487–5498, 2016.

[69] C. Corradino, P. Jouve, A. La Spina, and C. Del Negro, "Monitoring earth's atmosphere with sentinel-5 tropomi and artificial intelligence: Quantifying volcanic so2 emissions," *Remote Sensing of Environment*, vol. 315, p. 114 463, 2024.

[70] E. Dualeh and J Biggs, "Separating magmatic and hydrothermal deformation using insar timeseries: Independent component analysis at corbetti caldera, ethiopia," *Journal of Geophysical Research: Solid Earth*, vol. 130, no. 4, e2024JB030974, 2025.

[71] J. Biggs, N. Anantrasirichai, K. Anderson, *et al.*, "Rapid advances in volcano monitoring driven by the first decade of sentinel-1 observations," *Available at SSRN 5196731*,

[72] A. Hooper, M. Gaddes, C. N. Lizama, *et al.*, "Machine learning for volcano deformation: Detection, interpretation and forecasting," Copernicus Meetings, Tech. Rep., 2025.

[73] E. Ferrentino, C. Bignami, F. Nunziata, S. Stramondo, and M. Migliaccio, "On the ability of dual-polarimetric sar measurements to observe lava flows under different volcanic environments," *International Journal of Applied Earth Observation and Geoinformation*, vol. 123, p. 103 471, 2023.

[74] E. Ferrentino, C. Bignami, L. Beccaro, and M. Polcari, "Dual-polarimetric sar data to detect volcanic activities," in *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2024, pp. 3976–3979.

[75] T. Liu, J. Zhang, G. Gao, J. Yang, and A. Marino, "Cfar ship detection in polarimetric synthetic aperture radar images based on whitening filter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 1, pp. 58–81, 2019.

[76] C. Kidd, V. Levizzani, and P. Bauer, "A review of satellite meteorology and climatology at the start of the twenty-first century," *Progress in Physical Geography*, vol. 33, no. 4, pp. 474–489, 2009.

[77] T. Holloway, D. Miller, S. Anenberg, *et al.*, "Satellite monitoring for air quality and health," *Annual review of biomedical data science*, vol. 4, no. 1, pp. 417–447, 2021.

[78] K. Hu, X. Feng, Q. Zhang, *et al.*, "Review of satellite remote sensing of carbon dioxide inversion and assimilation," *Remote Sensing*, vol. 16, no. 18, p. 3394, 2024.

[79] S. Carn, L. Clarisse, and A. J. Prata, "Multi-decadal satellite measurements of global volcanic degassing," *Journal of Volcanology and Geothermal Research*, vol. 311, pp. 99–134, 2016.

[80] L. Clarisse, D. Hurtmans, C. Clerbaux, J. Hadji-Lazaro, Y. Ngadi, and P.-F. Coheur, "Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (iasi)," *Atmospheric Measurement Techniques*, vol. 5, no. 3, pp. 581–594, 2012.

[81] N. A. Krotkov, S. A. Carn, A. J. Krueger, P. K. Bhartia, and K. Yang, "Band residual difference algorithm for retrieval of so/sub 2/from the aura ozone monitoring instrument (omi)," *IEEE Transactions on geoscience and remote sensing*, vol. 44, no. 5, pp. 1259–1266, 2006.

[82] D. Ridley, S. Solomon, J. Barnes, *et al.*, "Total volcanic stratospheric aerosol optical depths and implications for global climate change," *Geophysical Research Letters*, vol. 41, no. 22, pp. 7763–7769, 2014.

[83] S. A. Carn, A. J. Krueger, N. A. Krotkov, K. Yang, and K. Evans, "Tracking volcanic sulfur dioxide clouds for aviation hazard mitigation," *Natural hazards*, vol. 51, no. 2, pp. 325–343, 2009.

[84] C. Clerbaux, A. Boynard, L. Clarisse, *et al.*, "Monitoring of atmospheric composition using the thermal infrared iasi/metop sounder," *Atmospheric Chemistry and Physics*, vol. 9, no. 16, pp. 6041–6054, 2009.

[85] V. Fioletov, C. A. McLinden, D. Griffin, *et al.*, "Anthropogenic and volcanic point source so 2 emissions derived from tropomi on board sentinel-5 precursor: First results," *Atmospheric Chemistry and Physics*, vol. 20, no. 9, pp. 5591–5607, 2020.

[86] K. L. Chan, P. Valks, K.-P. Heue, *et al.*, "Global ozone monitoring experiment-2 (gome-2) daily and monthly level-3 products of atmospheric trace gas columns," *Earth System Science Data*, vol. 15, no. 4, pp. 1831–1870, 2023.

[87] L. Clarisse, P.-F. Coheur, A. J. Prata, *et al.*, "Tracking and quantifying volcanic so 2 with iasi, the september 2007 eruption at jebel at tair," *Atmospheric Chemistry and Physics*, vol. 8, no. 24, pp. 7723–7734, 2008.

[88] E. Carboni, R. G. Grainger, T. A. Mather, *et al.*, "The vertical distribution of volcanic so 2 plumes measured by iasi," *Atmospheric Chemistry and Physics*, vol. 16, no. 7, pp. 4343–4367, 2016.

[89] J.-P. Vernier, L. W. Thomason, J.-P. Pommereau, *et al.*, "Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade," *Geophysical Research Letters*, vol. 38, no. 12, 2011.

[90] S Self, R Gertisser, T. Thordarson, M. Rampino, and J. Wolff, "Magma volume, volatile emissions, and stratospheric aerosols from the 1815 eruption of tambora," *Geophysical Research Letters*, vol. 31, no. 20, 2004.

[91] M. P. McCormick, L. W. Thomason, and C. R. Trepte, "Atmospheric effects of the mt pinatubo eruption," *Nature*, vol. 373, no. 6513, pp. 399–404, 1995.

[92] P Zoogman, X Liu, R. Suleiman, *et al.*, "Tropospheric emissions: Monitoring of pollution (tempo)," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 186, pp. 17–39, 2017.

[93] E. Malina, K. W. Bowman, V. Kantchev, *et al.*, "Joint spectral retrievals of ozone with suomi npp cris augmented by s5p/tropomi," *Atmospheric Measurement Techniques*, vol. 17, no. 17, pp. 5341–5371, 2024.

[94] A. J. Illingworth, H. Barker, A Beljaars, *et al.*, "The earthcare satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation," *Bulletin of the American Meteorological Society*, vol. 96, no. 8, pp. 1311–1332, 2015.

[95] L. Guanter, H. Kaufmann, K. Segl, *et al.*, "The enmap spaceborne imaging spectroscopy mission for earth observation," *Remote Sensing*, vol. 7, no. 7, pp. 8830–8857, 2015.

[96] P. Rosen, S. Hensley, S. Shaffer, *et al.*, "The nasa-isro sar (nisar) mission dual-band radar instrument preliminary design," in *2017 IEEE international geoscience and remote sensing symposium (IGARSS)*, IEEE, 2017, pp. 3832–3835.

[97] S. Valade, A. Ley, F. Massimetti, *et al.*, "Towards global volcano monitoring using multisensor sentinel missions and artificial intelligence: The mounts monitoring system," *Remote Sensing*, vol. 11, no. 13, p. 1528, 2019.

[98] F. Marchese and N. Genzano, "Global volcano monitoring through the normalized hotspot indices (nhi) system," *Journal of the Geological Society*, vol. 180, no. 1, jgs2022–014, 2023.

[99] S. Cariello, A. B. Malaguti, C. Corradino, and C. Del Negro, "V-star: A cloud-based tool for satellite detection and mapping of volcanic thermal anomalies," *GeoHazards*, vol. 6, no. 2, p. 24, 2025.

[100] F. Pardini, S. Corradini, A. Costa, *et al.*, "Ensemble-based data assimilation of volcanic ash clouds from satellite observations: Application to the 24 december 2018 mt. etna explosive eruption," *Atmosphere*, vol. 11, no. 4, p. 359, 2020.

[101] G. Fu, A. Heemink, S. Lu, A. Segers, K. Weber, and H.-X. Lin, "Model-based aviation advice on distal volcanic ash clouds by assimilating aircraft in situ measurements," *Atmospheric Chemistry and Physics*, vol. 16, no. 14, pp. 9189–9200, 2016.

[102] A. Capponi, N. J. Harvey, H. F. Dacre, *et al.*, "Refining an ensemble of volcanic ash forecasts using satellite retrievals: Raikoke 2019," *Atmospheric Chemistry and Physics*, vol. 22, no. 9, pp. 6115–6134, 2022.

[103] S Scollo, A Folch, and A Costa, "A parametric and comparative study of different tephra fallout models," *Journal of Volcanology and Geothermal Research*, vol. 176, no. 2, pp. 199–211, 2008.

[104] M. Woodhouse, A. Hogg, J. Phillips, and R. Sparks, "Interaction between volcanic plumes and wind during the 2010 eyjafjallajökull eruption, iceland," *Journal of Geophysical Research: Solid Earth*, vol. 118, no. 1, pp. 92–109, 2013.

[105] C. Bonadonna, L. Connor, C. B. Connor, and L. M. Courtland, "Tephra2," 2010.

[106] H. F. Schwaiger, R. P. Denlinger, and L. G. Mastin, "Ash3d: A finite-volume, conservative numerical model for ash transport and tephra deposition," *Journal of Geophysical Research: Solid Earth*, vol. 117, no. B4, 2012.

[107] G. Chen, Q. Xiong, P. J. Morris, E. G. Paterson, A. Sergeev, and Y Wang, "Openfoam for computational fluid dynamics," *Notices of the AMS*, vol. 61, no. 4, pp. 354–363, 2014.

[108] F. Legros, "Minimum volume of a tephra fallout deposit estimated from a single isopach," *Journal of volcanology and geothermal research*, vol. 96, no. 1-2, pp. 25–32, 2000.

[109] V. Zago, G. Bilotta, A. Hérault, *et al.*, "Semi-implicit 3d sph on gpu for lava flows," *Journal of Computational Physics*, vol. 375, pp. 854–870, 2018.

[110] V. Zago, G. Bilotta, A. Cappello, *et al.*, "Sph model for the simulation of lava-buildings interactions," in *EGU General Assembly Conference Abstracts*, 2021, EGU21–14 717.

[111] C. Del Negro, L. Fortuna, A. Herault, and A. Vicari, "Simulations of the 2004 lava flow at etna volcano using the magflow cellular automata model," *Bulletin of Volcanology*, vol. 70, no. 7, pp. 805–812, 2008.

[112] G. M. Crisci, R. Rongo, S. Di Gregorio, and W. Spataro, "The simulation model sciara: The 1991 and 2001 lava flows at mount etna," *Journal of Volcanology and Geothermal Research*, vol. 132, no. 2-3, pp. 253–267, 2004.

[113] L. Cui, X. Song, and G. Zhong, "Comparative analysis of three methods for hysplit atmospheric trajectories clustering," *Atmosphere*, vol. 12, no. 6, p. 698, 2021.

[114] M. Kanakidou, J. Seinfeld, S. Pandis, *et al.*, "Organic aerosol and global climate modelling: A review," *Atmospheric Chemistry and Physics*, vol. 5, no. 4, pp. 1053–1123, 2005.

[115] N. S. Diffenbaugh and F. Giorgi, "Climate change hotspots in the cmip5 global climate model ensemble," *Climatic change*, vol. 114, no. 3, pp. 813–822, 2012.

[116] R. Laprise, "Regional climate modelling," *Journal of computational physics*, vol. 227, no. 7, pp. 3641–3666, 2008.

[117] M. Rummukainen, "State-of-the-art with regional climate models," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 1, no. 1, pp. 82–96, 2010.

[118] G. A. Meehl, G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer, "The coupled model intercomparison project (cmip)," *Bulletin of the American Meteorological Society*, vol. 81, no. 2, pp. 313–318, 2000.

[119] V Buchard, C. Randles, A. Da Silva, *et al.*, "The merra-2 aerosol reanalysis, 1980 onward. part ii: Evaluation and case studies," *Journal of climate*, vol. 30, no. 17, pp. 6851–6872, 2017.

[120] S. Kobayashi, Y. Ota, Y. Harada, *et al.*, "The jra-55 reanalysis: General specifications and basic characteristics," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 93, no. 1, pp. 5–48, 2015.

[121] H. Hersbach, B. Bell, P. Berrisford, *et al.*, "The era5 global reanalysis," *Quarterly journal of the royal meteorological society*, vol. 146, no. 730, pp. 1999–2049, 2020.

[122] E. O. Pyzer-Knapp, J. W. Pitera, P. W. Staar, *et al.*, "Accelerating materials discovery using artificial intelligence, high performance computing and robotics," *npj Computational Materials*, vol. 8, no. 1, p. 84, 2022.

[123] M. Reichstein, G. Camps-Valls, B. Stevens, *et al.*, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.

[124] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geoscience frontiers*, vol. 7, no. 1, pp. 3–10, 2016.

[125] C. Caudron, T. Girona, A. Jolly, *et al.*, "A quest for unrest in multiparameter observations at whakaari/white island volcano, new zealand 2007–2018," *Earth, Planets and Space*, vol. 73, no. 1, pp. 1–21, 2021.

[126] F. Torrisi, C. Corradino, S. Cariello, and C. Del Negro, "Enhancing detection of volcanic ash clouds from space with convolutional neural networks," *Journal of Volcanology and Geothermal Research*, vol. 448, p. 108 046, 2024.

[127] P. Rey-Devesa, J. Carthy, M. Titos, J. Prudencio, J. M. Ibáñez, and C. Benítez, "Universal machine learning approach to volcanic eruption forecasting using seismic features," *Frontiers in Earth Science*, vol. 12, p. 1 342 468, 2024.

[128] R. Wright and L. P. Flynn, "Space-based estimate of the volcanic heat flux into the atmosphere during 2001 and 2002," *Geology*, vol. 32, no. 3, pp. 189–192, 2004.

[129] T. Beker, H. Ansari, S. Montazeri, Q. Song, and X. X. Zhu, "Deep learning for subtle volcanic deformation detection with insar data in central volcanic zone," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–20, 2023.

[130] M. Gaddes, A. Hooper, and F. Albino, "Simultaneous classification and location of volcanic deformation in sar interferograms using a convolutional neural network," *Earth and Space Science*, vol. 11, no. 6, e2024EA003679, 2024.

[131] A. Piscini, M Picchiani, M Chini, *et al.*, "A neural network approach for the simultaneous retrieval of volcanic ash parameters and so 2 using modis data," *Atmospheric Measurement Techniques*, vol. 7, no. 12, pp. 4023–4047, 2014.

[132] J. Pathak, S. Subramanian, P. Harrington, *et al.*, "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators," *arXiv preprint arXiv:2202.11214*, 2022.

[133] A. Sether, "Cloud computing benefits," *Available at SSRN 2781593*, 2016.

[134] N. Mungoli, "Adaptive feature fusion: Enhancing generalization in deep learning models," *arXiv preprint arXiv:2304.03290*, 2023.

[135] N. Rane, S. P. Choudhary, and J. Rane, "Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions," *Studies in Medical and Health Sciences*, vol. 1, no. 2, pp. 18–41, 2024.

[136] G. Ramesh *et al.*, "Advanced water quality assessment with iot wsn and multivariate machine learning techniques," in *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS)*, IEEE, 2025, pp. 1–6.

[137] D. Rolnick, P. L. Donti, L. H. Kaack, *et al.*, "Tackling climate change with machine learning," *ACM Computing Surveys (CSUR)*, vol. 55, no. 2, pp. 1–96, 2022.

[138] R. Bommasani, "On the opportunities and risks of foundation models," *arXiv preprint arXiv:2108.07258*, 2021.

[139] Y. Cong, S. Khanna, C. Meng, *et al.*, "Satmae: Pre-training transformers for temporal and multi-spectral satellite imagery," *Advances in Neural Information Processing Systems*, vol. 35, pp. 197–211, 2022.

[140] K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, "Accurate medium-range global weather forecasting with 3d neural networks," *Nature*, vol. 619, no. 7970, pp. 533–538, 2023.

[141] F. Gascon, C. Bouzinac, O. Thépaut, *et al.*, "Copernicus sentinel-2a calibration and products validation status," *Remote Sensing*, vol. 9, no. 6, p. 584, 2017.

[142] J. Li, Y. Li, L. He, J. Chen, and A. Plaza, "Spatio-temporal fusion for remote sensing data: An overview and new benchmark," *Science China Information Sciences*, vol. 63, no. 4, p. 140 301, 2020.

[143] N. Joshi, M. Baumann, A. Ehammer, *et al.*, *A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. remote sens. 8, 70*, 2016.

[144] F. Samadzadegan, A. Toosi, and F. Dadrass Javan, "A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: Current status and prospects," *International journal of remote sensing*, vol. 46, no. 3, pp. 1327–1402, 2025.

[145] S. Salcedo-Sanz, P. Ghamisi, M. Piles, *et al.*, "Machine learning information fusion in earth observation: A comprehensive review of methods, applications and data sources," *Information Fusion*, vol. 63, pp. 256–272, 2020.

[146] J. Li, D. Hong, L. Gao, *et al.*, "Deep learning in multimodal remote sensing data fusion: A comprehensive review," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102 926, 2022.

[147] S. Murphy, R Wright, C Oppenheimer, and C. Souza Filho, "Modis and aster synergy for characterizing thermal volcanic activity," *Remote sensing of environment*, vol. 131, pp. 195–205, 2013.

[148] S. Carn, K Yang, A. Prata, and N. Krotkov, "Extending the long-term record of volcanic so2 emissions with the ozone mapping and profiler suite nadir mapper," *Geophysical Research Letters*, vol. 42, no. 3, pp. 925–932, 2015.

[149] E. Weisz and W. P. Menzel, "Monitoring the 2021 cumbre vieja volcanic eruption using satellite multisensor data fusion," *Journal of Geophysical Research: Atmospheres*, vol. 128, no. 2, e2022JD037926, 2023.

[150] X. Zhu, F. Cai, J. Tian, and T. K.-A. Williams, "Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions," *Remote Sensing*, vol. 10, no. 4, p. 527, 2018.

[151] P. Ghamisi, N. Yokoya, J. Li, *et al.*, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, 2018.

[152] A. Karpatne, G. Atluri, J. H. Faghmous, *et al.*, "Theory-guided data science: A new paradigm for scientific discovery from data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 10, pp. 2318–2331, 2017.

[153] T. Schneider, S. Lan, A. Stuart, and J. Teixeira, "Learned representations for climate and earth system science: Challenges and opportunities," *Nature Reviews Earth & Environment*, vol. 3, no. 8, pp. 634–651, 2022.

[154] C. M. Gevaert, "Explainable ai for earth observation: A review including societal and regulatory perspectives," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102 869, 2022.

[155] Z. Zhao, L. Alzubaidi, J. Zhang, Y. Duan, and Y. Gu, "A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations," *Expert Systems with Applications*, vol. 242, p. 122 807, 2024.

[156] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[157] T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," in *The elements of statistical learning: Data mining, inference, and prediction*, Springer, 2008, pp. 485–585.

[158] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[159] G. Sumbul, M. Charfuelan, B. Demir, and V. Markl, "Bigearthnet: A large-scale benchmark archive for remote sensing image understanding," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 5901–5904.

[160] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "Sen12ms–a curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion," *arXiv preprint arXiv:1906.07789*, 2019.

[161] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote sensing of environment*, vol. 204, pp. 509–523, 2018.

[162] C. Pelletier, G. I. Webb, and F. Petitjean, "Temporal convolutional neural network for the classification of satellite image time series," *Remote Sensing*, vol. 11, no. 5, p. 523, 2019.

[163] M. Rußwurm, C. Pelletier, M. Zollner, S. Lefèvre, and M. Körner, "Breizhcrops: A time series dataset for crop type mapping," *arXiv preprint arXiv:1905.11893*, 2019.

[164] H. Yin, C. Ma, L. Weng, M. Xia, and H. Lin, "Bitemporal remote sensing image change detection network based on siamese-attention feedback architecture," *Remote Sensing*, vol. 15, no. 17, p. 4186, 2023.

[165] S. Hislop, S. Jones, M. Soto-Berelov, A. Skidmore, A. Haywood, and T. H. Nguyen, "Using landsat spectral indices in time-series to assess wildfire disturbance and recovery," *Remote sensing*, vol. 10, no. 3, p. 460, 2018.

[166] C. Révillion, P. Mouquet, J. Commins, *et al.*, "Sen2chain: An open-source toolbox for processing sentinel-2 satellite images and producing time-series of spectral indices," *arXiv preprint arXiv:2404.04305*, 2024.

[167] Z. Zhong, B. He, L. Guo, and Y. Zhang, "Performance of various forms of the palmer drought severity index in china from 1961 to 2013," *Journal of Hydrometeorology*, vol. 20, no. 9, pp. 1867–1885, 2019.

[168] A. Xiao, W. Xuan, J. Wang, *et al.*, "Foundation models for remote sensing and earth observation: A survey," *IEEE Geoscience and Remote Sensing Magazine*, 2025.

[169] D. Hong, B. Zhang, X. Li, *et al.*, "Spectralgpt: Spectral remote sensing foundation model," *arXiv preprint arXiv:2311.07113*, 2023.

[170] J. Jakubik, S. Roy, C. Phillips, *et al.*, "Foundation models for generalist geospatial artificial intelligence," *arXiv preprint arXiv:2310.18660*, 2023.

[171] J. Jakubik, F. Yang, B. Blumenstiel, *et al.*, "Terramind: Large-scale generative multimodality for earth observation," *arXiv preprint arXiv:2504.11171*, 2025.

[172] J. Schmude, S. Roy, W. Trojak, *et al.*, "Prithvi wxc: Foundation model for weather and climate," *arXiv preprint arXiv:2409.13598*, 2024.

[173] N. Ravi, V. Gabeur, Y.-T. Hu, *et al.*, "Sam 2: Segment anything in images and videos," *arXiv preprint arXiv:2408.00714*, 2024.

[174] P. Shi, J. Qiu, S. M. D. Abaxi, H. Wei, F. P.-W. Lo, and W. Yuan, "Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation," *Diagnostics*, vol. 13, no. 11, p. 1947, 2023.

[175] G. Sumbul, A. De Wall, T. Kreuziger, *et al.*, "Bigearthnet-mm: A large-scale, multimodal, multilabel benchmark archive for remote sensing image classification and retrieval [software and data sets]," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 3, pp. 174–180, 2021.

[176] R. Karthik, T. S. Vaichole, S. K. Kulkarni, O. Yadav, and F. Khan, "Eff2net: An efficient channel attention-based convolutional neural network for skin disease classification," *Biomedical Signal Processing and Control*, vol. 73, p. 103 406, 2022.

[177] I. Demir, K. Koperski, D. Lindenbaum, *et al.*, "Deepglobe 2018: A challenge to parse the earth through satellite images," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 172–181.

[178] A. Lacoste, N. Lehmann, P. Rodriguez, *et al.*, "Geo-bench: Toward foundation models for earth monitoring," *Advances in Neural Information Processing Systems*, vol. 36, pp. 51 080–51 093, 2023.

[179] N. Papadopoulos, N. I. Bountos, M. Sdraka, A. Karavias, and I. Papoutsis, "Hephaestus minicubes: A global, multi-modal dataset for volcanic unrest monitoring," *arXiv preprint arXiv:2505.17782*, 2025.

[180] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.

[181] P. Saunders-Shultz, *Deep Learning Detection and Quantification of Volcanic Thermal Signals in Infrared Satellite Data*. University of Alaska Fairbanks, 2024.

[182] G. Boudoire, S Calabrese, A Colacicco, *et al.*, "Scientific response to the 2021 eruption of nyiragongo based on the implementation of a participatory monitoring system," *Scientific reports*, vol. 12, no. 1, p. 7488, 2022.

[183] O. Chevrel, A. Harris, A. Peltier, *et al.*, "Volcanic crisis management supported by near real-time lava flow hazard assessment at piton de la fournaise, la réunion," *Volcanica*, vol. 5, no. 2, pp. 313–334, 2022.

[184] Q. Liu, L. Gui, J. Liu, *et al.*, "Multi-satellite detection of long-range transport and transformation of atmospheric emissions from the hunga tonga-hunga ha'apai volcano," *Remote Sensing*, vol. 15, no. 10, p. 2661, 2023.

[185] L. M. Watson, A. M. Iezzi, L. Toney, *et al.*, "Volcano infrasound: Progress and future directions," *Bulletin of Volcanology*, vol. 84, no. 5, p. 44, 2022.

[186] J. Cole-Dai, "Volcanoes and climate," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 1, no. 6, pp. 824–839, 2010.

[187] G. J. Bluth, S. D. Doiron, C. C. Schnetzler, A. J. Krueger, and L. S. Walter, "Global tracking of the so2 clouds from the june, 1991 mount pinatubo eruptions," *Geophysical Research Letters*, vol. 19, no. 2, pp. 151–154, 1992.

[188] G. A. Zielinski, "Climatic impact of volcanic eruptions," *The Scientific World Journal*, vol. 2, no. 1, pp. 869–884, 2002.

[189] C. F. Mass and D. A. Portman, "Major volcanic eruptions and climate: A critical evaluation," *Journal of Climate*, vol. 2, no. 6, pp. 566–593, 1989.

[190] D. A. Portman and D. S. Gutzler, "Explosive volcanic eruptions, the el niño–southern oscillation, and us climate variability," *Journal of climate*, vol. 9, no. 1, pp. 17–33, 1996.

[191] A. D. Wirakusumah and H. Rachmat, "Impact of the 1815 tambora eruption to global climate change," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, vol. 71, 2017, p. 012 007.

[192] L. Stowe, R. Carey, and P. Pellegrino, "Monitoring the mt. pinatubo aerosol layer with noaa/11 avhrr data," *Geophysical research letters*, vol. 19, no. 2, pp. 159–162, 1992.

[193] R. Holasek, S Self, and A. Woods, "Satellite observations and interpretation of the 1991 mount pinatubo eruption plumes," *Journal of Geophysical Research: Solid Earth*, vol. 101, no. B12, pp. 27 635–27 655, 1996.

[194] M. Abrams, L. Glaze, and M. Sheridan, "Monitoring colima volcano, mexico, using satellite data," *Bulletin of Volcanology*, vol. 53, no. 7, pp. 571–574, 1991.

[195] J. Higgins and A. Harris, "Vast: A program to locate and analyse volcanic thermal anomalies automatically from remotely sensed data," *Computers & Geosciences*, vol. 23, no. 6, pp. 627–645, 1997.

[196] V. Tramutoli, C. Filizzola, N. Genzano, and M. Lisi, "Robust satellite techniques for detecting preseismic thermal anomalies," *Pre-earthquake processes: a multidisciplinary approach to earthquake prediction studies*, pp. 241–258, 2018.

[197] C. Corradino, G. Ganci, A. Cappello, G. Bilotta, S. Calvari, and C. Del Negro, "Recognizing eruptions of mount etna through machine learning using multiperspective infrared images," *Remote Sensing*, vol. 12, no. 6, p. 970, 2020.

[198] J. Gama and P. Brazdil, "Cascade generalization," *Machine learning*, vol. 41, no. 3, pp. 315–343, 2000.

[199] S Supreeth and K. K. Patil, "Virtual machine scheduling strategies in cloud computing-a review," *International Journal on Emerging Technologies*, vol. 10, no. 3, pp. 181–188, 2019.

[200] S. Ray and A. De Sarkar, "Execution analysis of load balancing algorithms in cloud computing environment," *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, vol. 2, no. 5, pp. 1–13, 2012.

[201] M. Ripepe, E. Marchetti, D. Delle Donne, *et al.*, "Infrasonic early warning system for explosive eruptions," *Journal of Geophysical Research: Solid Earth*, vol. 123, no. 11, pp. 9570–9585, 2018.

[202] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

[203] H. J. Lee, I. Ullah, W. Wan, Y. Gao, and Z. Fang, "Real-time vehicle make and model recognition with the residual squeezenet architecture," *Sensors*, vol. 19, no. 5, p. 982, 2019.

[204] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[205] J. Huang, X. Lu, L. Chen, H. Sun, S. Wang, and G. Fang, "Accurate identification of pine wood nematode disease with a deep convolution neural network," *Remote Sensing*, vol. 14, no. 4, p. 913, 2022.

[206] W. Lin, Z. Wu, L. Lin, A. Wen, and J. Li, "An ensemble random forest algorithm for insurance big data analysis," *Ieee access*, vol. 5, pp. 16 568–16 575, 2017.

[207] S. Cariello, C. Corradino, F. Torrisi, and C. Del Negro, "Cascading machine learning to monitor volcanic thermal activity using orbital infrared data: From detection to quantitative evaluation," *Remote Sensing*, vol. 16, no. 1, p. 171, 2023.

[208] A. S. Eesa and W. K. Arabo, "A normalization methods for backpropagation: A comparative study," *Science Journal of University of Zakho*, vol. 5, no. 4, pp. 319–323, 2017.

[209] S. Calvari and G. Nunnari, "Comparison between automated and manual detection of lava fountains from fixed monitoring thermal cameras at etna volcano, italy," *Remote Sensing*, vol. 14, no. 10, p. 2392, 2022.

[210] S. Calvari and G. Nunnari, "Reawakening of voragine, the oldest of etna's summit craters: Insights from a recurrent episodic eruptive behavior," *Remote Sensing*, vol. 16, no. 22, p. 4278, 2024.

[211] E. Amato, C. Corradino, F. Torrisi, and C. Del Negro, "Combined use of satellite data and machine learning for detecting, measuring, and monitoring active lava flows at etna volcano," *Authorea Preprints*, 2022.

[212] F. Giudicepietro, S. Calvari, L. D'Auria, *et al.*, "Changes in the eruptive style of stromboli volcano before the 2019 paroxysmal phase discovered through som clustering of seismo-acoustic features compared with camera images and gbinsar data," *Remote Sensing*, vol. 14, no. 5, p. 1287, 2022.

[213] H. N. Lechner and M. D. Rouleau, "Should we stay or should we go now? factors affecting evacuation decisions at pacaya volcano, guatemala," *International Journal of Disaster Risk Reduction*, vol. 40, p. 101 160, 2019.

[214] W. I. Rose, J. L. Palma, R. E. Wolf, and R. O. M. Gomez, "A 50 yr eruption of a basaltic composite cone: Pacaya, guatemala," 2013.

[215] F. Massimetti, D. Coppola, M. Laiolo, S. Valade, C. Cigolini, and M. Ripepe, "Volcanic hot-spot detection using sentinel-2: A comparison with modis–mirova thermal data series," *Remote Sensing*, vol. 12, no. 5, p. 820, 2020.

[216] I. J. Hamling, S. Cevuard, and E. Garaebiti, "Large-scale drainage of a complex magmatic system: Observations from the 2018 eruption of ambrym volcano, vanuatu," *Geophysical Research Letters*, vol. 46, no. 9, pp. 4609–4617, 2019.

[217] T. Shreve, R. Grandin, M. Boichu, *et al.*, "From prodigious volcanic degassing to caldera subsidence and quiescence at ambrym (vanuatu): The influence of regional tectonics," *Scientific Reports*, vol. 9, no. 1, p. 18 868, 2019.

[218] D. Andronico, A. Cannata, G. Di Grazia, and F. Ferrari, "The 1986–2021 paroxysmal episodes at the summit craters of mt. etna: Insights into volcano dynamics and hazard," *Earth-Science Reviews*, vol. 220, p. 103 686, 2021.

[219] E. De Beni, M. Cantarero, M. Neri, and A. Messina, "Lava flows of mt etna, italy: The 2019 eruption within the context of the last two decades (1999–2019)," *Journal of Maps*, vol. 17, no. 3, pp. 65–76, 2021.

[220] A. B. Malaguti, C. Corradino, A. La Spina, S. Branca, and C. Del Negro, "Machine learning insights into the last 400 years of etna lateral eruptions from historical volcanological data," *Geosciences*, vol. 14, no. 11, p. 295, 2024.

[221] S Cariello, C Corradino, and C Del Negro, "How artificial intelligence can enhance monitoring of volcanoes from space," *Il Nuovo Cimento*, vol. 100, p. 47, 2024.

[222] S. Carn, N. Krotkov, B. Fisher, and C Li, "Out of the blue: Volcanic so2 emissions during the 2021–2022 eruptions of hunga tonga—hunga ha'apai (tonga)," *Frontiers in Earth Science*, vol. 10, p. 976 962, 2022.

[223] G. Sumbul, R. G. Cinbis, and S. Aksoy, "Fine-grained object recognition and zero-shot learning in remote sensing imagery," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, IEEE, 2017, pp. 770–779.

[224] Y. Qu, R. K. Baghbaderani, and H. Qi, "Few-shot hyperspectral image classification through multitask transfer learning," in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 2019, pp. 1–5.

[225] S. Liu, Q. Shi, and L. Zhang, "Few-shot hyperspectral image classification with unknown classes using multitask deep learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 6, pp. 5085–5102, 2020.

[226] L. Wang, R. Li, C. Duan, C. Zhang, X. Meng, and S. Fang, "A novel transformer based semantic segmentation scheme for fine-resolution remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[227] Y. Li, P. Zhang, X. Xu, *et al.*, "Few-shot prototype alignment regularization network for document image layout segementation," *Pattern Recognition*, vol. 115, p. 107 882, 2021.

[228] X. Zhang, Y. Liu, Y. Lin, Q. Liao, and Y. Li, "Uv-sam: Adapting segment anything model for urban village identification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, 2024, pp. 22 520–22 528.

[229] S. Pekhale, R. Sathish, S. Basavaraju, and D. Sharma, "Sa-modified: A foundation model-based zero-shot approach for refining noisy land-use land-cover maps," in *2024 IEEE India Geoscience and Remote Sensing Symposium (InGARSS)*, IEEE, 2024, pp. 1–4.

[230] Z. Yan, J. Li, X. Li, *et al.*, "Ringmo-sam: A foundation model for segment anything in multimodal remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–16, 2023.

[231] F. Pardini, M. Burton, M. d. Vitturi, *et al.*, "Retrieval and intercomparison of volcanic so2 injection height and eruption time from satellite maps and ground-based observations," *Journal of Volcanology and Geothermal Research*, vol. 331, pp. 79–91, 2017.

[232] F. Pardini, M. Burton, F. Arzilli, G. La Spina, and M. Polacci, "So2 emissions, plume heights and magmatic processes inferred from satellite data: The 2015 calbuco eruptions," *Journal of Volcanology and Geothermal Research*, vol. 361, pp. 12–24, 2018.

[233] E. Kalnay, M. Kanamitsu, R. Kistler, *et al.*, "The ncep/ncar 40-year reanalysis project," in *Renewable energy*, Routledge, 2018, Vol1_146–Vol1_194.

[234] D. P. Dee, S. Uppala, A. J. Simmons, *et al.*, "The era-interim reanalysis: Configuration and performance of the data assimilation system," *Quarterly Journal of the royal meteorological society*, vol. 137, no. 656, pp. 553–597, 2011.

[235] R. Gelaro, W. McCarty, M. J. Suárez, *et al.*, "The modern-era retrospective analysis for research and applications, version 2 (merra-2)," *Journal of climate*, vol. 30, no. 14, pp. 5419–5454, 2017.

[236] G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, *et al.*, "The twentieth century reanalysis project," *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 654, pp. 1–28, 2011.

[237] A. E. Bourassa, A. Robock, W. J. Randel, *et al.*, "Large volcanic aerosol load in the stratosphere linked to asian monsoon transport," *Science*, vol. 337, no. 6090, pp. 78–81, 2012.

[238] J. Liang and J. Haywood, "Future changes in atmospheric rivers over east asia under stratospheric aerosol intervention," *Atmospheric Chemistry and Physics*, vol. 23, no. 2, pp. 1687–1703, 2023.

[239] M. Van Der Meer, S. de Roda Husman, and S. Lhermitte, "Deep learning regional climate model emulators: A comparison of two downscaling training frameworks," *Journal of Advances in Modeling Earth Systems*, vol. 15, no. 6, e2022MS003593, 2023.

[240] A. Doury, S. Somot, S. Gadat, A. Ribes, and L. Corre, "Regional climate model emulator based on deep learning: Concept and first evaluation of a novel hybrid downscaling approach," *Climate Dynamics*, vol. 60, no. 5, pp. 1751–1779, 2023.