Pumice Raft Detection Using Machine-Learning on Multispectral Satellite Imagery

Maggie Zheng[1], Tushar Mittal[1], Kristen E. Fauria[2], Ajit Subramaniam[3], Martin Jutzeler[4]

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[2]Department of Earth and Environmental Sciences, Vanderbilt University, Nashville, Tennessee, USA
[3]Lamont-Doherty Earth Observatory, Palisades, New York, USA
[4]Centre for Ore Deposit and Earth Sciences (CODES), University of Tasmania, Hobart, Australia

Key Points:

- We use Google Earth Engine and a Random Forest classifier for an online pumice raft detection algorithm in high resolution satellite imagery
- We find that $\sim 16\%$ of Sentinel-2 images from Rabaul over 2017-2020 show pumice rafts. We interpret these to form by pumice remobilization.
- Pumice raft remobilization is potentially a common process for many near shore volcanic systems with historical pumice-forming eruptions.

Corresponding author: Maggie Zheng, mzhengxi@mit.edu

Abstract

Most of Earth's volcanic eruptions occur underwater, and these submarine eruptions can significantly impact large-scale earth systems. In this study, we develop a new semi-automated analysis framework to detect submarine eruptions through the supervised classification of satellite images on Google Earth Engine (GEE). We present a case study from the Rabaul caldera region in Papua New Guinea and find a large number of new unreported pumice rafts (in ∼16% images from 2017–present). After analysis of the spatial pattern of raft sightings and ancillary observations, we interpret that these rafts are not the result of a new eruption. Instead, we posit that the observed rafts represent remobilization of pumice clasts from previous historical eruptions. This novel process of raft remobilization may be common at near-shore/partially submarine caldera systems (e.g., Rabaul, Krakatau) and has significant implications for new submarine eruption detection, volcanic stratigraphy, and biological dispersal by rafts.

Plain Language Summary

Submarine volcano eruptions can significantly impact large-scale earth systems, but are challenging to detect. In this paper, we describe a new methodology to detect pumice rafts, a key signature of submarine volcano eruptions, in satellite imagery using a machine-learning classification algorithm. We apply our methodology to the Rabaul caldera region in Papua New Guinea and find a large number of unreported rafts in the last 5 years. After analyzing additional datasets such as seismicity, presence of discolored water, and reports from the local volcano observatory, we find that the our raft detections are most likely not generated from new eruptions, but instead are likely secondary rafts mobilized tens or hundreds of years after the original eruption. This secondary raft process is a novel process that has not been fully documented in the modern/satellite era, but is likely relevant to the dispersal of eruptive products in many coastal volcanic systems.

1 Introduction

Submarine volcanism is an important driver for Earth's climate and geochemical cycles (Embley et al., 2004; Kelley, 2017; Santana-Casiano et al., 2013; Tilstone et al., 2014; Mittal & Delbridge, 2019). For instance, submarine eruptions inject ash, pumice, and magmatic volatiles (with nutrients such as Fe) into the water column and the atmosphere (J. D. L. White et al., 2015). Pumice rafts, one of the key signatures of some submarine eruptions, can transport volcanic products and marine organisms across thousands of kilometers, potentially dispersing nutrients and increasing microbial biomass and biodiversity in areas geographically distant from the site of the eruption (Risso et al., 2002; Bryan et al., 2012). Pumice rafts can also be a significant hazard for human maritime activities, especially by clogging harbors and affecting near shore sealife and fisheries, disrupting local economies (Bryan et al., 2012; Jutzeler et al., 2020, Also the ongoing raft dispersal from the Fukutoku-Okanoba 2021 eruption).

Modern day submarine volcanism includes both mid-ocean ridge and ocean-island volcanism as well as fully/partially submarine subduction zone volcanism (e.g., Kermadec-Tonga Arc, Izu-Bonin Arc, Papua New Guinea Arc) (S. M. White et al., 2006; GVP, 2013). Overall, submarine volcanism potentially represents the majority ($> 70\%$) of Earth's present-day volcanism (S. M. White et al., 2006; J. D. L. White et al., 2015; ERUPT, 2017; Rubin et al., 2012). However, we have historically detected only a very small fraction of expected underwater eruptions. In fact, only ∼10% of all eruptions in the Smithsonian Global Volcanism database (GVP, 2013) over the past 100 years are submarine (Fig 1A, S. M. White et al. (2006)) and the majority of detected eruptions are shallow (< 100 meters water depth; Fig 1A).

One of the key reasons for this strong bias in our submarine eruption detection ability is the remote location of submarine volcanoes. Although multiple new approaches have been proposed to improve submarine volcanism detection, including hydro-acoustics (e.g., Tepp et al., 2019; Heaney et al., 2013), seismic and ground deformation with ocean bottom seismometers (e.g., Wilcock et al., 2016; Cesca et al., 2020; Matsumoto et al., 2019; Tepp & Dziak, 2021), and ocean thermal anomalies (e.g., Mittal & Delbridge, 2019; Baker et al., 1989), challenges remain due to limited global instrumental coverage. In this study, we describe another dataset — satellite imagery — that can be used to efficiently detect and characterize products of submarine volcanism.

Automated satellite image analysis has already proven to be very useful for global **subaerial** eruption detections based on thermal anomalies and ash-rich subaerial plumes (e.g., Wright et al., 2004; Furtney et al., 2018; Poland et al., 2020; Engwell et al., 2021). However these methods are not adapted for submarine eruptions where the presence of water obscures/reduces these signatures. Satellite imagery has been used to map the eruptive products (e.g., pumice, ash, hydrothermal fluids) from submarine eruptions (e.g., O'Malley et al., 2014; Whiteside et al., 2021; Sakuno, 2021; Jutzeler et al., 2020, 2014; Bryan et al., 2004) on an event-by-event basis. For example, the pumice raft from the recent Tonga 2019 eruption was tracked in near-real-time by Sentinel-2 ($\sim$ 10 m/pixel) and Landsat 8 ($\sim$ 30 m/pixel) satellite imagery (Jutzeler et al., 2020). However, this was done by manual hand-tracing and visual tracking by sifting through various satellite images. Although this process is fairly accurate for large rafts, it introduces subjectivity in tracing, especially for smaller rafts. Consequently, it is difficult to quantify uncertainties and biases across different studies. An ancillary challenge with using satellite imagery is the large data volume associated with extensive satellite collections. For example, a single day in the Tonga region is composed of about forty individual 100 km x 100 km image granules, each containing about 600 MB of data. Thus, analyzing entire global collections over extended time periods, and for multiple different satellites, would require handling enormous amounts of data and requisite computing resources.

Our study aims to address these challenges of submarine eruption detection by developing a semi-automated Machine-Learning (ML) based methodology using global, publicly available, high resolution ($\sim<$ 30 m/pixel) satellite data products (Fig 1C). This method utilizes Google Earth Engine (Gorelick et al., 2017), in order to remove the large data storage need that is typical for analyzing satellite collections. Our primary focus is on detecting rafts formed from floating pumices emitted by intermediate to silicic volcanism, but our approach can be applied to other signatures of submarine eruptions (e.g., discolored water from hydrothermal fluids). Our analysis is complementary to recent work on detection of large submarine eruptions using specific global, low resolution ($>$ 250 m/pixel) satellite products (Whiteside et al., 2021; O'Malley et al., 2014; Qi et al., 2020).

As a complementary question, we also seek to examine whether individual pumice raft detections necessarily indicate a new eruption. Previous work has examined the remobilization of pumice clasts, which may have been deposited in the area immediately surrounding the vent, or stranded a distance away after traveling as a raft, following large eruptions (Mandeville et al., 1996; Manville et al., 2002; Jutzeler et al., 2020; Shane et al., 1998). Using our automated detection algorithms, we can improve raft detection. This allows us to assess how long rafts can persist after initial pumice formation and the mechanisms that would enable pumice to be rafted tens to hundreds of years after the original eruption (Brasier et al., 2011; Bryan et al., 2012).

In Section 2, we describe our detection algorithm and its implementation in Google Earth Engine. In Section 3, we illustrate our method's accuracy using satellite imagery from the 2019 Tonga submarine eruption (Jutzeler et al., 2020) and then use our method to analyze pumice rafts in a region close to the Rabaul volcano in Papua New Guinea (partially submerged caldera). In Section 4, we discuss what our new pumice raft detections from Rabaul suggest in regards to suspension of pumice material, potentially

from pumice clasts or rafts previously washed up on shores or eroded on riverbanks or coastal cliffs. Finally, we briefly discuss areas for future algorithmic improvements.

2 Methods

2.1 Google Earth Engine

We developed and implemented our Machine Learning (ML) detection algorithm for pumice raft detection on the Google Earth Engine platform (GEE, Gorelick et al., 2017). GEE is a web-based, publicly available platform that enables access to a vast catalog of satellite images and the resources to run global-scale analyses without the need to download or export large amounts of data. There are various satellite collections offered through GEE, such as low resolution (MODIS, Sentinel-2) and medium-high resolution imagery (Landsat, Sentinel-2). Although some other super-high-resolution image collections are available outside of GEE (e.g., Planet labs - 3m/pixel, Digital Globe - 50 cm/pixel), they are typically not publicly available without commercial licenses. Thus, for this study, we have primarily focused on using GEE resources for the ML algorithm.

Specifically, we use GEE collections from the Sentinel-2 Multi-Spectral Instrument (MSI) as our baseline satellite product. Sentinel-2 (a pair of two satellites, each with MSI instrumentation) offers both high-resolution imagery (10-60 m/pixel), good coverage in regions of interest, and a relatively frequent repeat time ($\sim$ global 5-day revisits; See Supplementary Text S6). Sentinel-2 data products are also freely available through the European Space Agency's Copernicus Open Access Hub as well as other cloud environments.

For our study, we chose to use Sentinel-2 as its high resolution imagery could be used to detect much smaller rafts than a lower resolution satellite (e.g. MODIS). In addition, Sentinel-2's MSI collects data across 13 different spectral bands, with finer spectral coverage than other high resolution satellite image collections (e.g., Landsat 7 and 8) (See spectral response curve for Landsat 8 image of Puyehue-Cordón Caulle pumice in Fig 1B). An initial method using thresholds on only the visual bands to detect pumice rafts was insufficient, so the additional spectral bands are necessary in our ML algorithm (Supplement S1). As illustrated by the variable importance in the Random Forest classifier (Supplementary Figure S11), the multi-wavelength information is critical for accurate classification with a dominant role of the visible bands. Our overall methodology is general and can be applied to other satellite collections in the future (Supplement S10).

2.2 Machine-Learning Algorithm

To identify spectral characteristics that can be used to classify Sentinel-2 image pixels as pumice rafts, we generated spectral response curves for pumice and other categories of interest (Fig 1B). Spectral response curves record the mean reflectance or brightness of an image pixel for a range of wavelengths. We used the Tonga pumice raft from August 11, 2019 to generate the spectral response curves (Fig 1B), as the particular eruption and the associated raft has been extensively analyzed by previous work (Brandl et al., 2020; Jutzeler et al., 2020). We also show the variance around the mean spectral response curve calculated for all of the pixels for each class (pumice, water, light clouds).

A key result from this analysis is that there is a significant difference between the spectral response curves of pumice, water, and light clouds (Fig 1B). Additionally, we find relatively minor (compared to differences with other classes) variation in the reflectance from pumice pixels within a single geo-temporal area, such as a specific day in Tonga (Fig 1B) or comparing across multiple days for the same raft (Supplement Fig S4). Although there is some variation in pumice spectral response curves when comparing rafts from different chemical compositions, sources, and times (Fig 1B, comparison with Rabaul raft and Puyehue-Cordón Caulle raft), the general shape of the reflectance curve remains

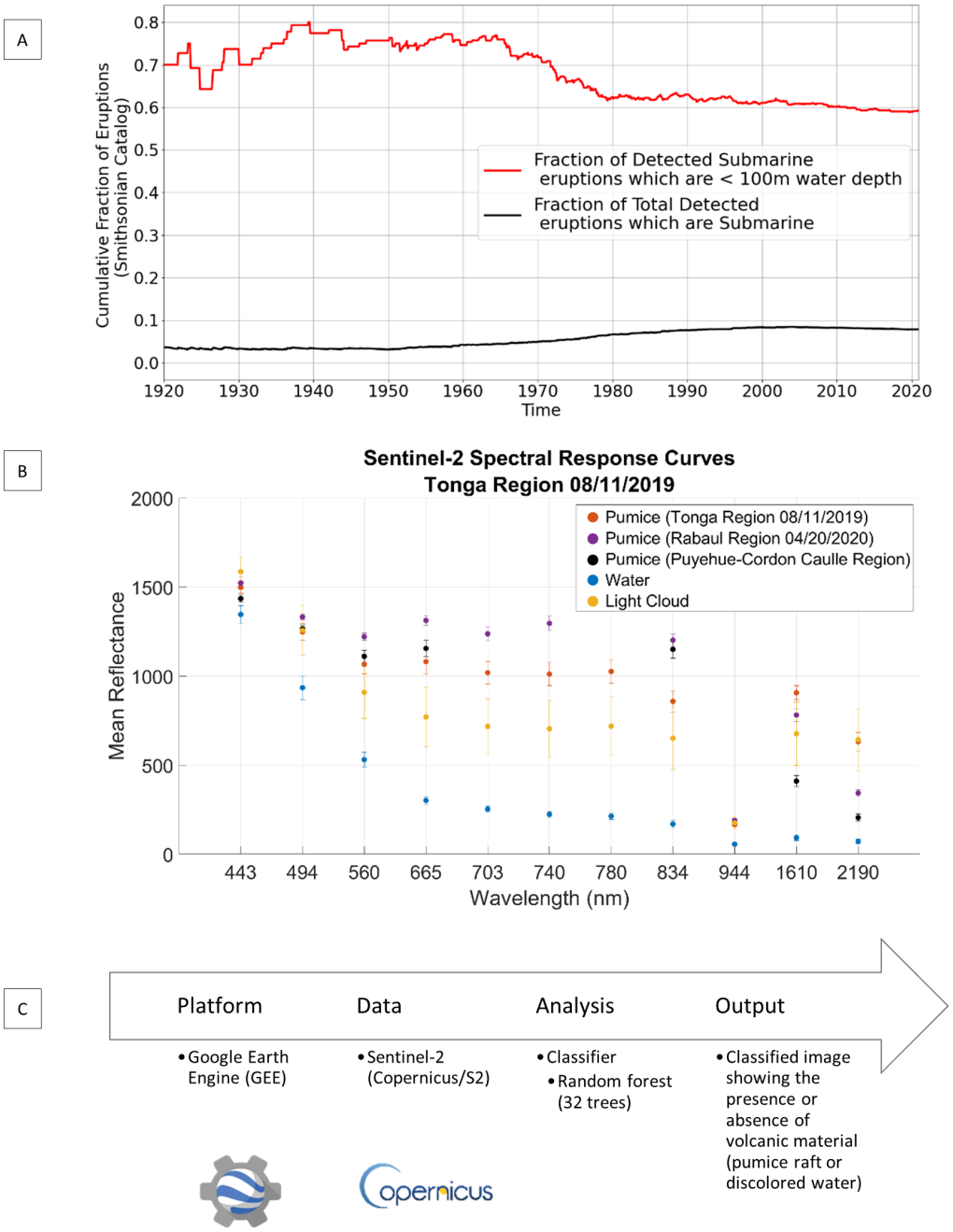

Figure 1. (A) Fraction of all detected submarine eruptions out of total eruptions and fraction of detected shallow submarine eruptions (less than 100 meter water depth) out of all detected submarine eruptions (GVP, 2013). Only a small fraction of submarine eruptions are pumice-forming. (B) Mean spectral response curves generated for a Sentinel-2 raft image in the Tonga region (August 11, 2019). Error bars are generated from the standard deviation measured for each wavelength. Spectral response curves for pumice from the Rabaul region (April 20, 2020) and a lake near Puyehue-Cordón Caulle are also provided for comparison. For Puyehue-Cordón Caulle, Sentinel-2 imagery was not available, so Landsat 8 imagery was used instead, and mean reflectance values were averaged between two dates (June 19, 2013 and October 5, 2013). (C) Schematic of workflow used in this study.

very similar. This characteristic shape of the spectral response curve for pumice pixels allows for an algorithm to identify pumice and differentiate from other classes (e.g., water, clouds) across a broad range of regions and time periods. Details for the Puyehue-Cordón Caulle raft are provided in Supplement Text S13.

Our machine-learning algorithm uses a Random Forest (RF) classifier to read in an RGB Sentinel-2 image and return a classified image, where each pixel is colored according to the assigned class. The algorithm specifics are detailed in Supplementary Text S4. Since RF is a supervised learning algorithm, we need to train it on a set of manually demarcated and classified pixels. Our primary training data for pumice, ocean water, light cloud cover, and heavy cloud cover was sampled from the Tonga raft on August 11, 2019 (Fig 2A, only a small part of the raft pixels were used for training). We also included additional data from a Sentinel-2 scene of Rabaul, Papua New Guinea, on April 20, 2020 (Fig 2B). This image includes a large, distinct pumice raft as well as ocean water, light cloud cover, pumice mixture classes, and two different discolored water classes (additional information for the discolored water classes are included in Supplementary Text S9). Since the discolored water classes are not the primary focus of this study, our primary optimization for the RF algorithm was to ensure accurate detection of pumice rafts.

3 Results

3.1 Single Image Analysis Results

We applied our classification algorithm to Sentinel-2 images from different geo-temporal regions to test model accuracy (Fig 2). In the Tonga area on August 11, 2019 (Fig 2A), the classifier displays pumice pixels in red, water in blue, light cloud cover in orange, and heavy cloud cover in white. The shape of the large raft is distinctly visible in the classified image. In the Rabaul region, on April 20, 2020 (Fig 2B), the classifier also includes additional classes: mixed/faint pumice — a mixture of water and pumice — shown in light blue, and two different classes of discolored water shown in turquoise and magenta. Overall, our algorithm is efficient at identifying pumice from other backgrounds. Algorithm validation methods and results are included in Supplementary Text S5.

3.2 Regional Results

To assess the utility of our algorithm for new submarine eruption detection, we applied the classifier over a single region for an extended period of time. We focused on Rabaul, a partially submarine volcano located on the Gazelle Peninsula's tip at the northeast end of New Britain in Papua New Guinea (Fig 3A). The Rabaul caldera ($\sim$ 8 x 14 km size) was formed as a consequence of multiple large explosive eruptions in the past few hundred thousand years, with the present day shape due to an eruption $\sim$ 1400 years ago (GVP, 1994b). The caldera is mostly shallow submarine ($<$ 200 m water depth) and is connected to the sea on the east through a wide opening (Blanche Bay). The main raft-forming eruptions for this volcano occurred in 1878, 1937, and 1994, and no raft formation has been recorded since 1994 (GVP, 1994b, 1994a, 2006). No activity has been recorded at either of the main vents (Vulcan and Tavurvur) since 2014 (Bernard & Bouvet de Maisonneuve, 2020). More detailed eruption history is provided in Supplementary Text S8.

In the Rabaul area, we applied our algorithm from November 2015 (start of the Sentinel-2 coverage for the Rabaul region) to August 2020 — a total of 239 distinct days with images. More details on our algorithm application method are included in Supplementary Text S7.

Of these 239 days, we found that 74 days were too cloudy for the classifier to detect any pumice meaningfully. Cloudy days were filtered out by manually examining clas-

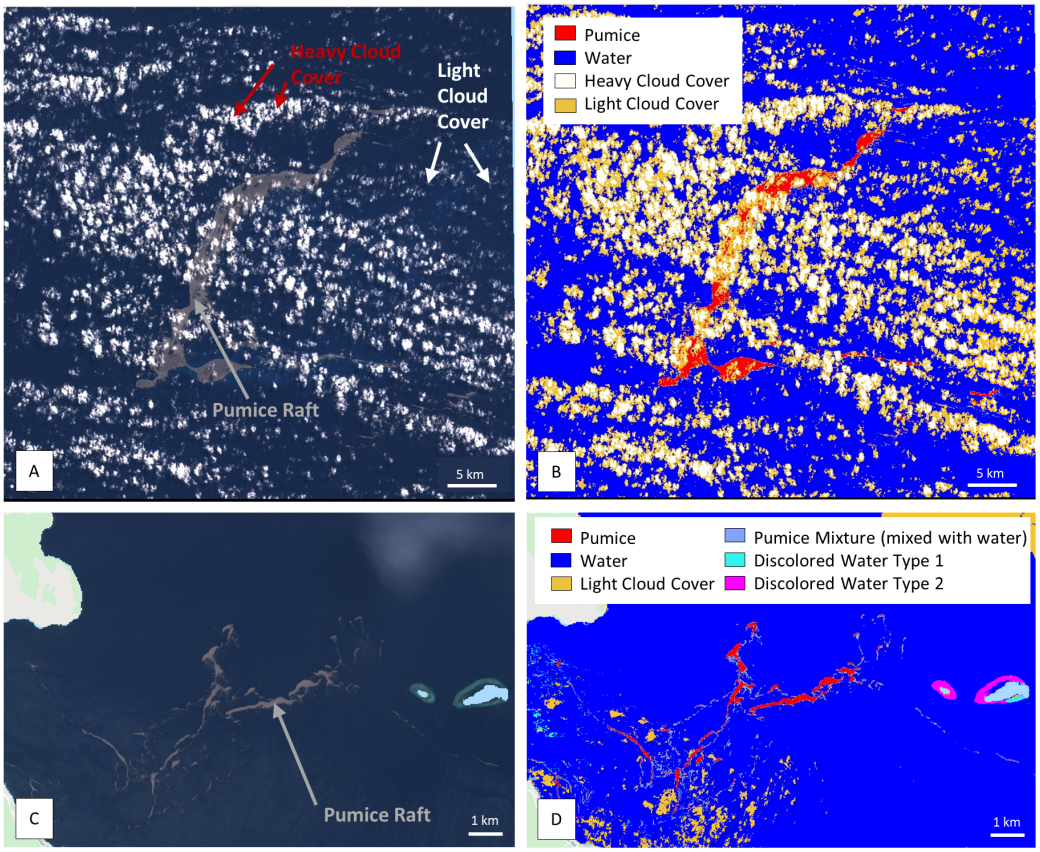

Figure 2. (A) Land-masked RGB image of Tonga region on 08/11/2019 (B) Classified image of Tonga region on 08/11/2019 (C) Land-masked RGB image of Rabaul on 04/20/2020 (D) Land-masked classified image of Rabaul on 04/20/2020

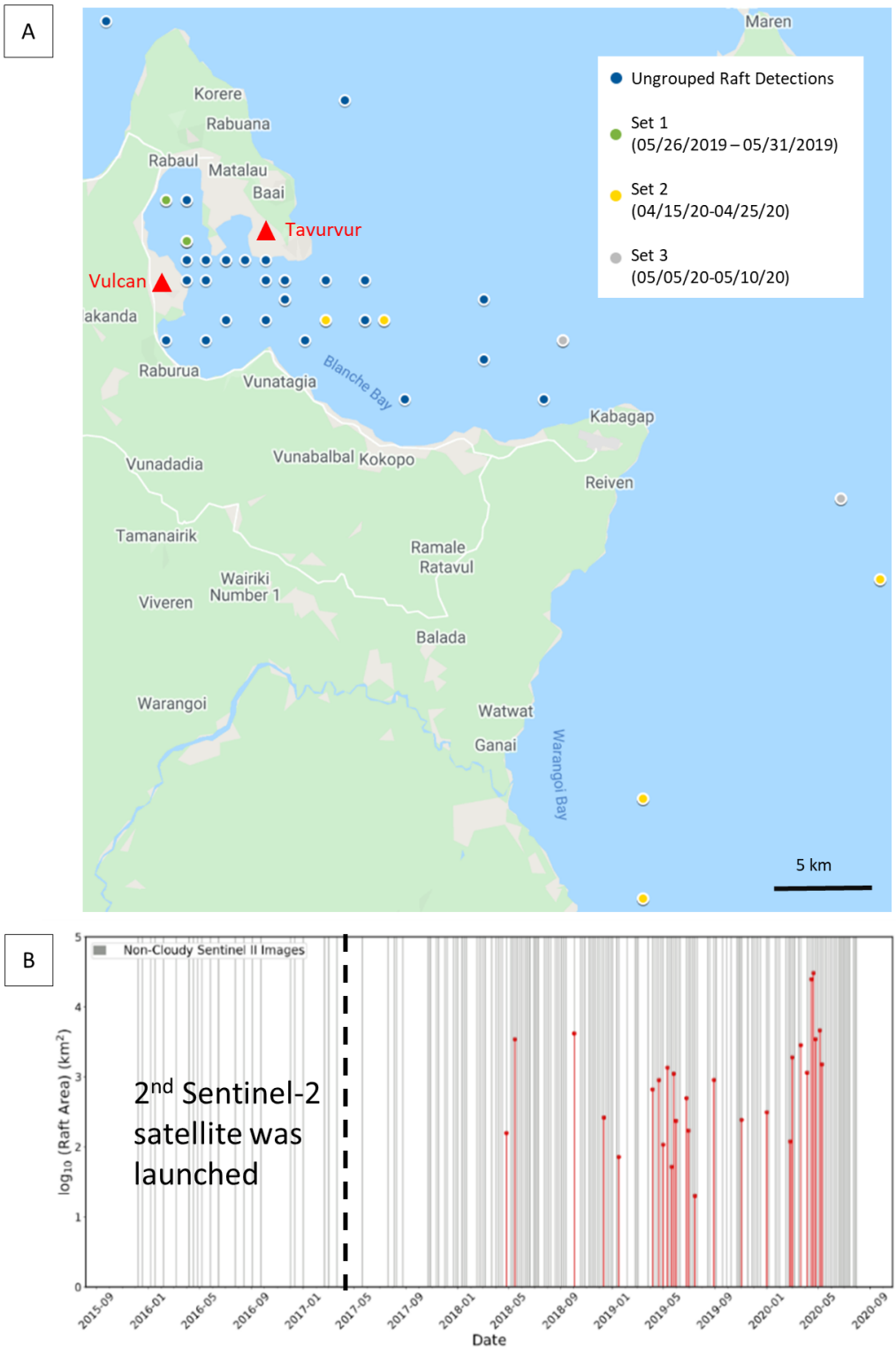

Figure 3. (A) Map of pumice raft detection locations in Rabaul area. Main vent locations are marked with red triangles. Groupings of raft detections (by proximity in time and location) are indicated by the different colored icons (B) Non-cloudy days during Sentinel-2 coverage period, with raft detections in red (height indicating raft area)

sified images and removing images in which every pixel was labeled as heavy or light cloud cover. In the future, this step can be automated by explicitly filtering the images based on the classified heavy cloud fraction. We detected rafts in 28 (red lines, Fig 3B) of the remaining 165 days (gray lines, Fig 3B), leading to a detection rate of 16.97%. As illustrated in Fig 3B, most of our raft detections were after January 2018 (Fig 3B). This is likely a consequence of increased revisit frequency ($\sim$ 5-day) after the second Sentinel-2 satellite launch. Before 2018, when only one Sentinel-2 satellite was in operation, there are significantly fewer images available. It is noteworthy that none of the pumice rafts detected in our analysis had been previously reported in the scientific literature (to the best of our knowledge) or the Smithsonian Global Volcanism Catalog (GVP, 2013). The sizes of our detected rafts varied greatly, with raft areas as small as 20 km^2 to as great as 10,000 km^2.

3.2.1 Source of New Rafts in Rabaul

Given our new raft detections, there is a natural follow-up question — do these rafts represent previously unreported submarine eruptions, or are they suspended pumice remobilized from known previous eruption deposits from Rabaul? These are the two primary end-member models, with the latter being a process that has been documented during the dispersal of large pumice rafts (e.g., the interaction of Tonga 2019 raft with islands in Fiji (Jutzeler et al., 2020)). Redeposition and remobilization of volcanic products such as ash fall (Del Bello et al., 2021; Etyemezian et al., 2019) and subaqueous pyroclastic material has been recorded after initial deposition (Mandeville et al., 1996; Manville et al., 2002; Park & Schmincke, 2020). However, most of this work has focused on a short time frame—on the order of days to weeks or months after the eruption (Bryan et al., 2012, 2004). Here, if our hypothesis is correct, the rafts we are seeing can form tens or even hundreds of years after the original eruption since the last major raft forming eruption in Rabaul was in 1994.

We assess the likelihood of new submarine eruptions by analyzing the reported volcanic activity for Rabaul in the Smithsonian Volcano Catalog (GVP, 2013). The Rabaul Volcano Observatory has recorded no large eruptions since 2014 (GVP, 2013) and/or any significant submarine activity besides hydrothermal discharge near the Tavurvur vent. Because rafts initiating from point sources can indicate new eruptions, we test this further by recording each of our raft detections' spatial location and considering the spread of each sighting. We have tried to manually aggregate three sets of raft locations together. These sets are of sequential images, in which the raft detections were somewhat close, not only in time, but in location as well. Conclusively tracking the rafts as they are advected around by local ocean currents is challenging due to repeat frequency (5-day gap between images), cloud cover, and complex shallow-water ocean currents in the regions. In aggregate, the detections are scattered over a broad area in the caldera and surrounding sea, rather than primarily located near any known vents (Fig 3A).

We also used Sentinel-2 imagery as well as ancillary datasets (e.g., higher spatial and temporal resolution Planet Labs imagery (Planet Team, 2018–2021) to check if the rafts are associated with any other eruptive signatures expected for shallow submarine eruptions (e.g., aerial plumes, discolored water). We did not find any aerial plumes and, while there was some discolored water around the Tavurvur vent location, we did not find any relationship between the days with raft detections and days with discolored water around the vent (Supplement Text S14). Thus, we interpret that the detected rafts are not actually products from a new submarine eruption. Instead, we propose that they are secondary rafts (Osborne et al., 1991) that have been suspended after being deposited on surrounding shores and riverbanks following their initial eruptions tens to hundreds of years ago (see Discussion section for the potential process). *This is a new, novel physical process that has not been fully documented before, especially in the modern/satellite era. Our analysis is the first study, to the best of our knowledge, to carefully document*

the secondary raft process on timescales of years or longer using satellite imagery, further validating the importance of this process as suggested by previous studies (Bryan et al., 2012; Jutzeler et al., 2014; Shane et al., 1998; Osborne et al., 1991; Pullar et al., 1977). This secondary raft process is likely relevant for the dispersal of eruptive products from many volcanic systems in coastal regions (e.g., Krakatau in Indonesia, Loisel pumice deposits in New Zealand (Shane et al., 1998)). Without any specimens of the pumice that we detected, we are presently unable to ascertain a specific source eruption of the rafts. In addition, even with samples, it may still be difficult to determine the original source eruption or eruptions, as pumices from the 1878 eruption and subsequent eruptions have very similar overall composition and texture (Bernard & Bouvet de Maisonneuve, 2020).

4 Discussion

4.1 Source of Pumice Remobilization: Influence of Weather Factors

Considering our interpretation that our detected rafts in Rabaul are secondary rafts (Osborne et al., 1991), an important question to consider is what potential physical mechanisms are responsible for the pumice mobilization. One possibility is that delayed suspension is a consequence of local climatological conditions, e.g., high rainfall events, high wind conditions that dislodge pumice along coastlines and riverbanks back into the water. Local weather can lead to landslides and dislodgement of small pumice rafts (e.g., local pumice raft from Rockslide in the Askja caldera, Iceland on July 21 2014 (Icelandic Meteorological Office, 2014)). Using ERA5 Daily Aggregate Reanalysis Product (Hersbach et al., 2020) (directly accessible through GEE), we generated time series of various atmospheric properties — daily mean air temperature, wind magnitude, wind direction, and precipitation. These time series were all sampled from the same location, directly on top of one of Rabaul's vents, and the time series spanned the entire Sentinel-2 coverage period in the area. We did not observe any significant correlation between the daily mean air temperature and the detection of pumice rafts in the area (See Supplement Fig S3). We also explored potential correlations with weather parameters up to 10 days before raft detection to allow for some unknown advection time (See Supplementary Data file). Overall, we did not find significantly different results across these windows. The main statistically robust relationships in our dataset are between raft detection and wind and precipitation.

4.1.1 Precipitation

To compare the impact of wind, precipitation, and other weather parameters on raft detection, we construct and compare probability density distributions (PDF). A PDF is a function that provides the relative likelihood of an event (raft detection) given another parameter (e.g. wind speed, recorded rainfall). We find that the PDF for the days with sighted pumice rafts (red curve, Fig 4A) were slightly different from the curves for the total days in the coverage period (blue curve, Fig 4A) and the days where no rafts were detected (gray curve, Fig 4A) (using 5 day rolling window, other windows have similar results). However, this difference is not statistically significant when using either the Anderson-Darling (AD) test statistic (Scholz & Stephens, 1987) or the Epps-Singleton (ES) test statistic (Epps & Singleton, 1986). The medians of the raft vs. non-raft precipitation PDFs are potentially different, as shown by the lower p-value for the Kruskal-Wallis test (Kruskal & Wallis, 1952)). We also do not find any clear correlation between precipitation values and raft area.

We analyzed the long-term precipitation history in the Rabaul area to help elucidate the remobilization process. We used ERA5 data to consider 3-day rolling sums of precipitation values in Rabaul since 1990. From the long-term history, we observe the peak precipitation occurred in February 2018. Our algorithm only detected rafts after

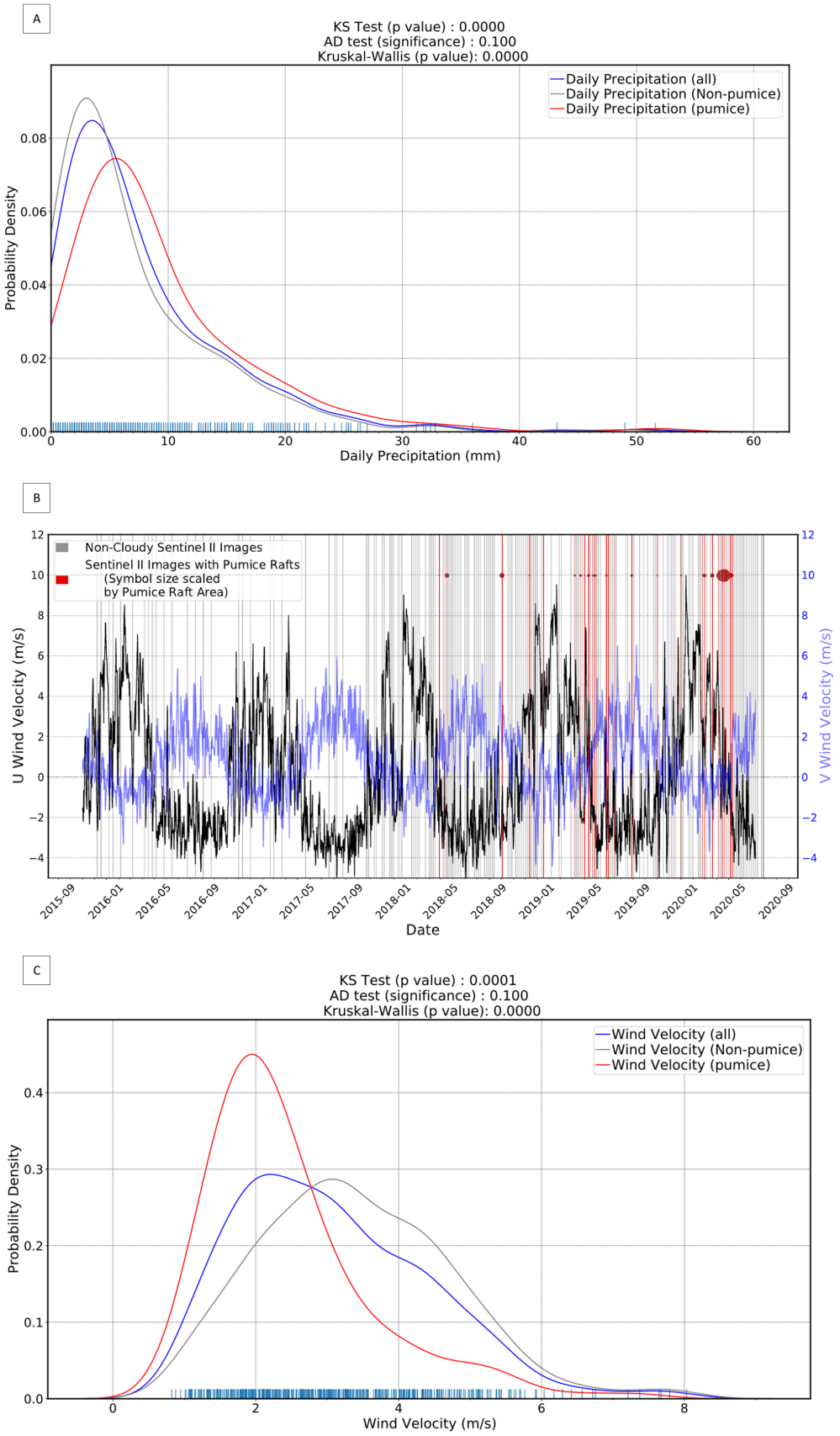

Figure 4. (A) Probability density function for the daily precipitation in the Rabaul region (5 day rolling window) (B) Time series of daily wind directionality in the Rabaul region—U wind velocity is the eastward component of wind while the V wind velocity is the northward component (C) Probability density function for daily wind magnitude in the Rabaul region (5 day rolling window)

2018. Some of this can be attributed to the increased frequency of sampling in the area after the second Sentinel-2 satellite launch in March of 2017. However, the precipitation peak in early 2018 potentially suggests that a large storm or significant weather event may have made rafts easier to remobilize post-2017. Further analysis using Landsat imagery is included in Supplement Text S15.

Overall, we find that there is only a weak correlation between precipitation and raft detection. We posit that the slightly higher values for precipitation before raft detection compared to non-raft days suggest a role for higher precipitation to increase erosion and consequently encourage raft remobilization. However, it is clear that precipitation is not a unique factor since days of high precipitation are not always followed by raft detections (Fig 4A, Supplement Fig S2).

4.1.2 Wind

In addition to precipitation, we also considered the role of wind in raft formation. Fig 4B shows the daily wind direction (U (eastward), V (northward) components) in Rabaul along with red vertical lines highlighting days with raft detections. We find that the general wind direction in Rabaul has a strong seasonal cycle which is relatively stable over the past five years. Interestingly, most of our pumice raft sightings were around the March-May window despite having a number of non-cloudy images for other months. This suggests that certain wind directions are more favorable for rafts to be advected off from the shores, or eroded from riverbanks and into the ocean.

In order to assess the role of overall wind magnitude, we show the probability distribution curves for wind velocity for all days in the Sentinel-2 coverage period (blue curve, Fig 4C), days without raft sightings (gray curve, Fig 4C), and days with raft sightings (red curve, Fig 4C) (using 5 day rolling window, other windows have similar results). We find that days where rafts were detected produced a significantly different probability distribution curve (Fig 4C). There is also a high correlation between wind amplitude and raft area (Supplement Fig S3). However, since there are not many high raft area data points, the correlation may be biased by outliers. Overall, we see most of our raft sightings are in the distinct range of wind velocities (1 m/s to 4 m/s) compared to the overall distribution. Even when accounting for different sample sizes, this difference is statically significant (Anderson-Darling (AD) test statistic (Scholz & Stephens, 1987); the Epps-Singleton (ES) test statistic (Epps & Singleton, 1986); Kolmogorov-Smirnov (KS) test (Hodges, 1958)). We conclude that the high wind velocities likely break up and disperse the secondary rafts too rapidly for Sentinel-2 to capture.

4.2 Open Challenges for Global Pumice Detection Algorithm

Although our ML algorithm is reasonably successful for pumice raft detection, it is not fully automated. The classification process requires manual checks to filter out incorrect classifications of pumice and cloud cover. In particular, the light cloud cover with a flat spectral response curve can at times be misclassified as pumice (and vice versa). Also, the satellite's viewing geometry may create a "sun glint" in certain images, where all of the pixels in the RGB rendering are affected and off-colored. The classifier subsequently has difficulty correctly identifying the correct class of each pixel. There are some ways these issues can be addressed. Better atmospheric corrected products, specifically for oceanic regions, would help improve detection. For instance, in some cases, using the atmosphere corrected Surface Reflectance (Level-2A) product can allow us to detect pumice rafts on images discolored due to atmospheric effects (Supplement Fig S10). Alternatively, more stringent data filtering for satellite viewing angle and cloudiness bounds can help reduce potential false positives. Additional potential options for algorithmic improvement are described in Supplementary Text S11.

4.3 Conclusions

In this study, we show that GEE and RF classifiers can be successfully used to detect pumice rafts. This can be useful to efficiently track pumice rafts, which can pose as hazards and disruptions to boats and harbors (Jutzeler et al., 2014, 2020) and thus help with hazard mitigation and coordination services along populated shorelines (e.g., ongoing raft arrival on mainland Japan from the August 2021 Fukutoku-Okanoba eruption). Our methodology can help address our current strong bias in eruption detection and improve the detection of submarine eruptions globally. Using GEE removes the large data storage requirement and allows for a semi-automated, easily scalable classification with minimal subjective biases. Using the Rabaul caldera regions in Papua New Guinea as a regional case study, we show that new raft detections do not necessarily indicate a recent submarine eruption. Indeed, in some coastal regions, raft remobilization is likely to be a widespread phenomenon and can affect the spatial pattern of how products from an eruption are deposited. Since these spatial patterns serve as the basis for estimating volcanic eruptive histories and eruptive volumes, as well as long-distance stratigraphic correlations (Shane et al., 1998; Mouginis-Mark & Zimbelman, 2020; Freundt et al., 2021), the raft remobilization process can introduce significant distortions in our understanding of these histories. Understanding raft remobilization can also contribute to our knowledge about population connectivity (Bryan et al., 2012, 2004) and the long distance transport of microbial populations through the remobilization process.

Acknowledgments

M. Zheng acknowledges support from the MIT UROP program, T. Mittal acknowledges funding support from the Crosby Postdoc Fellowship at MIT, and K. Fauria acknowledges funding support from the NASA Grant 80NSSC20K1450. We thank Amber Madden-Nadeau, Samantha L. Engwell, Sebastian Watt, Michael Cassidy, Ralf Bennartz, Ashok Gupta, Liam Kelly, John Rausch for useful discussions and suggestions for the manuscript text. We thank (reviewers and editors) for their valuable comments and suggestions. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. There is no experimental or observational data associated to this study. Codes for the Google Earth Engine analysis, data files for the statistical analysis, and all the classified Sentinel-2 images can be downloaded from `https://figshare.com/projects/Pumice_Raft_Detection_Using_Machine-Learning_on_Multispectral_Satellite_Imagery/126466`. We thank Planet Labs, Sentinel-2, Landsat 7/8, and Google Earth Engine platform for providing the satellite imagery and computational tools.

References

Baker, E., Lavelle, J., Feely, R., Massoth, G., Walker, S., & Lupton, J. (1989). Episodic venting of hydrothermal fluids from the juan de fuca ridge. *Journal of Geophysical Research: Solid Earth*, *94*(B7), 9237–9250.

Bernard, O., & Bouvet de Maisonneuve, C. (2020). Controls on eruption style at rabaul, papua new guinea – insights from microlites, porosity and permeability measurements. *Journal of Volcanology and Geothermal Research*, *406*, 107068. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0377027320303346` doi: https://doi.org/10.1016/j.jvolgeores.2020.107068

Brandl, P. A., Schmid, F., Augustin, N., Grevemeyer, I., Arculus, R. J., Devey, C. W., ... Hannington, M. D. (2020). The 6–8 aug 2019 eruption of 'volcano f'in the tofua arc, tonga. *Journal of Volcanology and Geothermal Research*, *390*, 106695.

Brasier, M. D., Matthewman, R., McMahon, S., & Wacey, D. (2011). Pumice as a remarkable substrate for the origin of life. *Astrobiology*, *11*(7), 725–735.

Bryan, S. E., Cook, A., Evans, J., Colls, P., Wells, M., Lawrence, M., ... Leslie,

R. (2004). Pumice rafting and faunal dispersion during 2001–2002 in the southwest pacific: record of a dacitic submarine explosive eruption from tonga. *Earth and Planetary Science Letters*, *227*(1), 135-154. Retrieved from https://www.sciencedirect.com/science/article/pii/S0012821X0400500X doi: https://doi.org/10.1016/j.epsl.2004.08.009

Bryan, S. E., Cook, A. G., Evans, J. P., Hebden, K., Hurrey, L., Colls, P., ... Firn, J. (2012). Rapid, long-distance dispersal by pumice rafting. *PloS one*, *7*(7), e40583.

Cesca, S., Letort, J., Razafindrakoto, H. N., Heimann, S., Rivalta, E., Isken, M. P., ... others (2020). Drainage of a deep magma reservoir near mayotte inferred from seismicity and deformation. *Nature geoscience*, *13*(1), 87–93.

Del Bello, E., Taddeucci, J., Merrison, J. P., Rasmussen, K. R., Andronico, D., Ricci, T., ... Iversen, J. J. (2021). Field-based measurements of volcanic ash resuspension by wind. *Earth and Planetary Science Letters*, *554*, 116684.

Embley, R., Baker, E., Chadwick, W., Lupton, J., Resing, J., Massoth, G., & Nakamura, K. (2004). Explorations of mariana arc volcanoes reveal new hydrothermal systems. *Eos, Transactions American Geophysical Union*, *85*(4), 37–40.

Engwell, S., Mastin, L., Tupper, A., Kibler, J., Acethorp, P., Lord, G., & Filgueira, R. (2021). Near-real-time volcanic cloud monitoring: insights into global explosive volcanic eruptive activity through analysis of volcanic ash advisories. *Bulletin of Volcanology*, *83*(2), 1–17.

Epps, T., & Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistical Computation and Simulation*, *26*(3-4), 177–203.

ERUPT. (2017). *Volcanic eruptions and their repose, unrest, precursors, and timing*. Washington, DC: The National Academies Press. Retrieved from https://www.nap.edu/catalog/24650/volcanic-eruptions-and-their-repose-unrest-precursors-and-timing doi: 10.17226/24650

Etyemezian, V., Gillies, J., Mastin, L. G., Crawford, A., Hasson, R., Van Eaton, A. R., & Nikolich, G. (2019). Laboratory experiments of volcanic ash resuspension by wind. *Journal of Geophysical Research: Atmospheres*, *124*(16), 9534–9560.

Freundt, A., Schindlbeck-Belo, J. C., Kutterolf, S., & Hopkins, J. L. (2021). Tephra layers in the marine environment: A review of properties and emplacement processes. *Geological Society, London, Special Publications*, *520*.

Furtney, M. A., Pritchard, M. E., Biggs, J., Carn, S. A., Ebmeier, S. K., Jay, J. A., ... Reath, K. A. (2018). Synthesizing multi-sensor, multi-satellite, multi-decadal datasets for global volcano monitoring. *Journal of Volcanology and Geothermal Research*, *365*, 38–56.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18–27.

GVP. (1994a). *Report on rabaul (papua new guinea) (wunderman, r., ed.), bulletin of the global volcanism network, 19:10. smithsonian institution.* https://doi.org/10.5479/si.GVP.BGVN199408-252140. (Downloaded : 1st Feb 2021)

GVP. (1994b). *Report on rabaul (papua new guinea) (wunderman, r., ed.), bulletin of the global volcanism network, 19:8. smithsonian institution.* https://doi.org/10.5479/si.GVP.BGVN199408-252140. (Downloaded : 1st Feb 2021)

GVP. (2006). *Report on rabaul (papua new guinea) (wunderman, r., ed.), bulletin of the global volcanism network, 31:9. smithsonian institution.* https://doi.org/10.5479/si.GVP.BGVN200609-252140. (Downloaded : 1st Feb 2021)

GVP. (2013). *Global volcanism program, 2013. volcanoes of the world, v. 4.6.2. venzke, e (ed.). smithsonian institution.* http://dx.doi.org/10.5479/si.GVP.VOTW4-2013. (Downloaded : 10 Oct 2017)

Heaney, K. D., Campbell, R. L., & Snellen, M. (2013). Long range acoustic measure-

ments of an undersea volcano. *The Journal of the Acoustical Society of America*, *134*(4), 3299–3306.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... others (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049.

Hodges, J. L. (1958). The significance probability of the smirnov two-sample test. *Arkiv för Matematik*, *3*(5), 469–486.

Icelandic Meteorological Office. (2014). *Rockslide in askja, july 21 2014 - preliminary results of observations.* Retrieved from `https://en.vedur.is/avalanches/articles/nr/2929`.

Jutzeler, M., Marsh, R., Carey, R. J., White, J. D. L., Talling, P. J., & Karlstrom, L. (2014). On the fate of pumice rafts formed during the 2012 havre submarine eruption. *Nature communications*, *5*.

Jutzeler, M., Marsh, R., van Sebille, E., Mittal, T., Carey, R. J., Fauria, K. E., ... McPhie, J. (2020). Ongoing dispersal of the 7 august 2019 pumice raft from the tonga arc in the southwestern pacific ocean. *Geophysical Research Letters*, *47*(5), e1701121.

Kelley, D. (2017). Volcanology: Vulcan rule beneath the sea. *Nature Geoscience*.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, *47*(260), 583–621.

Mandeville, C., Carey, S., & Sigurdsson, H. (1996). Sedimentology of the krakatau 1883 submarine pyroclastic deposits. *Bulletin of Volcanology*, *57*, 512-529. doi: 10.1007/BF00304436

Manville, V., Segschneider, B., & White, J. (2002). Hydrodynamic behaviour of taupo 1800a pumice: Implications for the sedimentology of remobilized pyroclasts. *Sedimentology*, *49*, 955 - 976. doi: 10.1046/j.1365-3091.2002.00485.x

Matsumoto, H., Zampolli, M., Haralabus, G., Stanley, J., Mattila, J., & Özel, N. M. (2019). Interpretation of detections of volcanic activity at ioto island obtained from in situ seismometers and remote hydrophones of the international monitoring system. *Scientific reports*, *9*(1), 1–11.

Mittal, T., & Delbridge, B. (2019). Detection of the 2012 havre submarine eruption plume using argo floats and its implications for ocean dynamics. *Earth and Planetary Science Letters*, *511*, 105–116.

Mouginis-Mark, P. J., & Zimbelman, J. R. (2020). Rafted pumice: A new model for the formation of the medusae fossae formation, mars. *Icarus*, *343*, 113684.

O'Malley, R. T., Behrenfeld, M. J., Westberry, T. K., Milligan, A. J., Reese, D. C., & Halsey, K. H. (2014). Improbability mapping: a metric for satellite-detection of submarine volcanic eruptions. *Remote sensing of environment*, *140*, 596–603.

Osborne, N. M., Enright, N. J., & Parnell, K. E. (1991). The age and stratigraphic significance of sea-rafted loisels pumice in northern new zealand. *Journal of the Royal Society of New Zealand*, *21*(4), 357-371. Retrieved from `https://doi.org/10.1080/03036758.1991.10420833` doi: 10.1080/03036758.1991.10420833

Park, C., & Schmincke, H.-U. (2020). Multistage damming of the rhine river by tephra fallout during the 12,900 bp plinian laacher see eruption (germany). syn-eruptive rhine damming i. *Journal of Volcanology and Geothermal Research*, *389*, 106688. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0377027319300502` doi: https://doi.org/10.1016/j.jvolgeores.2019.106688

Planet Team. (2018–2021). *Planet application program interface: In space for life on earth.* Retrieved from `https://api.planet.com`

Poland, M. P., Lopez, T., Wright, R., & Pavolonis, M. J. (2020). Forecasting, detecting, and tracking volcanic eruptions from space. *Remote Sensing in Earth Systems Sciences*, *3*(1), 55–94.

Pullar, W., Kohn, B., & Cox, J. (1977). Air-fall kaharoa ash and taupo pumice, and sea-rafted loisels pumice, taupo pumice, and leigh pumice in northern and eastern parts of the north island, new zealand. *New Zealand journal of geology and geophysics*, *20*(4), 697–717.

Qi, L., Hu, C., Mikelsons, K., Wang, M., Lance, V., Sun, S., ... Van der Zande, D. (2020). In search of floating algae and other organisms in global oceans and lakes. *Remote Sensing of Environment*, *239*, 111659.

Risso, C., Scasso, R. A., & Aparicio, A. (2002). Presence of large pumice blocks on tierra del fuego and south shetland islands shorelines, from 1962 south sandwich islands eruption. *Marine Geology*, *186*(3-4), 413–422.

Rubin, K. H., Soule, S. A., Chadwick Jr, W. W., Fornari, D. J., Clague, D. A., Embley, R. W., ... Dziak, R. P. (2012). Volcanic eruptions in the deep sea. *Oceanography*, *25*(1), 142–157.

Sakuno, Y. (2021). Trial of chemical composition estimation related to submarine volcano activity using discolored seawater color data obtained from gcom-c sgli. a case study of nishinoshima island, japan, in 2020. *Water*, *13*(8), 1100.

Santana-Casiano, J. M., González-Dávila, M., Fraile-Nuez, E., De Armas, D., González, A. G., Domínguez-Yanes, J. F., & Escánez, J. (2013). The natural ocean acidification and fertilization event caused by the submarine eruption of el hierro. *Scientific reports*, *3*.

Scholz, F. W., & Stephens, M. A. (1987). K-sample anderson–darling tests. *Journal of the American Statistical Association*, *82*(399), 918–924.

Shane, P., Froggatt, P., Smith, I., & Gregory, M. (1998). Multiple sources for sea-rafted loisels pumice, new zealand. *Quaternary Research*, *49*(3), 271–279.

Tepp, G., Chadwick Jr, W. W., Haney, M. M., Lyons, J. J., Dziak, R. P., Merle, S. G., ... Young III, C. W. (2019). Hydroacoustic, seismic, and bathymetric observations of the 2014 submarine eruption at ahyi seamount, mariana arc. *Geochemistry, Geophysics, Geosystems*, *20*(7), 3608–3627.

Tepp, G., & Dziak, R. P. (2021). The seismo-acoustics of submarine volcanic eruptions. *Journal of Geophysical Research: Solid Earth*, *126*(4), e2020JB020912.

Tilstone, G. H., Miller, P. I., Brewin, R. J., & Priede, I. G. (2014). Enhancement of primary production in the north atlantic outside of the spring bloom, identified by remote sensing of ocean colour and temperature. *Remote sensing of environment*, *146*, 77–86.

White, J. D. L., Schipper, C. I., & Kano, K. (2015). Chapter 31 - submarine explosive eruptions. In H. Sigurdsson (Ed.), *The encyclopedia of volcanoes (second edition)* (Second Edition ed., p. 553 - 569). Amsterdam: Academic Press. doi: http://doi.org/10.1016/B978-0-12-385938-9.00031-6

White, S. M., Crisp, J. A., & Spera, F. J. (2006). Long-term volumetric eruption rates and magma budgets. *Geochemistry, Geophysics, Geosystems*, *7*(3).

Whiteside, A., Dupouy, C., Singh, A., Frouin, R., Menkes, C., & Lefèvre, J. (2021). Automatic detection of optical signatures within and around floating tonga-fiji pumice rafts using modis, viirs, and olci satellite sensors. *Remote Sensing*, *13*(3), 501.

Wilcock, W. S., Tolstoy, M., Waldhauser, F., Garcia, C., Tan, Y. J., Bohnenstiehl, D. R., ... Mann, M. E. (2016). Seismic constraints on caldera dynamics from the 2015 axial seamount eruption. *Science*, *354*(6318), 1395–1399.

Wright, R., Flynn, L. P., Garbeil, H., Harris, A. J., & Pilger, E. (2004). Modvolc: near-real-time thermal monitoring of global volcanism. *Journal of Volcanology and Geothermal Research*, *135*(1-2), 29–49.