ANNALS OF GEOPHYSICS, 68, 2, V224, 2025; doi:10.4401/ag-9189
OPEN ACCESS

Integrated machine learning approach for volcanic cloud tracking: A Case Study of Etna's Lava Fountains (2020-2022)

Federica Torrisi[*,1,2]

[(1)] Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo, Catania, Italy
[(2)] University of Catania, Department of Electrical, Electronic and Computer Science Engineering, Catania, Italy

Article history: received October 20, 2024; accepted March 14, 2025

Abstract

Between December 2020 and February 2022, Mt. Etna produced extraordinary lava fountains which developed into eruptive columns rising several kilometers above the vent. It is crucial to monitor the volcanic clouds produced during these eruptions to assess their impact on the environment, human health, and aviation. Geostationary satellite missions provide high-frequency thermal infrared data, which are crucial for monitoring volcanic clouds during intense explosive eruptions. However, the large volume of satellite data necessitates automatic and accurate processing algorithms, especially when dealing with global-scale observations every 5 minutes. In this work, a robust machine learning approach is developed to identify and track volcanic clouds using images from the EUMETSAT MSG SEVIRI (Meteosat Second Generation – Spinning Enhanced Visible and InfraRed Imager). This approach combines two distinct machine learning models: a deep learning (DL) model for volcanic cloud detection and a supervised machine learning (ML) model for identifying its primary components. The DL model segments volcanic clouds in SEVIRI images by analyzing both the spatial and spectral intensity data. The supervised ML model is able to distinguish the main components of a volcanic cloud by classifying the pixels as ash-rich, SO_2-rich, or characterized by mixed components. Once an accurate mask of the volcanic cloud is obtained, the volcanic plume height is retrieved from satellite observations for further characterization. This integrated ML approach was applied to characterize the volcanic clouds produced during some of the lava fountains occurred at Etna volcano (Italy) between 2020 and 2022.

Keywords: Volcanic clouds; Geostationary satellite sensors; Machine learning; Convolutional Neural Networks; Etna 2020-2022 lava fountains

1. Introduction

Etna is one of the most active volcanoes in the world, characterized by a frequent explosive activity (Branca and Carlo, 2005; Freret-Lorgeril et al., 2018; Calvari and Nunnari, 2024; Corsaro et al., 2024). At Etna volcano, between December 2020 and February 2022, a total of 66 lava fountains occurred (Calvari and Nunnari, 2022), producing

lava flows and large volcanic clouds that spread over the surrounding areas. Most of these clouds caused disruptions, frequently leading to the closure of Catania airport. Volcanic clouds are produced during explosive volcanic eruptions, which release significant amount of silicate particles and gases, primarily composed of water vapor (H_2O), carbon dioxide (CO_2) and sulphur dioxide (SO_2) (Scollo et al., 2012). These emissions can affect human health and ecosystems (Mather et al., 2003; Durant et al., 2010), and are one of the most important natural sources of pollutants in the atmosphere (Filonchyk et al., 2022). Silicate particles, in particular, can damage aircrafts (Guffanti et al., 2009), crops and infrastructure (Wilson et al., 2014) and also cause respiratory diseases (Gudmundsson, 2011). SO_2, on the other hand, can lead to acid rain and, if it reaches the stratosphere, can alter the aerosol optical depth, potentially causing climatic disturbances that may last for several years following major volcanic eruptions (Pitari and Mancini, 2002). Large volcanic eruptions are not only known as significant hazards to aviation and health but also play a crucial role in the climate's natural variability (Ge et al., 2016; Marshall et al., 2025). Volcanic aerosols are a significant yet unpredictable climate forcing factor, capable of triggering major climatic anomalies in the years following an eruption, with potentially devastating effects on the environment and societies. Volcanic sulfate aerosols can persist in the stratosphere for months or even years, depending on factors such as the altitude of SO_2 injection, total mass loading, latitude, and dispersion patterns (Carn et al., 2016). Given the significant impacts of ash and SO_2, accurate monitoring of the volcanic clouds and their components is essential. Numerous studies have been conducted to monitor volcanic clouds (Prata, 1989a; Pavolonis et al., 2015; Corradini et al., 2016; Biondi et al., 2017).

In recent years, satellites have emerged as a promising tool for monitoring volcanic clouds due to several advantages over ground-based methods. These advantages include global coverage, low costs, timely and continuous observations, and integration of multiple sensors (Poland et al., 2020). The most useful sensors for detecting volcanic ash and SO_2 operate in the infrared (IR) spectral regions (Dubuisson et al., 2014). Specifically, the IR region between 8 and 12 μm is optimal for this purpose as SO_2 gas is absorbed in the band around 8.6 μm, while ash particles exhibit dispersive behavior in the same range (Prata and Lynch, 2019). The Geostationary Meteorological Satellite (GMS), a geosynchronous weather satellite imagery, was the first employed by Sawada (1987) to record volcanic eruptions and track ash cloud dispersion in the western Pacific region in 1987. Sawada found that the detection rate was quite low (approximately 13.7%), primarily due to restrictions on spatial and spectral sampling, as well as interference from meteorological clouds (Sawada, 1996). Prata was the first to develop a theoretical approach in 1989 for distinguishing volcanic ash clouds from meteorological clouds using IR satellite imagery (Prata, 1989a; Prata, 1989b). This method relies on calculating the difference between two images captured at slightly different wavelengths within the IR window, specifically between 8 and 12 μm. Ash clouds primarily consist of silicate particles, which absorb, scatter, and re-emit infrared radiation differently from ice and water, the main components of meteorological clouds (Prata, 2009). Over time, advancements in technologies have led to the development of satellite instrument with improved resolutions, allowing for more precise monitoring of volcanic clouds. The geostationary satellite MSG (Meteosat Second Generation) is a series of climate and forecasting missions developed by ESA and EUMETSAT, launched on August 28, 2002. It carries the Spinning Enhanced Visible and InfraRed Imager (SEVIRI), a radiometer known for its good spectral resolution, from visible (VIS) to thermal infrared (TIR), and a high temporal resolution (with a 5- or 15-minute repeat cycle). These capabilities enable continuous monitoring of the entire evolution of a volcanic cloud.

Accurate automatic volcanic cloud detection by means of satellite data is a challenging task, and innovative strategies are essential to address this difficulty. One such strategy is the application of machine learning (ML), which offer new perspectives to learn complex patterns and trends from data, allowing a faster and more precise process to monitor volcanic phenomena in real-time (Gaddes et al., 2019; Biggs et al., 2022; Corradino et al., 2024a, 2024b; Di Bella et al., 2024; Malaguti et al., 2024; Rey-Devesa et al., 2024). Where the feature extraction phase is integrated in the ML technique rather than being performed manually (i.e. feature engineering), it is referred to as deep learning (DL), widely used for image segmentation and pattern recognition (Anantrasirichai et al., 2019; Amato et al., 2023; Corradino et al., 2023; Saunders-Shultz et al., 2024; Tan et al., 2024). Currently, new advanced ML approaches to process the huge amount of satellite data and retrieve important information about the formation and dispersion of a volcanic cloud are being developed (Del Negro et al., 2022; Petracca et al., 2022; Torrisi, 2022; Torrisi et al., 2024a). They eliminate the need to process images individually or set thresholds to determine whether a pixel belong to a volcanic cloud. Accurate detection of a volcanic cloud is essential for precisely determining its location in space, but it is also important to obtain with greater precision additional parameters that are useful for fully characterizing the evolution of the volcanic cloud. One of the key parameters to retrieve is the plume height. This parameter is essential to evaluate the Volcanic Explosive Index (Newhall and Self, 1982) or it can be used as input for volcanic cloud retrieval procedures (Pugnaghi et al., 2016). Different studies have identified an

empirical power law relationship that correlates plume height with the mass eruption rate (MER) (Aubry et al., 2023). This relationship allows for the rapid and simple estimation of the MER, which is essential for real-time volcanic hazard management and the reconstruction of past explosive eruptions. Moreover, the correlation between volcanic plume height and volcanic radiative power (VRP) is currently being investigated, with initial results suggesting a positive relationship between these two parameters that could facilitate forecasting the peak of thermal activity, specifically the maximum VRP value of an event (Torrisi et al., 2025).

The aim of this work is to propose an approach that combine two distinct ML models applied to SEVIRI data: a DL model to detect volcanic clouds (Torrisi et al., 2024a) and a supervised ML model to identify their main components (Torrisi et al., 2022). The choice to train two different ML models arises from the challenge of distinguishing volcanic clouds from meteorological clouds due to their similar spectral features. A DL model is ideal for detection, as it is able to segment the volcanic clouds in the images by analyzing both spectral and spatial patterns, including cloud shape. Meanwhile, a classical supervised ML model is well suited for identifying different volcanic cloud components based on their distinct spectral characteristics and it is used to classify the pixels of the images as ash-rich, SO_2-rich, or containing mixed components. After obtaining an accurate mask of the volcanic cloud, the volcanic plume height is determined from satellite observations, providing additional insights for a more detailed characterization. This integrated ML approach allows for the detection and tracking of entire volcanic events using SEVIRI high-temporal-resolution satellite images, enabling near real-time monitoring of volcanic clouds evolution and therefore the identification of regions that can be affected by their impacts. It was used to characterize the volcanic clouds generated by several lava fountains at Mt. Etna (Italy) between 2020 and 2022.

2. Materials and method

The proposed ML approach consists of two main steps, as is shown in Fig. 1:

- Input Feature Preparation: SEVIRI TIR bands are opportunely combined in order to create an Ash RGB (Red, Green, Blue) composite, designed to detect ash and SO_2 from volcanic eruptions. This composite serves as the input data for the subsequent analysis.
- Integrated ML approach: a DL model and a support vector machine (SVM) model are used in cascade to respectively detect a volcanic cloud and identify its main components.

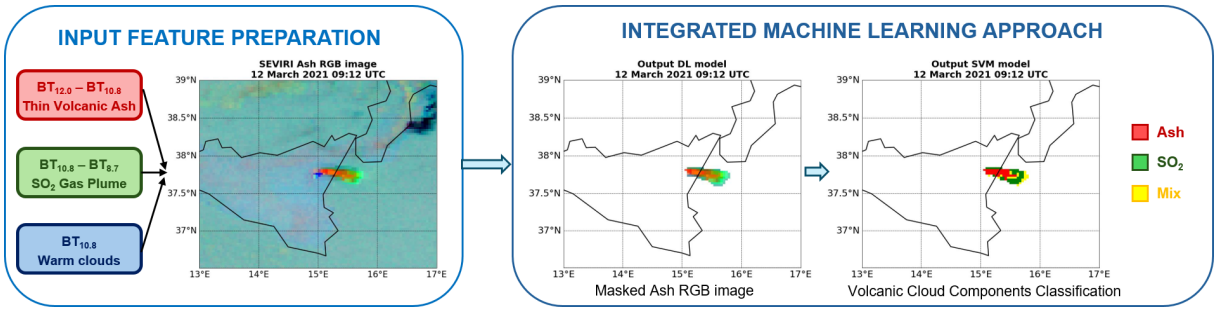

Figure 1. Scheme of the proposed machine learning (ML) approach: the first step is the realization of SEVIRI Ash RGB (Red, Green, Blue) images used as input data, while the second is the combination of a deep learning (DL) model and a support vector machine (SVM) model to respectively detect a volcanic cloud and classify its main components.

2.1 Input Feature Preparation

Data collected by the SEVIRI instrument were considered for this study. SEVIRI is a multispectral imaging radiometer used primarily to monitor weather and climate-related phenomena across Europe and Africa (Schmetz et al., 2007). SEVIRI has 12 spectral channels ranging from VIS to TIR, with a nadir spatial resolution of 3 km at sub-satellite point. It offers a temporal resolution of 15 min for Earth Full Disk and 5 min for Rapid Scan Mode (Aminou et al., 2002). The MSG Level 1.5 Image Data product was used, providing image data that has been corrected for all undesired radiometric and geometric effects, geolocated using a standardized projection, calibrated

and radiance-linearized. The Level 1.5 images, provided in a geostationary projection (GEOS Projection), were georeferenced to the EPSG:4326—WGS 84 reference system. SEVIRI images of Sicily were collected, covering the region between 36.5°-39.5° north latitude and 12°-18° east longitude. This area includes Etna and its surroundings, which are often affected by volcanic clouds produced during explosive eruptions.

As input features of the proposed ML approach, we used SEVIRI Ash RGB images. These images are RGB composites created by combining the brightness temperature (BT) of the bands centered at 8.7, 10.8, and 12.0 μm (respectively, $BT_{8.7}$, $BT_{10.8}$ and $BT_{12.0}$), producing a false-color image that highlights the presence of volcanic ash and SO_2 from volcanic activity. The methodology to realize these images is described in detail in Torrisi et al. (2024a). To interpret the different colors in a SEVIRI Ash RGB image:

- Volcanic ash pixels typically appear red when minimal SO_2 is present and yellow when SO_2 concentrations are high;
- In the absence of ash, SO_2 appears green.

2.2 Integrated machine learning approach

Designing an accurate ML model with a high generalization capability is a challenging task. To address this, combining multiple ML models can improve performance by leveraging their strengths. Here, we propose an integrated ML approach which combines in cascade two distinct models: a hybrid DL model, called VGG16-UNet, and a SVM. The VGG16-UNet model uses spatial and spectral intensity data derived from SEVIRI Ash RGB images to segment volcanic clouds. Its output is a masked Ash RGB image, where the mask represents the selection of all the pixels within the volcanic clouds. This masked Ash RGB image then serves as input for the SVM, which generates a classification image of the volcanic clouds components, i.e. an image where the pixels within the volcanic clouds are categorized as ash-rich, SO_2-rich, or containing mixed components. One of the key advantages of using ML techniques is the ability to process large volumes of Ash RGB images efficiently, enabling the automatic detection of volcanic clouds and the identification of their main components without the need for manual, image-by-image analysis. Given that SEVIRI captures images every 5/15 minutes, these techniques enable rapid and automated tracking of volcanic cloud formation and propagation. The choice to train two different ML models is essential because detecting volcanic clouds is challenging; their spectral features are really similar to those of meteorological clouds, making difficult the discrimination between the two. Using a DL model for detection is effective because it can identify complex patterns within the data while considering both spectral and spatial features, including the shape of a volcanic cloud. On the other hand, a classical supervised ML model is well-suited for distinguishing the various components based on their distinct spectral characteristics.

2.2.1 VGG16-UNet: Deep learning model to detect volcanic clouds

The VGG16-UNet is an hybrid DL architecture developed to detect volcanic clouds using satellite images (Torrisi et al., 2024a). This model was realized combining the UNet architecture (Long et al., 2015) with the VGG16 (Simonyan and Zisserman, 2015). UNet is popular for fast and precise segmentation of images, and 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. The encoder block of the U-Net is replaced with a VGG16 due to its similarity to U-Net's encoder path, the smaller number of parameters, and the easy accessibility to parameter weights, which are used in the new model (Kanaeva and Ivanova, 2021). Instead of training from scratch, we adopt transfer learning to enhance the accuracy of the output, saving training time and requiring less training data. The VGG16-UNet model has been trained with a learning rate of 0.001. The Adaptive Moment Estimator (Adam) was used as the optimizer to minimize the focal loss, which aims to address the class imbalance problem commonly encountered in satellite image segmentation tasks (Corradino et al., 2023), as volcanic clouds often occupy a significantly smaller portion of the image compared to the background. The architecture of the VGG16-UNet was implemented and described in details in Torrisi et al. (2024a). The VGG16-UNet was trained on SEVIRI Ash RGB images, with each pixel labeled as either "volcanic cloud" or "background", to detect volcanic clouds at Etna, achieving an accuracy of 0.90. It was found that the VGG16-UNet model outperformed traditional threshold-based methods like the Brightness Temperature Difference (BTD). The BTD technique calculates the difference between the $BT_{10.8}$ and the $BT_{12.0}$, classifying pixels below a set threshold as "volcanic cloud" and those above as "background". This is due to the ability of VGG16-UNet

to integrate spatial information alongside spectral features, enhancing its capability to accurately identify volcanic clouds. A preliminary attempt to train this model on different satellite data was conducted in Torrisi et al. (2024b), where the VGG16-UNet was trained on SEVIRI images of Etna and Advanced Baseline Imager (ABI) images of Shishaldin, a volcano in Alaska, using the Ash RGB composite as input in both instances. ABI is the primary instrument onboard the geostationary GOES-R (Geostationary Operational Environmental Satellite) series, designed to observe the Western Hemisphere. This approach of training the same model on two different satellite datasets is still under investigation; therefore, in this paper we use the model trained exclusively on SEVIRI images.

Figure 2 illustrates the application steps of the VGG16-UNet DL model. The SEVIRI Ash RGB image of the 12 March 2021 at 09:12 UTC was used as input for the model. The outcome is a probability map that assigns each pixel a probability of belonging to the "volcanic cloud" class. Using a validation dataset, we identified an optimal threshold t>0.8. Applying this threshold results in a binary image where pixels within the cloud are set to 1, while background pixels are set to 0. Consequently, the final output of the VGG16-UNet is a masked Ash RGB image of the volcanic clouds.

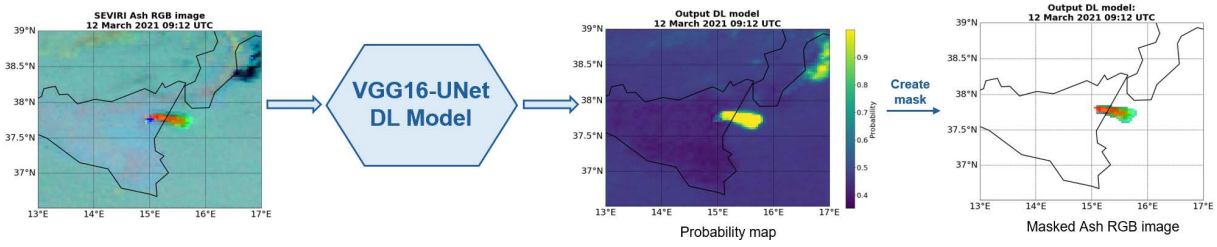

Figure 2. Scheme of the VGG16-UNet DL model used to segment a volcanic cloud. The input consists of a SEVIRI Ash RGB image, while the output is a probability map that assigns each pixel a probability of belonging to the "volcanic cloud" class. After applying a threshold to the probability map ("Create mask" step), the final output of the VGG16-UNet is a masked Ash RGB image of the volcanic clouds.

2.2.2 SVM: Supervised machine learning model to identify the main volcanic cloud components

A SVM model was trained for a multi-class classification problem to identify the main components of a volcanic cloud. This model classifies each pixel in a SEVIRI Ash RGB image into one of four classes: ash-rich (class ash), SO_2-rich (class SO_2), mixed components (class mix), or background (class background) (Torrisi et al., 2022). The SVM technique has been successfully applied to volcano monitoring tasks, demonstrating effective results (Amato et al., 2021; Corradino et al., 2021b; Del Negro et al., 2022), thanks to its high performance and lower computation costs. The SVM classifier was designed using a radial basis function (RBF) as kernel, which measures the similarity between pairs of data points, based on their distance in the feature space, and performs effectively with non-linear and high-dimensional data. The parameter γ represents the spread of the kernel and was carefully tuned for the problem, while c is the cost penalty for misclassifying a data point. γ and c were set to 0.5 and 10, respectively. The SVM was implemented and described in detail in Torrisi et al. (2022).

Here, the SVM is applied to the masked Ash RGB image produced by the VGG16-UNet DL model. Since the detection task has been already accomplished with a good accuracy, the objective now is to classify the pixels within the volcanic cloud according to their respective components. The SVM can also classify pixels into the background class. Identifying background pixels provides an additional verification step to assess the quality of the detection and enhance its accuracy. Pixels within the volcanic cloud that are classified as background undergo further examination. The probabilities assigned to these pixels by the VGG16-UNet DL model are evaluated. If a pixel's probability exceeds 0.95, it is considered part of the volcanic cloud and reclassified as "mix", while if the probability is 0.95 or lower, that pixel is excluded from the detection mask. Figure 3 illustrates an example of the SVM model applied to the masked Ash RGB image of the 12 March 2021 at 09:12 UTC, produced by the VGG16-UNet DL model. As first result, the number of pixels within the volcanic cloud classified as ash, SO_2 and mix are 41, 46 and 9, respectively. Additionally, 21 pixels are classified as background. However, after verifying the probabilities assigned to these pixels by the VGG16-UNet DL model, 9 pixels are excluded from the detection mask, while the remaining 12

pixels are reclassified as mix. Therefore, the final output of the SVM is a classification image of the volcanic clouds components, with a total of 41, 46 and 21 pixels belonging to the ash, SO_2 and mix classes, respectively.

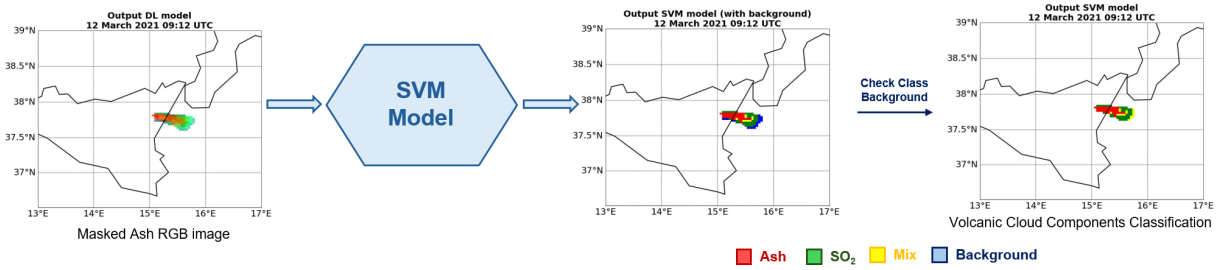

Figure 3. Scheme of the SVM model used to identify the main components of a volcanic cloud. The input is the masked Ash RGB image produced by the VGG16-UNet DL model, while the output is an image where each pixel within the volcanic cloud is classified as "ash", "SO_2", "mix" or "background". Pixels within the volcanic cloud that are classified as background undergo further examination ("Check Class Background" step). The probabilities assigned to these pixels by the VGG16-UNet DL model are evaluated, and if they are below a threshold, they are excluded; otherwise, they are reclassified as the "mix" class.

3. Case study: 2020-2022 Etna eruptions

Etna is one of the most active volcanoes in the world, situated on the east coast of Sicily. It is renowned for continuous activity at its summit craters, which includes degassing, explosive events, and rapid lava flows (Del Negro et al., 2019). Since the start of the new millennium, Etna's eruptive activity has been primarily characterized by sequences of lava fountain events, mainly originating from the South-East Crater (SEC) in the summit area. Paroxysmal episodes occur typically close together in time and are more impulsive, featuring rapid waxing and waning phases (Andronico et al., 2021).

On 13 December 2020 Etna volcano entered a new phase characterized by several lava fountains episodes, occurred from the SEC (Calvari et al., 2021). This event was followed by three others (21 and December 2020 and 18 January 2021). Then, from 16 February to 31 March 2021, Etna experienced a series of 17 paroxysms of short-duration (lasting a few hours) and high-intensity, occurring at regular intervals of approximately 30 hours (Torrisi et al., 2023). Afterward, Etna entered a period of inactivity before initiating a new phase of lava fountains from May to July, which occurred in relatively quick succession but were generally less energetic than the earlier events. The 2021 sequence concluded with two episodes in August (on the 9th and 29th), one in September (on the 21st), and the final eruption on 23 October 2021. In 2022, only two lava fountains were observed, occurring on 10 and 21 February. Most of these lava fountains were accompanied by ash plumes rising several kilometers high, reaching up to 9-10 km above sea level, causing significant concern among residents of towns surrounding the volcano. The ash fallout had a major impact on road safety, roof stability, air traffic (with Catania's international airport being a key hub), agriculture, water quality, and the health of the local population (Andronico and Del Carlo, 2016; Guerrieri et al., 2023). The complete list of lava fountains from December 2020 to February 2022 was derived from the bulletin emanated from the INGV-OE (Istituto Nazionale di Geofisica e Vulcanologia) and is available in Calvari and Nunnari (2022).

The integrated ML approach was applied to the series of lava fountains that occurred from December 2020 to February 2022 and, starting from these results, the maximum volcanic plume height (VPH) of each event was retrieved. The VPH is determined by using the darkest pixel (DP) technique (Corradino et al., 2021a; Guerrieri et al., 2023). This is a simplified approach that calculates the brightness temperature at 10.8 μm ($BT_{10.8}$) for the opaquest pixels of the volcanic plume, i.e. the minimum $BT_{10.8}$ value within the volcanic cloud (Prata and Grant, 2001). This value can be compared with a temperature profile, selected to be as close as possible in time and space, to determine the altitude at which the temperature best matches the plume-top temperature. This procedure is simply to apply but is reliable only when the cloud behaves as a black body and the atmospheric profile is representative. Therefore, the main weaknesses arise when the plume is not completely opaque and, to address this, the minimum $BT_{10.8}$ value was reduced by 2K (Prata and Grant, 2001).

Moreover, another drawbacks occurs for VPHs near the tropopause, where temperature variations with altitude are minimal (Corradini et al., 2018). For Etna, the atmospheric temperature profiles used are those derived from the LICT Trapani Station, provided by the Department of Atmospheric Science of the University of Wyoming (https://weather.uwyo.edu/upperair/sounding.html). The advantage of applying the DP technique after the integrated ML approach lies in the possibility to automatically search for the minimum $BT_{10.8}$ within the mask of the volcanic cloud, resulting in a more precise retrieval of this value. This method helps avoid considering erroneous values outside the mask, which are likely associated with meteorological clouds.

Figure 4 shows the VPH for all 2020-2022 Etna lava fountains, as determined using the DP technique. The VPH ranges from 3 km (22 May 2021 21:42 UTC) to 13.5 km (23 October 2021 23:12 UTC). The period February-March 2021 was marked by high volcanic plumes, with a maximum VPH of 12.8 km reached on 23 February 2021 at 00:42 UTC. The sequence of lava fountains between May and July 2021 was initially characterized by lower volcanic plumes, ranging from 3 km (22 May 2021 21:42 UTC) to 7.2 km (12 June 2021 21:27 UTC). Starting on 19 June 2021, the VPH increased, with several values exceeding 10 km. The final phase includes six paroxysms between August 2021 and February 2022, characterized by high plumes (between 9 and 11 km), with a maximum of 13.5 km reached on 21 October 2021 at 23:12 UTC.

The following subsections present the application of the integrated ML approach, showing the model's capability to detect volcanic clouds and identify their main components. Additionally, the time series of the VPH will be displayed, illustrating how the plume evolves vertically over time. It was considered one case study for each of the main sequences of 2020-2022 Etna lava fountains, and specifically the 16 February 2021, 23 June 2021 and 29 August 2021.

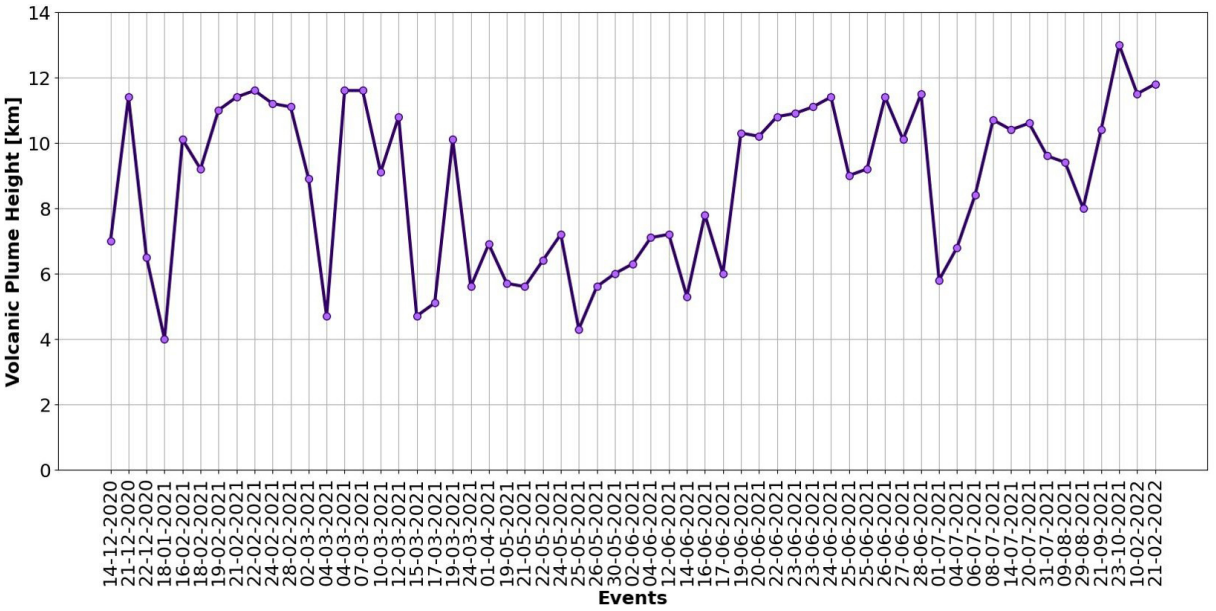

Figure 4. Time series of the volcanic plume height (VPH) calculated using the darkest pixel (DP) technique for all the 2020-2022 Etna lava fountains.

3.1 16 February 2021

During the morning of 15 February, the explosive activity of the SEC intensified, evolving into lava fountains around 16:00 UTC on February 16. This activity generated a lava flow of $1.45 \cdot 10^6$ m^2 (Amato et al., 2021) and produced a volcanic cloud that persisted in the atmosphere for approximately four hours.

Figure 5 illustrates the results of applying the integrated ML approach to SEVIRI images of the 16 February 2021. Specifically, Figs. 5a, 5b, 5c, and 5d show the input data of the proposed approach, i.e. the SEVIRI Ash RGB images at the following times: 16:42 UTC, 16:57 UTC, 17:12 UTC, and 17:27 UTC. The volcanic cloud is observed dispersing toward the southeast, impacting the area around Catania airport. Figures 5e, 5f, 5g,5h display the results of applying

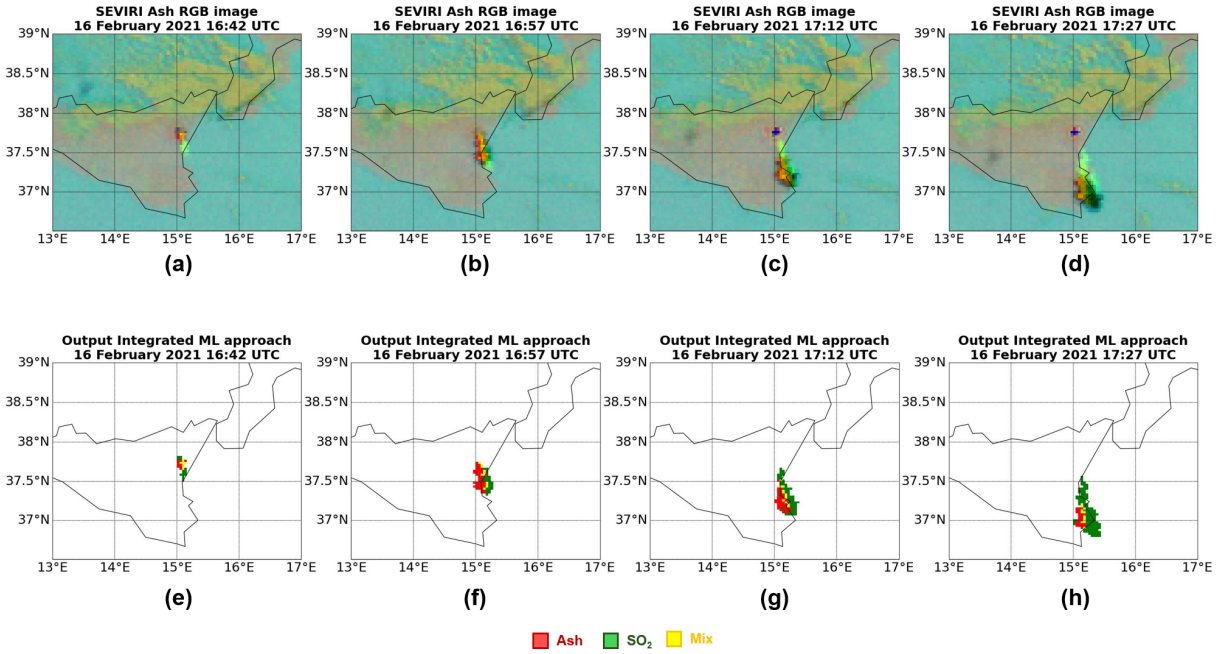

Figure 5. SEVIRI Ash RGB images used as input of the model and outcomes of the integrated ML approach for Etna on 16 February 2021 at 16:42 UTC (a, e), 16:57 UTC (b, f), 17:12 UTC (c, g) and 17:27 UTC (d, h).

the VGG16-UNet for the detection of volcanic clouds and subsequently the SVM for the identification of their main components. Initially, ash is the predominant component, in fact in Fig. 5f (16 February 2021 16:57 UTC) the model classifies 46 pixels as ash, 11 as mix and 27 as SO_2. Conversely, in the final frame in Fig. 5h, the percentage of SO_2 increases relative to ash, resulting in 27 pixels classified as ash, 12 as mix and 122 as SO_2.

A sequence of 7 SEVIRI images was processed, from 16 February 2021 16:27 UTC to 17:57 UTC. These images cover the entire period of formation and dispersion of the volcanic cloud into the atmosphere. The VPH was calculated using the DP technique and the results are shown in Fig. 6. At 16:27 UTC the volcanic cloud is detected for the first time with a height of 4 km. The VPH then begins to increase, reaching a maximum of 10.1 km at 17:12.

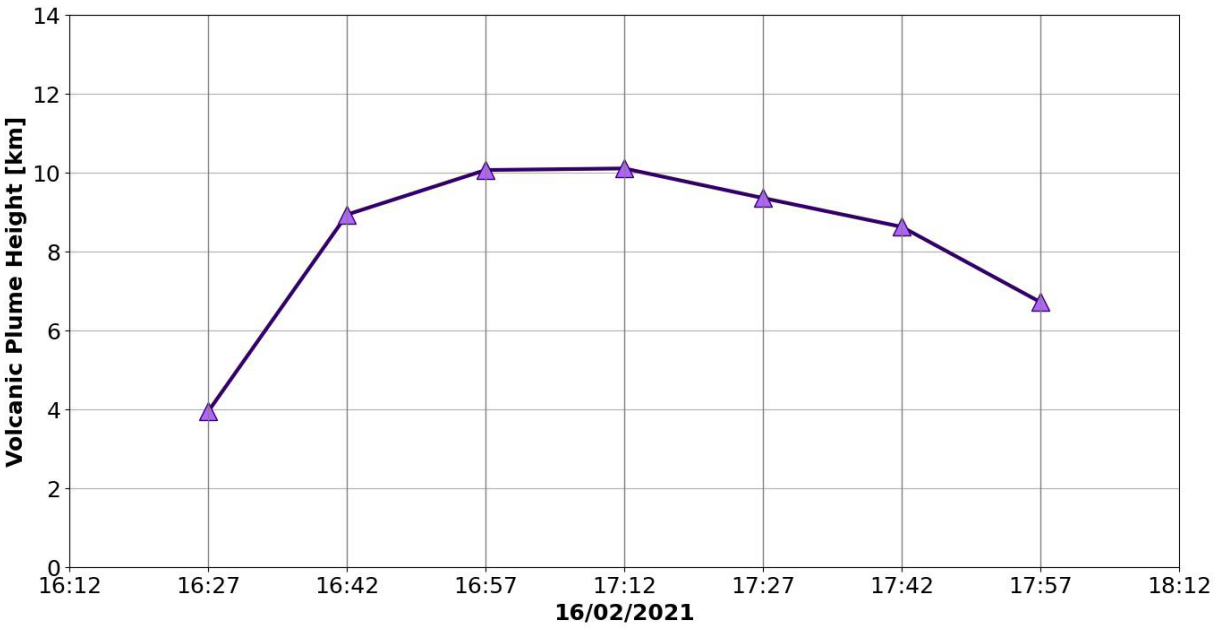

Figure 6. Time series of the VPH calculated using the DP technique for the event of 16 February 2021, from 16:27 UTC to 17:57 UTC.

3.2 23 June 2021

On 23 June 2021, Etna produced two lava fountains: one at the beginning of the day and a second at the end. The first lava fountain started at about 02:44 on 23 June 2021 and it erupted about $0.38 \cdot 10^6$ m^3 of volume (Calvari and Nunnari, 2022). A strong plume was produced, which was dispersed in the east direction. In Fig. 7 the results of applying the integrated ML approach to SEVIRI images from the early morning of 23 June 2021 are presented.

Figures 7a, 7b, 7c, and 7d show the SEVIRI Ash RGB images at the following times: 03:12 UTC, 03:27 UTC, 03:42 UTC, and 03:57 UTC. The volcanic cloud is observed dispersing toward the east, impacting the city of Catania and the nearby towns. In the lower part of the images, the presence of some meteorological clouds can be seen. This is one of the cases where the checking of the background class after applying the SVM proves useful for improving the accuracy of the detection performed by the VGG16-UNet. In Figs. 7e, 7f, 7g, 7h, we can observe the results of the SVM applied to identify the main components of the volcanic cloud, including the background class. The number of background pixels in these figures are 16, 8, 15 and 27, respectively. After checking the background class, i.e. verifying the probabilities assigned to these pixels by the VGG16-UNet DL model, some pixels were excluded from the detection masks: 11, 2, 10 and 9, respectively. The remaining pixels were reclassified as "mix" class, as shown in Figs. 7i, 7l, 7m, 7n, which represent the final output of the integrated ML approach. The main component of this volcanic cloud is ash, as evidenced by the fact that the largest number of pixels classified as ash is in Fig. 7m (66 pixels).

A sequence of 14 SEVIRI images was processed, spanning from 23 February 2021 at 02:42 UTC to 05:57 UTC. These images cover the entire period of formation and dispersion of the volcanic cloud into the atmosphere. The VPH was calculated using the DP technique and the results are shown in Fig. 8. At 02:42 UTC the volcanic cloud is detected for the first time, with a height of 8.1 km. The VPH then begins to increase, reaching a maximum of 11 km at 03:12.

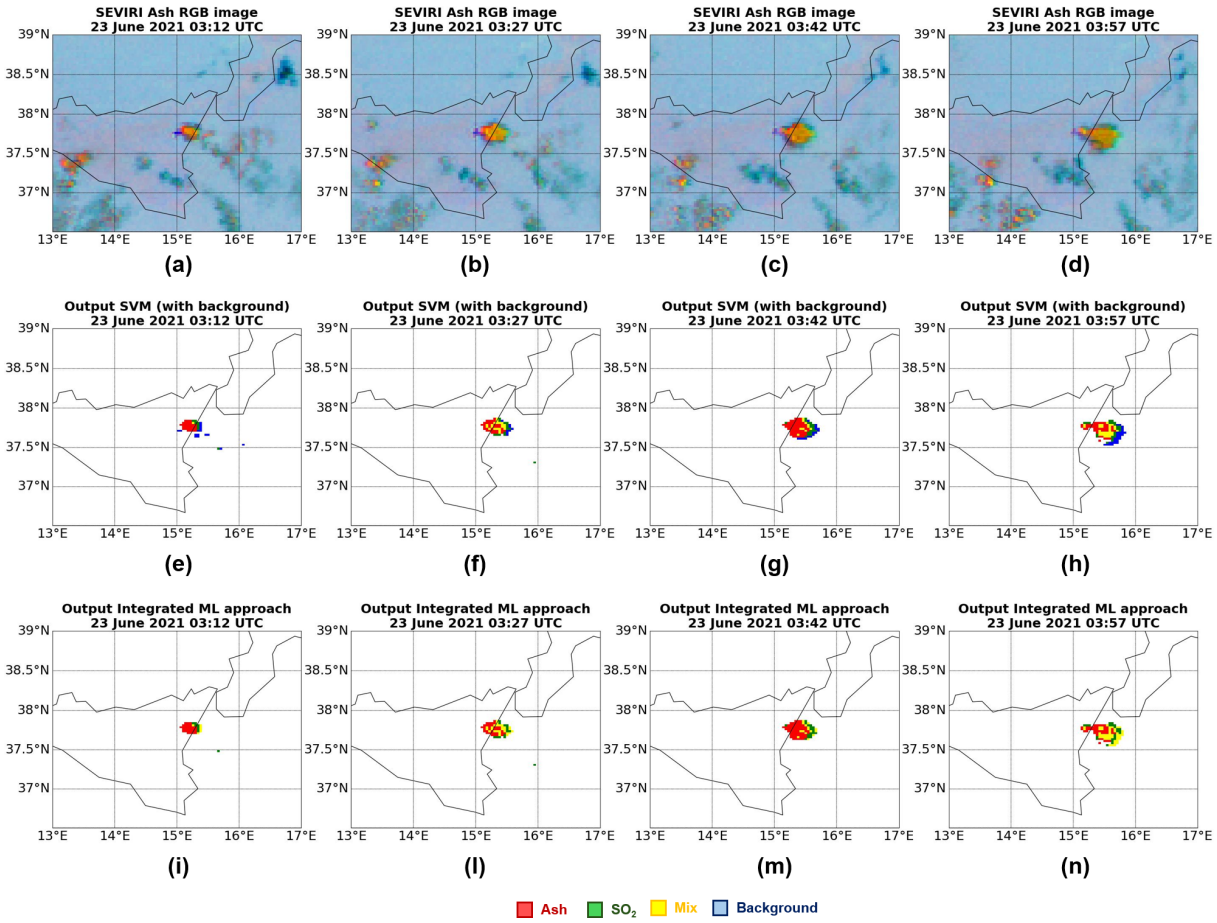

Figure 7. SEVIRI Ash RGB images used as input of the model, output of the SVM (including background pixels) and outcomes of the integrated ML approach for Etna on 23 June 2021 at 03:12 UTC (a, e, i), 03:27 UTC (b, f, l), 03:42 UTC (c, g, m) and 03:57 UTC (d, h, n).

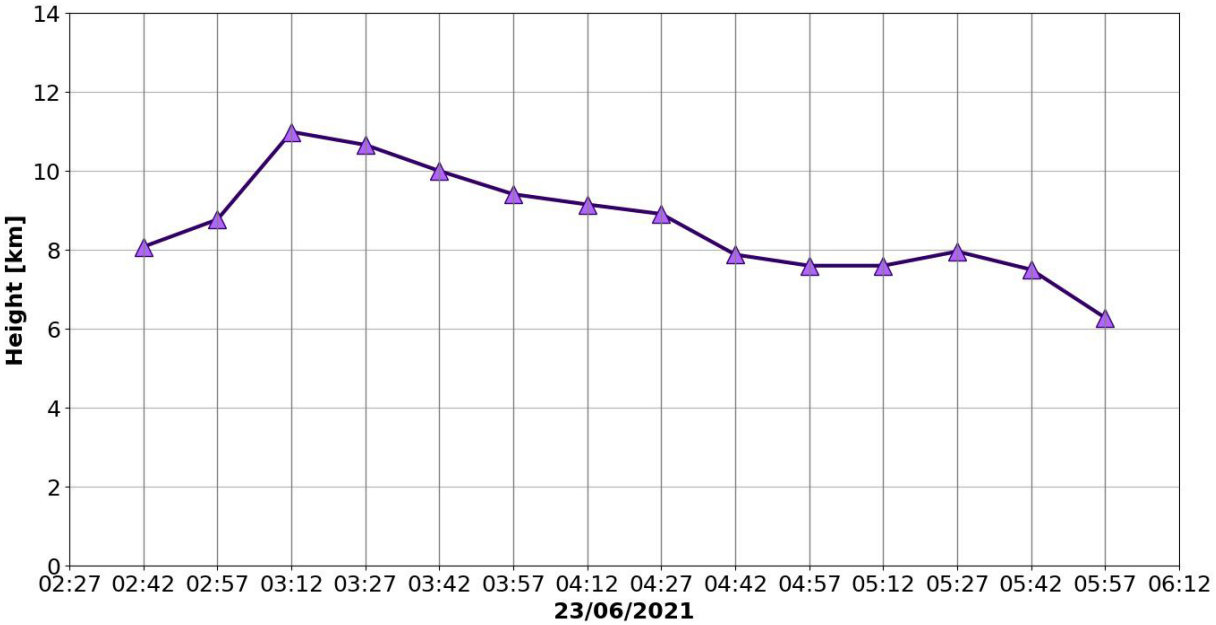

Figure 8. Time series of the VPH calculated using the DP technique method for the event of 23 June 2021, from 02:42 UTC to 05:57 UTC.

3.3 29 August 2021

Strombolian activity started in the early afternoon of 29 August 2021, and within an hour, a lava fountain was observed. A volcanic cloud was detected in SEVIRI images from 16:57 UTC and impacted the small towns located in the east side of the volcano. Figures 9a and 9b shows the input data of the integrated ML approach, i.e. the SEVIRI Ash RGB images of 29 August 2021 respectively at 16:57 UTC and 17:12 UTC, while Figs. 9c and 9d display the corresponding output. In Fig. 9c (29 August 2021 at 16:57 UTC) the pixels classified as ash, SO_2 and mix are respectively 12, 8, and 20, while in Fig. 9d (29 August 2021 at 17:12 UTC) they are respectively 26, 6, and 25. The maximum VPH is reached in the first SEVIRI image (29 August 2021 at 16:57 UTC), at approximately 8.2 km, as depicted in Fig. 9e. Moreover, Fig. 9e shows the time series of the VPH for the entire event, from 29 August 2021 at 16:57 UTC to 19:27 UTC. After reaching its peak, the VPH begins to decrease gradually.

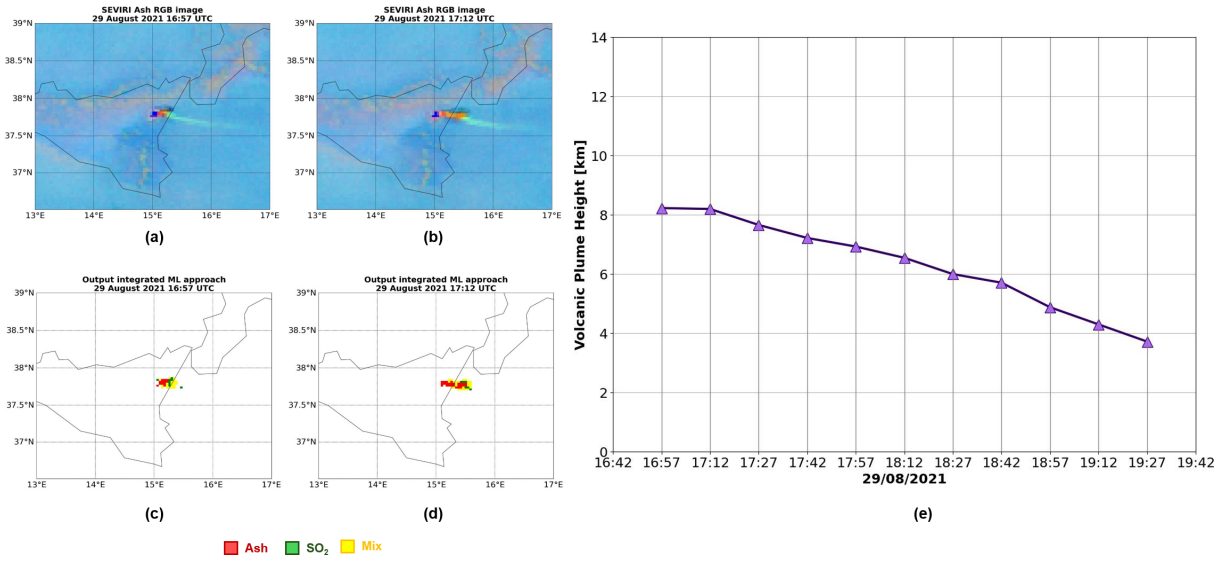

Figure 9. SEVIRI Ash RGB images used as input of the model and outcomes of the integrated ML approach for Etna on 29 August 2021 at 16:57 UTC (a, c) and 17:12 UTC (b, d). Time series of the VPH calculated using the DP technique for the event of 29 August 2021 (e).

4. Conclusions

The development of an integrated ML approach to detect and characterize volcanic clouds plays a crucial role in monitoring the emissions of silicate particles and gases during violent explosive eruptions, helping to assess the impact that these components can have on the environment, the human health and the aviation. Two ML models are combined in cascade: the VGG16-UNet DL model for volcanic cloud detection and the SVM model for identifying its primary components. By integrating the detection capabilities of the VGG16-UNet DL model with the characterization provided by the SVM, it becomes possible to detect and track entire volcanic events using SEVIRI's high temporal resolution satellite images. This approach facilitates near real-time monitoring of volcanic cloud evolution, aiding in the identification of high-risk areas, as volcanic ash and SO_2 emissions can pose significant health hazards. This method efficiently processes large satellite datasets, ensuring rapid availability of results. As a result, valuable insights into the dispersion of volcanic clouds can be obtained, aiding in assessing their potential effects on the environment and climate.

The integrated ML approach has been successfully applied to the lava fountains at Etna volcano from 2020 to 2022, enabling the processing of satellite data over a two-year period and automatically generating useful products to characterize the evolution of a volcanic cloud, including its components and height. Further investigations are conducting to apply the proposed ML approach to other volcanoes, particularly those with different eruption styles. To ensure the model's adaptability to various case studies, the training dataset have to include a collection of eruptions from different volcanoes, encompassing various eruption styles, locations, and atmospheric conditions. This will enable the model to learn from various scenarios, allowing the proposed ML approach to be applied globally and support near-real-time tracking of volcanic cloud evolution.

Acknowledgements. This work was developed within the framework of the Laboratory of Technologies for Volcanology (TechnoLab) at the Istituto Nazionale di Geofisica e Vulcanologia (INGV) in Catania (Italy). This research was funded by the ATHOS Research Program (INGV OB.FU. 0867.010), by the 2019 Strategic Project FIRST – ForecastIng eRuptive activity at Stromboli volcano: timing, eruptive style, size, intensity, and duration – of the INGV Volcanoes Department (Delibera n. 144/2020), and by Project INGV Pianeta Dinamico VT_ORME 2023-2025 (INGV OB.FU. 1020.010). I am grateful to EUMETSAT for the SEVIRI data (https://data.eumetsat.int, 19 October 2024).

References

Amato, E., C. Corradino, F. Torrisi and C. Del Negro (2021). Mapping lava flows at Etna Volcano using Google Earth Engine, open-access satellite data, and machine learning, 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME, 1-6, doi:10.1109/ICECCME52200.2021.9591110.

Amato, E., C. Corradino, F. Torrisi and C. Del Negro (2023). A Deep Convolutional Neural Network for Detecting Volcanic Thermal Anomalies from Satellite Images, Remote Sens., 15, 3718, doi:10.3390/rs15153718.

Aminou, D. M. A., A. Ottenbacher, B. Jacquet, F. Pasternak et al. (2002). Meteosat second generation: inflight calibration of the imaging radiometer SEVIRI, Proceedings Earth Observing Systems VI, SPIE, 4483, 248-257, doi:10.1117/12.453460.

Anantrasirichai, N., J. Biggs, F. Albino and D. Bull (2019). A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets, Remote Sens. Environ., 230, 111179, doi:10.1016/j.rse.2019.04.032.

Andronico, D. and P. Del Carlo (2016). PM10 measurements in urban settlements after lava fountain episodes at Mt. Etna, Italy: pilot test to assess volcanic ash hazard to human health, Nat. Hazards Earth Syst. Sci., 16, 29-40, doi:10.5194/nhess-16-29-2016.

Andronico, D., A. Cannata, G. Di Grazia and F. Ferrari (2021). The 1986-2021 paroxysmal episodes at the summit craters of Mt. Etna: Insights into volcano dynamics and hazard, Earth Sci. Rev., 220, 103686, doi:10.1016/j.earscirev.2021.103686.

Aubry, T. J., S. L. Engwell, C. Bonadonna, L. G. Mastin et al. (2023). New Insights Into the Relationship Between Mass Eruption Rate and Volcanic Column Height Based On the IVESPA Data Set, Geophys. Res. Lett., 50, e2022GL102633, doi:10.1029/2022GL102633.

Biggs, J., N. Anantrasirichai, F. Albino, M. Lazecky et al. (2022). Large-scale demonstration of machine learning for the detection of volcanic deformation in Sentinel-1 satellite imagery, Bull. Volcanol., 84, 100, doi:10.1007/s00445-022-01608-x.

Biondi, R., A. K. Steiner, G. Kirchengast, H. Brenot et al. (2017). Supporting the detection and monitoring of volcanic clouds: A promising new application of Global Navigation Satellite System radio occultation, Adv. Space Res., 60, 2707-2722. doi:10.1016/j.asr.2017.06.039.

Branca, S. and P. D. Carlo (2005). Types of eruptions of Etna volcano AD 1670-2003: implications for short-term eruptive behaviour, Bull. Volcanol., 67, 732-742, doi:10.1007/s00445-005-0412-z.

Calvari, S., A. Bonaccorso and G. Ganci (2021). Anatomy of a Paroxysmal Lava Fountain at Etna Volcano: The Case of the 12 March 2021, Episode, Remote Sens., 13, 3052, doi:10.3390/rs13153052.

Calvari, S. and G. Nunnari (2022). Comparison between Automated and Manual Detection of Lava Fountains from Fixed Monitoring Thermal Cameras at Etna Volcano, Italy, Remote Sens., 14, 2392, doi:10.3390/rs14102392.

Calvari, S. and G. Nunnari (2024). Reawakening of Voragine, the Oldest of Etna's Summit Craters: Insights from a Recurrent Episodic Eruptive Behavior, Remote Sens., 16, 22, 4278, doi:10.20944/preprints202409.2414.v1.

Carn, S. A., L. Clarisse and A. J. Prata (2016). Multi-decadal satellite measurements of global volcanic degassing, J. Volcanol. Geotherm. Res., 311, 99-134, doi:10.1016/j.jvolgeores.2016.01.002.

Corradini, S., M. Montopoli, L. Guerrieri, M. Ricci et al. (2016). A multi-sensor approach for volcanic ash cloud retrieval and eruption characterization: The 23 November 2013 Etna lava fountain, Remote Sens., 8, 1, 58, doi:10.3390/rs8010058.

Corradini, S., L. Guerrieri, V. Lombardo, L. Merucci et al. (2018). Proximal monitoring of the 2011-2015 Etna lava fountains using MSG-SEVIRI data, Geosci., 8, 4, 140, doi:10.3390/geosciences8040140.

Corradino, C., E. Amato, F. Torrisi, S. Calvari et al. (2021a). Classifying major explosions and paroxysms at Stromboli Volcano (Italy) from space, Remote Sens., 13, 20, 4080, doi:10.3390/rs13204080.

Corradino, C., G. Bilotta, A. Cappello, L. Fortuna et al. (2021b). Combining radar and optical satellite imagery with machine learning to map lava flows at Mount Etna and Fogo Island, Energies, 14, 1, 197, doi:10.3390/en14010197.

Corradino, C., M. S. Ramsey, S. Pailot-Bonnetat, A. J. L. Harris et al. (2023). Detection of subtle thermal anomalies: Deep learning applied to the ASTER global volcano dataset, IEEE Trans. Geosci. Remote Sens., 61, 1-15, doi:10.1109/TGRS.2023.3241085.

Corradino, C., P. Jouve, A. La Spina and C. Del Negro (2024a). Monitoring Earth's atmosphere with Sentinel-5 TROPOMI and artificial intelligence: Quantifying volcanic SO2 emissions, Remote Sens. Environ., 315, 114463, doi:10.1016/j.rse.2024.114463.

Corradino, C., Malaguti, A. B., Ramsey, M. S. and C. Del Negro (2024b). Quantitative assessment of volcanic thermal activity from space using an isolation forest machine learning algorithm, Remote Sens., 16, 11, 2001, doi:10.3390/rs16112001.

Corsaro, R. A., L. Miraglia, I. Arienzo and V. Di Renzo (2024). The 2020-2022 paroxysmal episodes of the South-East Crater (Mt. Etna): Insights into high-frequency eruptive activity from petrological monitoring, Bull. Volcanol., 86, 85, doi:10.1007/s00445-024-01770-4.

Del Negro, C., A. Cappello, G. Bilotta, G. Ganci et al. (2019). Living at the edge of an active volcano: Risk from lava flows on Mt. Etna, Geol. Soc. Am. Bull., 132, 7-8, 1615-1625, doi:10.1130/B35290.1.

Del Negro, C., E. Amato, F. Torrisi, C. Corradino et al. (2022). Support vector machine for volcano hazard monitoring from space at Mount Etna, in IEEE MELECON, IEEE, Palermo, 627-631, doi:10.1109/MELECON53508.2022.9842942.

Di Bella, G. S., C. Corradino, S. Cariello, F. Torrisi et al. (2024). Advancing volcanic activity monitoring: A near-real-time approach with remote sensing data fusion for radiative power estimation, Remote Sens., 16, 16, 2879, doi:10.3390/rs16162879.

Dubuisson, P., H. Herbin, F. Minvielle, M. Compiègne et al. (2014). Remote sensing of volcanic ash plumes from thermal infrared: A case study analysis from SEVIRI, MODIS, and IASI instruments, Atmos. Meas. Tech., 7, 2, 359-371, doi:10.5194/amt-7-359-2014.

Durant, A. J., C. Bonadonna and C. J. Horwell (2010). Atmospheric and environmental impacts of volcanic particulates, Elements, 6, 4, 235-240, doi:10.2113/gselements.6.4.235.

Filonchyk, M., M. P. Peterson, A. Gusev, F. Hu et al. (2022). Measuring air pollution from the 2021 Canary Islands volcanic eruption, Sci. Total Environ., 849, 157827, doi:10.1016/j.scitotenv.2022.157827.

Freret-Lorgeril, V., F. Donnadieu, S. Scollo, A. Provost et al. (2018). Mass eruption rates of tephra plumes during the 2011-2015 lava fountain paroxysms at Mt. Etna from Doppler radar retrievals, Front. Earth Sci., 6, 73, doi:10.3389/feart.2018.00073.

Gaddes, M. E., A. Hooper and M. Bagnardi (2019). Using machine learning to automatically detect volcanic unrest in a time series of interferograms, J. Geophys. Res. Solid Earth, 124, 12, 12304-12322, doi:10.1029/2019JB017519.

Ge, C., J. Wang, S. Carn, K. Yang et al. (2016). Satellite-based global volcanic SO2 emissions and sulfate direct radiative forcing during 2005-2012, J. Geophys. Res. Atmos., 121, 6, 3446-3464, doi:10.1002/2015JD023134.

Gudmundsson, G. (2011). Respiratory health effects of volcanic ash with special reference to Iceland: A review, Clin. Respir. J., 5, 1, 2-9, doi:10.1111/j.1752-699X.2010.00231.x.

Guerrieri, L., S. Corradini, N. Theys, D. Stelitano et al. (2023). Volcanic clouds characterization of the 2020-2022 sequence of Mt. Etna lava fountains using MSG-SEVIRI and products' cross-comparison, Remote Sens., 15, 8, 2055, doi:10.3390/rs15082055.

Guffanti, M., G. C. Mayberry, T. J. Casadevall and R. Wunderman (2009). Volcanic hazards to airports, Nat. Hazards, 51, 2, 287-302, doi:10.1007/s11069-008-9254-2.

Kanaeva, I. and J. Ivanova (2021). Road pavement crack detection using deep learning with synthetic data, IOP Conference Series, Materials Science and Engineering, 1019, 012036, doi:10.1088/1757-899X/1019/1/012036.

Long, J., E. Shelhamer and T. Darrell (2015). Fully convolutional networks for semantic segmentation, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431-3440, doi:10.1109/CVPR.2015.7298965.

Malaguti, A. B., C. Corradino, A. La Spina, S. Branca et al. (2024). Machine learning insights into the last 400 years of Etna lateral eruptions from historical volcanological data, Geosci., 14, 11, 295, doi:10.3390/geosciences14110295.

Marshall, L. R., A. Schmidt, A. P. Schurer, N. L. Abraham et al. (2025). Last-millennium volcanic forcing and climate response using SO2 emissions, Clim. Past, 21, 161-184, doi:10.5194/cp-21-161-2025.

Mather, T. A., D. M. Pyle and C. Oppenheimer (2003). Tropospheric volcanic aerosol, in Geophysical Monograph Series A. Robock and C. Oppenheimer (Eds.), American Geophysical Union, Washington, 189-212, doi:10.1029/139GM12.

Newhall, C. G. and S. Self (1982). The volcanic explosivity index (VEI): An estimate of explosive magnitude for historical volcanism, J. Geophys. Res. Oceans, 87, C2, 1231-1238, doi:10.1029/JC087iC02p01231.

Pavolonis, M. J., J. Sieglaff and J. Cintineo (2015). Spectrally enhanced cloud objects – A generalized framework for automated detection of volcanic ash and dust clouds using passive satellite measurements: 1. Multispectral analysis, J. Geophys. Res. Atmos., 120, 15, 7813-7841, doi:10.1002/2014JD022968.

Petracca, I., D. De Santis, M. Picchiani, S. Corradini et al. (2022). Volcanic cloud detection using Sentinel-3 satellite data by means of neural networks: The Raikoke 2019 eruption test case, Atmos. Meas. Tech., 15, 7195-7210, doi:10.5194/amt-15-7195-2022.

Pitari, G. and E. Mancini (2002). Short-term climatic impact of the 1991 volcanic eruption of Mt. Pinatubo and effects on atmospheric tracers, Nat. Hazards Earth Syst. Sci., 2, 91-108, doi:10.5194/nhess-2-91-2002.

Poland, M. P., T. Lopez, R. Wright and M. J. Pavolonis (2020). Forecasting, detecting, and tracking volcanic eruptions from space, Remote Sens. Earth Syst. Sci., 3, 55-94, doi:10.1007/s41976-020-00034-x.

Prata, A. J. (1989a). Infrared radiative transfer calculations for volcanic ash clouds, Geophys. Res. Lett., 16, 1293-1296, doi:10.1029/GL016i011p01293.

Prata, A. J. (1989b). Observations of volcanic ash clouds in the 10-12 μm window using AVHRR/2 data, Int. J. Remote Sens., 10, 4-5, 751-761, doi:10.1080/01431168908903916.

Prata, A. J. and I. F. Grant (2001). Retrieval of microphysical and morphological properties of volcanic ash plumes from satellite data: Application to Mt. Ruapehu, New Zealand, Q. J. R. Meteorol. Soc., 127, 2153-2179, doi:10.1002/qj.49712757615.

Prata, A. J. (2009). Satellite detection of hazardous volcanic clouds and the risk to global air traffic, Nat. Hazards, 51, 303-324, doi:10.1007/s11069-008-9273-z.

Prata, F. and M. Lynch (2019). Passive Earth observations of volcanic clouds in the atmosphere, Atmosphere, 10, 199, doi:10.3390/atmos10040199.

Pugnaghi, S., L. Guerrieri, S. Corradini and L. Merucci (2016). Real-time retrieval of volcanic cloud particles and SO_2 by satellite using an improved simplified approach, Atmos. Meas. Tech., 9, 3053-3062, doi:10.5194/amt-9-3053-2016.

Rey-Devesa, P., J. Carthy, M. Titos, J. Prudencio et al. (2024). Universal machine learning approach to volcanic eruption forecasting using seismic features, Front. Earth Sci., 12, doi:10.3389/feart.2024.1342468.

Saunders-Shultz, P., T. Lopez, H. Dietterich and T. Girona (2024). Automatic identification and quantification of volcanic hotspots in Alaska using HotLINK: The hotspot learning and identification network, Front. Earth Sci., 12, doi:10.3389/feart.2024.1345104.

Sawada, Y. (1987). Study on analyses of volcanic eruptions based on eruption cloud image data obtained by the Geostationary Meteorological Satellite (GMS), in Technical Report of Meteorological Research Institute, Meteorological Research Institute, Japan, 22, 335.

Sawada, Y. (1996). Detection of explosive eruptions and regional tracking of volcanic ash clouds with Geostationary Meteorological Satellite (GMS), in Monitoring and Mitigation of Volcano Hazards, Springer, Berlin, Heidelberg, 299-314, doi:10.1007/978-3-642-80087-0_9.

Schmetz, J., D. Klaes, M. Koenig and K. Holmlund (2007). Monitoring weather and climate with the Meteosat and Metop satellites, Revista de Teledetección, 27, 5-16.

Scollo, S., A. Boselli, M. Coltelli, G. Leto et al. (2012). Monitoring Etna volcanic plumes using a scanning LiDAR, Bull. Volcanol., 74, 2383-2395, doi:10.1007/s00445-012-0669-y.

Simonyan, K. and A. Zisserman (2015). Very deep convolutional networks for large-scale image recognition, doi:10.48550/arXiv.1409.1556.

Tan, D., D. Fee, A. Witsil, T. Girona et al. (2024). Detection and characterization of seismic and acoustic signals at Pavlof Volcano, Alaska, using deep learning, J. Geophys. Res. Solid Earth, 129, e2024JB029194, doi:10.1029/2024JB029194.

Torrisi, F., E. Amato, C. Corradino, S. Mangiagli et al. (2022). Automatic detection of volcanic ash clouds using MSG-SEVIRI satellite data and machine learning techniques, Nuovo Cim. C, 45, 1-10, doi:10.1393/ncc/i2022-22081-0.

Torrisi, F., E. Amato, C. Corradino, S. Mangiagli et al. (2022). Characterization of Volcanic Cloud Components Using Machine Learning Techniques and SEVIRI Infrared Images, Sensors, 22, 7712, doi:10.3390/s22207712.

Torrisi, F., E. Amato, C. Corradino and C. Del Negro (2023). The FastVRP automatic platform for the thermal monitoring of volcanic activity using VIIRS and SLSTR sensors: FastFRP to monitor volcanic radiative power, Ann. Geophys., 65, 1, doi:10.4401/ag-8823.

Torrisi, F., C. Corradino, S. Cariello and C. Del Negro (2024). Enhancing detection of volcanic ash clouds from space with convolutional neural networks, J. Volcanol. Geotherm. Res., 448, 108046, doi:10.1016/j.jvolgeores.2024.108046.

Torrisi, F., C. Corradino, S. Cariello, G. Di Bella et al. (2024). Challenges and Opportunities in Deep Learning and Geostationary Satellite Remote Sensing for Volcanic Cloud Monitoring, 2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI), Milano, 120-124, doi:10.1109/RTSI61910.2024.10761715.

Torrisi, F., C. Corradino and C. Del Negro (2025). Exploring the Relationship Between Volcanic Radiative Power and Plume Height Using Satellite Observations, Nuovo Cim., accepted for publication.

Wilson, G., T. M. Wilson, N. I. Deligne and J. W. Cole (2014). Volcanic hazard impacts to critical infrastructure: A review, J. Volcanol. Geotherm. Res., 286, 148-182, doi:10.1016/j.jvolgeores.2014.08.030.

***CORRESPONDING AUTHOR: Federica TORRISI**,
Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo, Catania, Italy
e-mail: federica.torrisi@ingv.it