Volcanoes show signs of large-scale thermal unrest for years prior to eruption

Társilo Girona[1,2], Vincent Realmuto[1], and Paul Lundgren[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA.

[2]Geophysical Institute, University of Alaska Fairbanks, Alaska 99775, USA (current affiliation).

Identifying the observables that warn of volcanic eruptions is a major challenge in natural hazard management. An important, but under-investigated, observable is the release of heat through volcano surfaces, which represents a major energy source at quiescent volcanism. However, it remains unclear whether surface heat emissions respond to pre-eruptive processes and vary before eruption. Here we show that the last magmatic and phreatic eruptions of five different volcanoes were preceded by subtle but significant long-term (~years), large-scale (10s of km^2), increases of their radiant heat flux (up to ~1 °C of median radiant temperature). This pre-eruptive thermal unrest is found through a new statistical analysis of satellite-based long-wavelength (10.780–11.280 μm) infrared data, and is attributed to the enhancement of underground hydrothermal activity. Large-scale thermal unrest is detected even before eruptions that were impossible to anticipate through other geophysical/geochemical methods (e.g., the 2014 phreatic eruption of Ontake, Japan; the 2015 magmatic eruption of Calbuco, Chile), thus opening new horizons to better constrain the thermal budget of volcanoes and improve eruption forecasts.

Volcanoes are major emitters of the Earth's internal heat. A large amount of heat is released diffusively through the soil, which represents a major energy source during quiescence[1]. For example, *Chiodini et al.*[2] found that the surface heat emissions at the Solfatara crater dominate the energy budget of Campi Flegrei caldera (Italy), and that it is one order of magnitude greater than the elastic energy released during recent seismic and deformation episodes. *Mannini et al.*[3] also reported that surface heat emissions account for >90% of the total heat supplied to the Vulcano's Fossa fumarole field (Italy). Heat emissions, however, are not constrained to fumarolic fields, but they extend over large active volcanic areas, as revealed by geothermal gradient determinations[4]. For example, *Corrado et al.*[5] found that the surface heat flux in some areas of Campi Flegrei (~0.149 W/m^2) is about twice above regional values (~0.085 W/m^2); and *Tanaka et al.*[6] reported that the heat flux in southern Kyushu (Japan) increases significantly with decreasing distance to volcanoes (with values of up to ~0.5 W/m^2). These studies stimulate the following question: Do the subsurface processes driving volcanic eruptions yield significant large-scale thermal footprints at the surface?

Detection of radiant heat emissions and case studies

This work addresses the aforementioned question by analyzing 16.5-year records (between 04-July-2002 and 31-January-2019) of the thermal infrared radiance measured by the Moderate Resolution Imaging Spectroradiometers (MODIS) aboard NASA's Terra and Aqua satellites. Specifically, we examine day- and night-time scenes from MODIS Band 31 (10.780-11.280 μm), which captures surface radiance variations with minimal effects from atmospheric

absorption and emissions[7]. MODIS is well suited for our study because it provides global coverage with high temporal resolution (~2 observations per day at tropical latitudes); wide ground swaths with simultaneous information of large areas (~2,330 km at tropical latitudes); and spatial resolution of ~1 km^2 at nadir. Our MODIS analysis is based on the median anomaly ($\overline{\delta T}$), a new statistical metric that tracks the radiant heat emitted by volcanic flanks and that minimizes the effects of "outliers" in the data records due to cloud coverage, geolocation errors, pixel mosaicking, or overlap between scan lines[7]. In particular, $\overline{\delta T}$ captures the long-term (>1 year), large-scale (10s of km^2), variations of the minimum spectral radiance (in units of radiant –or brightness- temperature for simplicity) of a volcanic edifice with respect to its surroundings (Supplementary Materials).

This study explores the temporal evolution of $\overline{\delta T}$ for five volcanoes that are representative of many volcanoes around the world, cover a broad range of behaviors and characteristics, and have erupted/exploded over the past two decades (Ontake, Japan; Ruapehu, New Zealand; Calbuco, Chile; Redoubt, Alaska; and Pico do Fogo, Cape Verde). In particular, these volcanic systems cover different: (a) Eruption type (phreatic/magmatic) and magnitude (Volcanic Explosivity Index or VEI=0-4); (b) Evolution; (c) Degree of hydrothermal activity at the surface; (d) Levels of pre-eruptive unrest and post-eruptive behavior; (e) Ground covers (permanent glaciers or snow, seasonal snow, crater lake, arid areas, forest and rainforest); (f) Latitude (from $\sim 8°$ to $\sim 60°$); and (g) Tectonic setting. The only requirement imposed in the choice of the volcanoes is that our statistical algorithm must be applicable. This implies that the volcanoes: (i) Must have erupted only once or twice over the last two decades because our algorithm does not capture

short-term (<1 year) thermal processes. (ii) Must not be located in very small islands because our algorithm requires a sufficient number of pixels for the statistical analysis to be robust. (iii) Must have significant topography, such that volcano summits are colder than the surroundings due to an altitude effect; this is used by our algorithm to automatically identify cloud-free scenes.

Pre-eruptive thermal unrest

Our analysis reveals that the eruptions/explosions of our test cases occurred around the highest values of the median anomaly $\overline{\delta T}$, and that they were preceded by a subtle but significant long-term (>1 year) increase of radiant temperature ($\sim 0.1-1.0$ ℃) whose uncertainty is smaller (<40%) for large events (Fig. 1; Table 1). In addition, $\overline{\delta T}$ decreases right after small eruptions (VEI<3), or a few months after large (VEI$\geq$3) eruptions, probably due to the materials deposited around the vent. Below, we describe the results obtained for each volcano.

Ontake featured two unexpected events, a small (VEI=0) gas explosion in March 2007 and a large (VEI=3) phreatic eruption in September 2014 (which killed around 60 hikers[8] and became the worst volcanic disaster in Japan since 1926) (Fig. 1a). The median anomaly $\overline{\delta T}$ increased by ~0.20 ℃ between January 2004 and March 2007, when the small gas explosion occurred. After this explosion, $\overline{\delta T}$ decreased gradually at an average rate of ~ 0.14 ℃/*year*, until reaching the minimum value of the time series in mid-2012. Then, $\overline{\delta T}$ increased by 0.72±0.21 ℃ for about two years, reaching $\overline{\delta T}$ = 0.85±0.15 ℃ by the time of the large 2014 phreatic eruption[8]. The ~2-year precursory thermal unrest of the 2014 eruption is apparently linked neither to long-term

seismic and geodetic observations around the volcano[8] nor to helium anomalies[9]. No other well-defined long-term warning sign has been detected so far for the 2014 eruption.

Ruapehu featured two unheralded events, a small (VEI=1) gas explosion in October 2006 and a larger phreatomagmatic eruption (VEI=1) in September 2007 (in which a climber suffered serious injuries[10]) (Fig. 1b). The median anomaly $\overline{\delta T}$ increased by 0.65±0.23 ℃ during, at least, the 3-to-4 years preceding the 2006 explosion and the 2007 eruption. This pre-eruptive thermal unrest coincides with the enhancement of seismic attenuation[11], and partially with the tidal modulation of the shallow seismicity[12]. Since the 2007 eruption, the thermal anomaly has remained at high levels ($\overline{\delta T}$~$0.6-0.8$ ℃) compared to 2004 ($\overline{\delta T}$~$0.2-0.3$ ℃), concomitant with several unrest episodes. It is worth highlighting a decrease of $\overline{\delta T}$ from the end of 2011 to mid-2015, coinciding in part with a sustained period of low lake temperatures (from mid-2012 to mid-2013); and a small upturn of ~0.15 ℃ between 2014-2017, coinciding with strong tremor levels, anomalous earthquakes, and high degassing rates and lake temperatures[12].

Calbuco featured a large but unheralded magmatic eruption[13] (VEI=4) in April-May 2015 (Fig. 1c). The median anomaly rose slowly by 0.32±0.13 ℃ between 2008 and 2012; then, it remained roughly constant until the onset of the 2015 magmatic eruption. After the eruption, $\overline{\delta T}$ increased at much larger rates until reaching $\overline{\delta T}$~0.52 ℃ by the end of 2015; and since 2017, $\overline{\delta T}$ has been decreasing quickly. The pre-eruptive increase of $\overline{\delta T}$ suggests that the quick (<4 days) ascent of magma[13,14] expelled during the 2015 event was indeed preceded by a much longer-term build-up process (e.g., slow accumulation of magma or gases in the shallow crust)

that initiated about seven years earlier leaving a detectable thermal signature at the surface. A subtle long-term slow build-up to eruption may explain why no deformation was detected prior to the 2015 eruption[13].

Redoubt had a prominent (VEI=3) magmatic eruption[15] in March 2009 (Fig. 1d). The median anomaly increased by 0.47 ± 0.17 ℃ from mid-2006 to the 2009 magmatic eruption (~0.17 ℃/*year* on average); about one year after the event, $\overline{\delta T}$ started to decrease at accelerated rates for about five years. Since 2014, $\overline{\delta T}$ has remained at low levels ($\lesssim$0.3 ℃), slightly above the values found years before the 2009 eruption. Interestingly, $\overline{\delta T}$ began to increase ~1 year earlier than other precursory signals, including glacier melting, sulfur odors, increased gas emissions, deep seismicity, and deformation[15].

Pico do Fogo is a hotspot volcano that erupted magma (VEI=3) between November 2014 and February 2015 (Fig. 1e), and has been found to release diffuse magmatic CO_2 over extensive areas[16]. This eruption, which displaced more than 1,000 people and destroyed hectares of agricultural land[17], was preceded by an increase of the median anomaly of 0.82 ± 0.23 ℃ since mid-2010 (~0.20 ℃/*year* on average). $\overline{\delta T}$ increased quickly after the eruption and decreased a few months afterwards; in addition, thermal unrest initiated much earlier than other warning signs were detected (e.g., increasing diffuse CO_2 was detected only a few months before the event[17]). Our analysis also reveals a less prominent but significant thermal unrest between 2004 and 2008. However, there is limited information between the 1995 and the 2014-2015 eruptions to assess whether this thermal unrest was concomitant with other geophysical/geochemical warning signs or whether it culminated in some type of undetected activity at the surface.

Enhancement of underground hydrothermal activity

Below, we discuss four possible explanations for the pre-eruptive long-term (> 1 year) variations of the median anomaly ($\overline{\delta T}$) detected at the target volcanoes. First, the variations of the median anomaly are spurious signals produced by low signal-to-noise ratios. This is not feasible because the probability of obtaining the observed amplitudes of $\overline{\delta T}$ by chance is 2.1% at most (for Calbuco), and is $\leq 0.01\%$ for Ontake and Redoubt (Fig. 2f). Hence, our results represent actual variations of the thermal radiant flux beyond the 95% confidence level. Second, the variations of $\overline{\delta T}$ are dominated by a combination of atmospheric, MODIS sensor stability, or fire processes. Although these processes may affect the median anomaly, the systematic increase of $\overline{\delta T}$ prior to eruption suggests that it is dominated by subsurface volcanic processes. Third, the pre-eruptive variations of $\overline{\delta T}$ reflect the emergence of small-scale (< 1 km^2) volcano-related hotspots (i.e., lava domes or fumaroles). This is not feasible because gradual, long-term, $\overline{\delta T}$ variations reflect gradual shifts of the radiant temperature distribution of the ground. The emergence of small-scale volcano-related hotspots can produce step variations of the median of the distribution but not gradual long-term changes, unless they emerge gradually over the large areas explored, which is not realistic. Fourth, $\overline{\delta T}$ increases before eruptions because subsurface pre-eruptive processes change the radiant characteristics of the surface (i.e., skin temperature and/or emissivity) over extensive areas (~10s of km^2) around the eruptive vents. The latest is the most realistic explanation for our results.

Pre-eruptive variations of the radiant characteristics of volcano surfaces probably reflect

subsurface hydrothermal activity (Fig. 2), which in turn may be enhanced by diffuse degassing (i.e., the permeable transport of hot magmatic gases through the crust). Diffuse degassing operates over extensive areas, increases with volcanic activity, is known to control soil heating in fumarolic fields, and is controlled by magma ascent, magma reservoir dynamics, and crust permeability[1,16,18-20]. For example, diffuse degassing could facilitate the upflow of underground waters, the elevation of the water table, and the circulation of hydrothermal cells that modify the near-surface geothermal gradient and slightly warm the soil and snow cover over extensive areas of the volcanic flanks[21,22]. This possibility is supported by the fact that long-term (~years) surface temperature is highly coupled with groundwater temperature at several tens of meters depth[23]. Diffuse degassing could also trigger boiling of underground water, thus leading to the transport of massive amounts of steam towards the surface, H_2O condensation beneath the ground, and release of latent heat that is then transported diffusively to the uppermost part of the soil[1,2,19,24]. Alternatively, an increasing supply of fluids to the shallow subsurface may produce long-term subtle changes in soil moisture and vegetation, thus gradually increasing the ground emissivity with time over large areas[25-27]. Either of these processes, or a combination of them, is likely responsible for the pre-eruptive increases in radiant energy recorded by the median anomaly $\overline{\delta T}$.

This work shows that volcanoes can experiment thermal unrest for several years before eruption. This thermal unrest is dominated by a large-scale phenomenon operating over extensive areas of volcanic edifices, can be an early indicator of volcanic reactivation, can increase prior to different types of eruption, and can be tracked through a statistical analysis of little processed (i.e., radiance or radiant temperature) satellite-based remote sensing data with high temporal

resolution. Our findings can open new horizons to better constrain magma-hydrothermal interaction processes, to explore the thermal budget of volcanoes, and to improve the forecast of eruptions that are very difficult to anticipate through other geophysical/geochemical methods.

METHODS

1) Data retrieval

Data from 04-July-2002 to 31-January-2019 are retrieved from the MODIS instruments aboard the Terra and Aqua satellites, launched in December 1999 and May 2002, respectively. MODIS products are a major component of NASA's Earth Observing System, and their strengths include global coverage, high sampling rate (from ~1 to ~4 swaths per day per satellite, depending on latitude), high radiometric resolution, suitable dynamic range, and accurate calibration[7]. In particular, we investigate the thermal evolution of volcanic surfaces by using two products (https://earthdata.nasa.gov/): MODIS Terra/Aqua Calibrated Radiances 5-Min Level-1B Swath 1km V006 (MOD021KM/MYD021KM) and MODIS Terra/Aqua Geolocation Fields 5-Min Level-1A Swath 1km V006 (MOD03/MYD03). The Level-1B Radiance product provides accurate values of radiance, which have little or no long-term drift[28], are radiometrically calibrated, and are corrected for instrumental effects; the Level-1A Geolocation product provides the geographical coordinates of each pixel of the Level-1B scenes. Altogether, we analyze >300,000 MODIS files (52,554 for Ontake; 57,078 for Ruapehu; 60,676 for Calbuco; 93,792 for Redoubt; and 44,008 for Pico do Fogo), which correspond to >30 TB of memory. MODIS acquires radiance at spatial resolution of $\sim 1\, km^2$ at nadir in 16 thermal bands of the electromagnetic spectrum, but this study focuses only on band 31 (10.780-11.280 μm) because is more sensitive to surface temperature variations[7]. Note, however, that we use the Level-1B Radiance product, instead of Land Surface Temperature (LST) products[29], because we want to

explore the heat flux radiated by volcanic surfaces using data with little previous processing, and thus minimizing the potential uncertainties as much as possible.

2) Data Analysis

We implement a new statistical strategy to capture the long-term (~years) radiant (or brightness) temperature evolution of the coldest areas (~20 km^2 on average) of the volcanic flanks. This approach contrasts with previous algorithms aimed at detecting the emergence of hotspots[30-36] (i.e., pixels that are hotter than the surrounding pixels) associated with magma exposure at the surface or fumarolic activity. Our algorithm consists of six main steps, which are detailed below (Supplementary Fig. 1).

2.1) Choice of exploration areas

We choose latitude-longitude quadrangles (0.30° latitude by 0.48° longitude) centered at the geographical coordinates of the volcanoes under study (as provided by the Smithsonian Institution's Global Volcanism Program database; https://volcano.si.edu/). These quadrangles cover the volcanic edifices and their surroundings, and represent surface areas A_t on the order of ~$900-1,700\ km^2$ and a number of pixels at nadir in the range $N \approx 900-1,700$, depending on the latitude of the volcanoes (Supplementary Fig. 1a, 1b).

2.2) Automatic discarding of useless scenes

We observe that, in cloud-free scenes, the pixels with lowest spectral radiance in the target area A_t lie at the highest levels of the volcanic edifices because the ground is colder (and thus emits

less radiance) due to altitude. This fact is used to automatically identify cloud-free scenes as follows (Supplementary Fig. 1c-1e): We calculate different statistical estimators in the target areas A_t, including the minimum spectral radiance, defined as the median spectral radiance of the M pixels with lowest spectral radiance ($L_{c,M}$); and the median spectral radiance of the K pixels with largest spectral radiance ($L_{h,K}$). The median is the statistical estimator chosen because it minimizes the effect of outliers (e.g., due to cloud coverage, pixel mosaicking, geolocation errors, or overlap of scan lines[7]). (ii) For each of the MODIS scenes, we plot the location of the pixel with spectral radiance $L_{c,11}$. (iii) We create an auxiliary area A_{aux} that covers, for each volcano, the region with the highest concentration of pixels with median spectral radiance $L_{c,11}$ (Supplementary Fig. 1c). (iv) We discard scenes whose pixel with median spectral radiance $L_{c,11}$ is outside A_{aux} and/or whose pixel with median spectral radiance $L_{h,N-101}$ is inside the subarea A_{aux} (where $N-101$ represents the total number of pixels within the target area A_t except the 101 pixels with lowest radiance).

This approach yields a percentage of scenes discarded between ~62% for Ruapehu and ~84% for Calbuco. In the worst-case scenario (Calbuco), this implies 4,956 useful scenes for the 16.5-year period analyzed and thus 300 useful scenes per year on average. It is worth noting that our algorithm allows us to discard scenes with high levels of cloudiness, but also scenes captured during post-eruptive periods (if hot magmatic products cover large areas of the volcanic flanks) and during days with temperature inversion (i.e., if the mountain is warmer than the surroundings due to atmospheric conditions). Our algorithm therefore standardizes the scenes explored, and allows us to minimize the level of noise by identifying useless scenes and discarding them. It is

worth noting that our approach is robust because results are not highly sensitive to the choice of the auxiliary area A_{aux}. In fact, if enlarging or dwarfing the subareas A_{aux} by 10%, we obtain essentially the same results; and, even if cloudy scenes are not discarded, low-pass filtering (see subsection 2.5) allows retrieving pre-eruptive thermal unrest in three of the five volcanoes (Ontake, Ruapehu, Redoubt) (Supplementary Fig. 2).

2.3) Thermal contrast between coldest areas and regional background

For the scenes that are not discarded, we calculate the thermal contrast between a volcano and its surroundings using the parameters $L_{c,M}$ and $L_{h,K}$. The parameter $L_{c,M}$ is an estimator of the minimum spectral radiance (calculated as the median spectral radiance of the M coldest pixels) of the volcanic flanks, whereas $L_{h,K}$ is an estimator of the median spectral radiance of the regional background. In particular, we use $M = 11$ pixels ($\sim 18-22\ km^2$, on average, for the target volcanoes) and $K = N - 101$ pixels (with N the total number of pixels within the target area A_t) (Supplementary Table 1), but our overall conclusions remain unchanged for a broad range of values of M ($= 21,\ 31$) and K ($= N - 51,\ N - 201$). Once the statistical estimators $L_{c,M}$ and $L_{h,K}$ are calculated, we convert them to brightness temperature for simplicity ($T_{c,M}$ and $T_{h,K}$, respectively) using the Planck's function:

$$T_x = \frac{C_2}{\lambda\, ln\left(1+\frac{C_1}{\lambda^5 L_x}\right)} \qquad , \qquad (M1)$$

where L_x represents $L_{c,M}$ or $L_{h,K}$; T_x represents the median radiant temperatures $T_{c,M}$ or $T_{h,K}$; $C_1 = 1.19 \cdot 10^{-16}\ m^2\ W$; $C_2 = 1.44 \cdot 10^{-2}\ m\ K$; and λ is the central wavelength of band 31

(11.03 μm). Then, we calculate the difference $\Delta T_{M,K} = T_{c,M} - T_{h,K}$ (Supplementary Fig. 1f). Calculating the contrast of radiant temperature between a volcano and its surroundings pursues three following goals: (i) To highlight any variation of radiant temperature occurring in the coldest parts of the volcanic flanks and not occurring in the area surrounding the volcano. (ii) To minimize any local/regional atmospheric effect, as well as the possible artifacts of jointly combining daytime/nighttime scenes. (iii) To minimize the possible artifacts associated with the use of different sensors (MODIS/Terra and MODIS/Aqua).

2.4) Daily median brightness temperature difference

We compute the daily median brightness temperature difference ($\overline{\Delta T}_{M,K}(t)$, where t is time), i.e., the daily median of the statistical estimator $\Delta T_{M,K}$ (Supplementary Fig. 1g). This approach pursues the following goals: (i) To produce a regular and continuous sampling rate (i.e., 1 sample/day), which facilitates signal processing; and (ii) to minimize possible outliers associated with daytime/nighttime scenes. We use daytime and nighttime scenes to increase the sample population, and thus yield a more reliable statistical analysis (although the patterns obtained are essentially the same when using only daytime or nighttime scenes; see Supplementary Fig. 3). The resulting time series ($\overline{\Delta T}_{M,K}(t)$) contains a seasonal component, noise, and gaps (produced by the discarding of useless scenes; section 2.2) that are filled through linear interpolation. In the worst-case scenario (Calbuco), we find 55% gaps in $\overline{\Delta T}_{M,K}(t)$, with a largest gap of 20 days, and 165 days per year (on average) with useful data (Supplementary Table 1). Although the large

number of daily gaps hinders the detection of short-term (~days) anomalies through this method, the number of valid scenes per year is suitable to detect long-term (~years) trends.

2.5) Low-pass filtering

After interpolation, the time series $\overline{\Delta T}_{M,K}(t)$ contains a well-defined seasonal component and noise; our aim is to bring out any long-term (~years) trend that may be contained in $\overline{\Delta T}_{M,K}(t)$ (Supplementary Fig. 1h). To this outcome, we have designed an efficient low-pass filtering technique through >100,000 Monte Carlo experiments with synthetic time series $\overline{\Delta T}_{SYN}(t)$. The details of our numerical experiments are provided below:

1. We generate synthetic time series ($\overline{\Delta T}_{SYN}(t)$) combining a long-term trend ($\Delta T_{SYN}^{trend}(t)$), a seasonal component ($A_s sin\left(\frac{2\pi t}{365}\right)$), and a zero-mean Gaussian noise ($R(0,\sigma_d)$) with standard deviation σ_d:

$$\overline{\Delta T}_{SYN}(t) = \Delta T_{SYN}^{trend}(t) + A_s sin\left(\frac{2\pi t}{365}\right) + R(0,\sigma_d) \qquad . \qquad (M2)$$

We use $A_s = \sigma_d = 100$ and $\Delta T_{SYN}^{trend}(t) = \alpha_{imp} cos\left(\frac{2\pi(t-\tau_{imp})}{T_{imp}}\right)$, where α_{imp}, T_{imp}, and τ_{imp} are an imposed amplitude, period, and time lag, respectively (Supplementary Fig. 4a-4d).

2. For a given combination of α_{imp}, T_{imp}, and τ_{imp}, we filter the resulting time series $\overline{\Delta T}_{SYN}(t)$ through 480 different low-pass filtering methods (i.e., 24 filtering techniques that are applied

iteratively up to $k = 20$ times; Supplementary Table 2) aiming to retrieve the imposed trend $\Delta T_{SYN}^{trend}(t)$. Iterative filtering methods are commonly used to improve the performance of noise reduction techniques[37]. The filtering methods explored include a combination of trailing moving mean, trailing moving median, MODWT (maximal overlap discrete wavelet transform), DWT (discrete wavelet transform), and Savitzky-Golay filters. The goodness of every denoising technique is assessed by computing the sinusoidal curve that best fits the filtered time series (with the non-linear least squares method), and then by calculating the coefficient of determination R^2, amplitude α_{ret}, period T_{ret}, and time lag τ_{ret} of the best fit curve. The values of R^2, α_{ret}, T_{ret}, and τ_{ret} are retrieved for 100 computer-generated synthetic time series $\overline{\Delta T}_{SYN}(t)$ with the same imposed values of α_{imp}, T_{imp}, and τ_{imp}, which are then used to calculate the mean coefficient of determination ($\overline{R^2}$), mean retrieved amplitude ($\overline{\alpha}_{ret}$), mean retrieved period ($\overline{T}_{ret}$), mean retrieved time lag ($\overline{\tau}_{ret}$), and their uncertainties (standard deviation). Finally, the values of $\overline{R^2}$, $\overline{\alpha}_{ret}$, $\overline{T}_{ret}$, $\overline{\tau}_{ret}$, and their uncertainties are calculated for different combinations of imposed amplitude, period, and time lag in the range $\alpha_{imp} = 2 - 90$, $T_{imp} = 5 - 20$ years, $\tau_{imp} = 0$. Under this configuration, the signal-to-noise ratio (SNR; defined here as the variance of $\Delta T_{SYN}^{trend}(t)$ over the variance of $\overline{\Delta T}_{SYN}(t)$) varies in the range $SNR \approx 10^{-4} - 0.2$.

3. We assess which of the 480 low-pass filtering methods are the most appropriate to retrieve the imposed long-term (~years) trends (i.e., $\Delta T_{SYN}^{trend}(t)$) from the time series $\overline{\Delta T}_{SYN}(t)$. Our numerical experiments reveal that one of the most suitable methods is Technique 11 with $k = 10$

iterations (Supplementary Table 2), i.e., the combination of 10 MODWT filters (at level 10 and using symlets-8 wavelets, Donoho and Johnstone's universal soft thresholding, and level-dependent estimation of the noise) + 1 trailing moving median filter (of order 2 years) (Supplementary Fig. 5–8). For a vast combination of α_{imp} and T_{imp}, this method provides the highest values of the coefficient of determination, produces mean retrieved amplitudes ($\overline{\alpha}_{ret}$) and periods ($\overline{T}_{ret}$) that are compatible with the imposed values, and generates a mean time lag ($\overline{\tau}_{ret}$) on the order of ~1 year in the filtered signal that can be easily corrected. The filtering process, however, has several limitations that need to be taken into account: (a) If the signal-to-noise ratio is too low, our filtering process is not able to accurately retrieve hidden long-term trends from the $\overline{\Delta T_{SYN}}(t)$ time series. For example, for $T_{imp} = 15$ years, it is met that $\overline{R^2} \gtrsim 0.85$, $\overline{T}_{ret} \approx T_{imp}$, $\overline{\alpha}_{ret} \approx \alpha_{imp}$, and $\overline{\tau}_{ret} \approx 1\ year$ when $SNR \gtrsim 0.001$ only (Supplementary Fig. 9a-9d). (b) For decreasing values of the imposed periods T_{imp}, our method yields $\overline{\alpha}_{ret} < \alpha_{imp}$ and $\overline{\tau}_{ret} < 1\ year$. For example, for $T_{imp} = 7$ years, it is met that $\overline{R^2} \gtrsim 0.85$ and $\overline{T}_{ret} \approx T_{imp}$ when $SNR \gtrsim 0.004$, although $\overline{\alpha}_{ret} \approx 0.8\ \alpha_{imp}$ and $\overline{\tau}_{ret} \sim 320$ days (Supplementary Fig. 9e-9h). (c) For imposed periods $T_{imp} \lesssim 6-7$ years, our filtering method cannot accurately retrieve long-term trends, independently of the signal-to-noise ratio (Supplementary Fig. 9i-9l). (d) When our filtering process is applied to a synthetic signal without an imposed trend (or with very low signal-to-noise ratio), a spurious trend can be retrieved (Supplementary Fig. 4e-4f). In subsection 2.6, we present a methodology to discern between actual hidden long-term trends and spurious signals associated with high levels of noise.

4. We apply the efficient low-pass filtering technique (Technique 11 with $k = 10$ iterations) to the daily median brightness temperature difference ($\overline{\Delta T}_{M,K}(t)$) obtained for the five target volcanoes. This yields the so-called median anomaly or $\overline{\delta T}$, which is then 1-year time-shifted to account for the delay produced by the filtering process (Supplementary Fig. 1h). Note that $\overline{\delta T} < 0$ (because $T_{c,M} < T_{h,K}$), has units of temperature, captures the long-term (~years) variations of median radiant temperature of the coldest parts of volcanoes with respect to the regional background, and its amplitude of variation must be considered a minimum estimate because part of it may be missed during the filtering process (Supplementary Fig. 9g, 9k). It is also worth highlighting that $\overline{\delta T}$ is calculated using radiant temperatures as measurement variable (instead of real temperatures) because it is sufficient to detect radiant heat flux variations of the surface. In addition, calculating the median anomaly in terms of real temperatures requires accurate estimations of the pixel emissivity as well; this could lead to significant uncertainties in our analysis that may be critical to detect the subtle variations observed.

2.6) Uncertainty analysis

Allocating uncertainties is crucial to assess the significance of the variations of the median anomaly time series ($\overline{\delta T}(t)$). The major source of uncertainty in $\overline{\delta T}(t)$, associated with the filtering process, is assessed through Monte Carlo experiments (shaded bands of Fig. 1a-2e):

1. We generate synthetic time series with a bootstrapping method (commonly used in seismic data processing[11]) consisting of resampling the daily median brightness temperature difference

time series $\overline{\Delta T}_{M,K}(t)$. In particular, each element *j* of the synthetic time series is selected randomly from the elements *j*-1, *j*, *j*+1 of $\overline{\Delta T}_{M,K}(t)$ (the values of the first and last elements of the synthetic time series are chosen to be equal to the values of the second and penultimate values, respectively). This technique yields synthetic time series with the same mean and standard deviation as $\overline{\Delta T}_{M,K}(t)$.

2. We apply the best denoising technique designed in section 2.5 to the aforementioned synthetic time series. The result yields $\overline{\delta T}^{*}(t)$, which is then compared with $\overline{\delta T}(t)$ by calculating the difference between each data point of the time series.

3. We repeat 1,000 times the steps 1-to-2. This provides 1,000 values of the difference $\overline{\delta T}(t) - \overline{\delta T}^{*}(t)$ for each data point *j*, which is then used to determine the y-axis uncertainty of $\overline{\delta T}(t)$ at the 95% confidence interval (Supplementary Fig. 1g). Note that this approach also captures the uncertainty associated with the time lag produced by the filtering process.

4. For simplicity, we rescale $\overline{\delta T}$ (with its uncertainty), so it is always ≥ 0.

On the other hand, a spurious trend may appear in the filtered time series as a consequence of the high levels of noise (see section 2.5 and Supplementary Fig. 4e-4f). The probability that the observed trends are spurious is assessed in terms of the maximum amplitude of variation of $\overline{\delta T}$ (Fig. 1f):

1. We design a synthetic time series composed of a random noise with the same probability density function as the daily median brightness temperature difference time series ($\overline{\Delta T}_{M,K}(t)$).

2. We apply the best denoising technique designed in section 2.5 to the aforementioned synthetic signal.

3. We repeat 10,000 times the steps 1-to-2. This is used to calculate the probability of obtaining by chance a spurious trend with the same amplitude or larger than the amplitude of the median anomaly time series $\overline{\delta T}$.

Data availability. Data (MODIS Terra/Aqua Calibrated Radiances 5-Min Level-1B Swath 1km V006 and MODIS Terra/Aqua Geolocation Fields 5-Min Level-1A Swath 1km V006) are available fully, openly, and without restrictions at https://earthdata.nasa.gov/.

Code availability. Two Matlab scripts for data processing and analysis are available (contact the corresponding author for further details).

REFERENCES:

1. Chiodini, G. *et al.* Carbon dioxide diffuse degassing and estimation of heat release from volcanic and hydrothermal systems. *J. Geophys. Res. Solid Earth* **110**, B08204 (2005).

2. Chiodini, G. *et al.* CO2 degassing and energy release at Solfatara volcano, Campi Flegrei, Italy. *J. Geophys. Res. Solid Earth* **106**, 16213–16221 (2001).

3. Mannini, S., Harris, A. J. L. & Jessop, D. E. Combining Ground - and ASTER - Based Thermal Measurements to Constrain Fumarole Field Heat Budgets : The Case of Vulcano Fossa 2000 – 2019. *Geophys. Res. Lett.* 1–10 (2019). doi:10.1029/2019GL084013.

4. Carlino, S., Somma, R., Troise, C., De Natale, G. The geothermal exploration of Campanian volcanoes: Historical review and future development, *Renew. Sust. Energ. Rev.* **16**, 1004-1030 (2012).

5. Corrado, G., De Lorenzo, S., Mongelli, F., Tramacere, A., & Zito, G. Surface heat flow density at the Phlegrean Fields caldera (southern Italy). *Geothermics* **27**, 469-484 (1998).

6. Tanaka, A., Yamano, M., Yano, Y., & Sasada, M. Geothermal gradient and heat flow data in and around Japan (I): Appraisal of heat flow from geothermal gradient data. *Earth Planets Space* **56**, 1191-1194 (2004).

7. Ahmad, S. . P. *et al.* MODIS radiances and reflectances for earth system science studies and environmental applications. *Proc. 18th Int. Conf. Interact. Inf. Process. Syst. Meteorol. Oceanogr. Hydrol.* 188–192 (2002).

8. Murase, M. *et al.* Preparatory process preceding the 2014 eruption of Mount Ontake volcano, Japan: insights from precise leveling measurements. *Earth, Planets Sp.* **68**, (2016).

9. Sano, Y., Kagoshima, T., Takahata, N., Nishio, Y. & Roulleau, E. Ten-year helium anomaly prior to the 2014 Mt Ontake eruption. *Sci. Rep.* **5**, 13069 (2015).

10. Kilgour, G. *et al.* The 25 September 2007 eruption of Mount Ruapehu, New Zealand: Directed ballistics, surtseyan jets, and ice-slurry lahars. *J. Volcanol. Geotherm. Res.* **191**, 1–14 (2010).

11. Caudron, C. *et al.* Change in seismic attenuation as a long-term precursor of gas-driven eruptions. *Geology* (2019). doi:10 .1130 /G46107.1.

12. Girona, T., Huber, C. & Caudron, C. Sensitivity to lunar cycles prior to the 2007 eruption of Ruapehu volcano. **8**, 1–9 (2018).

13. Delgado, F., Pritchard, M. E., Ebmeier, S., González, P. & Lara, L. Recent unrest (2002–2015) imaged by space geodesy at the highest risk Chilean volcanoes: Villarrica, Llaima, and Calbuco (Southern Andes). *J. Volcanol. Geotherm. Res.* **344**, 270–288 (2017).

14. Morgado, E. *et al.* Old magma and a new, intrusive trigger: using diffusion chronometry to understand the rapid-onset Calbuco eruption, April 2015 (Southern Chile). *Contrib. Min.* (2019).

15. Bull, K. *et al.* The 2009 eruption of Redoubt volcano, Alaska. in (ed. Schaefer, J. R.)

(State of Alaska, Department of Natural Resources, 2012).

16. Dionis, S. M. *et al.* Diffuse CO_2 degassing and volcanic activity at Cape Verde Islands, West Africa. *Earth, Planets Sp.* **67**, (2015).

17. Program, G. V. *Report on Fogo (Cape Verde) — November 2014*. (Smithsonian Institution, 2014). doi:https://doi.org/10.5479/si.GVP.BGVN201411-384010.

18. Schwandner, F. M., Seward, T. M., Gize, A. P., Hall, P. A. & Dietrich, V. J. Diffuse emission of organic trace gases from the flank and crater of a quiescent active volcano (Vulcano, Aeolian Islands, Italy). *J. Geophys. Res. Atmos.* **109**, n/a-n/a (2004).

19. Brombach, T., Hunziker, C., Chiodini, G., Cardellini, C. & Marini, L. Soil diffuse degassing and thermal energy fluxes from the southern Lakki plain, Nisyros (Greece). *Geophys. Res. Lett.* **28**, 69–72 (2001).

20. Hernández, P. A. *et al.* Diffuse emission of CO_2 from Miyakejima volcano, Japan. *Chem. Geol.* **177**, 175–185 (2001).

21. Bibby, H.M., Caldwell, T.G., Davey, F.J., & Webb, T.H. Gephysical evidence on the structure of the Taupo Volcanic Zone and its hydrothermal circulation. *J. Volcanol. Geotherm. Res.* **68**, 29-58 (1995).

22. Jasim, A., Hemmings, B., Mayer, K., & Scheu, B. Groundwater flow and volcanic unrest. *Advs in Volcanology*, 83-99 (2019).

23. Benz, S.A., Bayer, P., Blum, P. Global patterns of shallow groundwater temperatures. *Environ. Res. Lett.* **12**, 034005 (2017).

24. Frondini, F. *et al.* Diffuse CO_2 degassing at Vesuvio, Italy. *Bull. Volcanol.* **66**, 642–651 (2004).

25. Mira, M., Valor, E., Boluda, R., Coll, C. Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. *J. Geophys. Earth. Surf.* **112**(F04003) (2007).

26. Bogue, R.R. *et al.* Plant responses to volcanically elevated CO_2 in two Costa Rican forests. *Biogeosciences* **16**, 1343-1360 (2019).

27. Gerhards, M., Schlerf, M., Mallick, K., & Udelhoven, T. Challenges and future perspectives of multi-/hyperspectral thermal infrared remote sensing for crop water-stress detection: A review. *Remote Sens.* **11**, 1240 (2019).

28. Wenny, B. N., Xiong, X., Madhavan, S., Wu, A. & Li, Y. Long-term band-to-band calibration stability of MODIS thermal emissive bands. *Ocean Sens. Monit. V* **8724**, 872412 (2013).

29. Li, Z. L. *et al.* Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **131**, 14–37 (2013).

30. Wright, R., Flynn, L. P., Garbeil, H., Harris, A. J. L. & Pilger, E. MODVOLC: near-real-time thermal monitoring of global volcanism. **135**, 29–49 (2004).

31. Coppola, D., Laiolo, M., Cigolini, C., Donne, D. D. & Ripepe, M. Enhanced volcanic hot-spot detection using MODIS IR data: results from the MIROVA system. *Geol. Soc. London, Spec. Publ.* **426**, 181–205 (2015).

32. Reath, K. A., Ramsey, M. S., Dehn, J. & Webley, P. W. Predicting eruptions from precursory activity using remote sensing data hybridization. *J. Volcanol. Geotherm. Res.* **321**, 18–30 (2016).

33. Van Manen, S. M. & Dehn, J. Satellite remote sensing of thermal activity at Bezymianny and Kliuchevskoi from 1993 to 1998. **37**, 983–986 (2009).

34. Lacava, T. *et al.* Thermal Monitoring of Eyjafjöll Volcano Eruptions by Means of Infrared MODIS Data. **7**, 3393–3401 (2014).

35. Laiolo, M. *et al.* Evidences of volcanic unrest on high-temperature fumaroles by satellite thermal monitoring: The case of Santa Ana volcano, El Salvador. *J. Volcanol. Geotherm. Res.* **340**, 170–179 (2017).

36. Marchese, F. *et al.* Inferring phases of thermal unrest at Mt . Asama (Japan) from infrared satellite observations. *J. Volcanol. Geotherm. Res.* **237**–**238**, 10–18 (2012).

37. Billings, S. A. & Lee, K. L. A smoothing algorithm for nonlinear time series. *Int. J. Bifurc. Chaos* **14**, 1037–1051 (2004).

Corresponding author. Correspondence and request for materials should be addressed to T. Girona. E-mail address: tarsilo.girona@gmail.com.

Acknowledgements. T. Girona is supported by an appointment to the NASA Postdoctoral Program at the Jet Propulsion Laboratory, California Institute of Technology, administered by Universities Space Research Association under contract with the National Aeronautics and Space Administration. Authors appreciate discussions on the matters of this paper with Á. Amigo and members of the Observatorio Volcanológico de Los Andes del Sur (OVDAS) - Servicio Nacional de Geología y Minería (Sernageomin); S. García and members of the Servicio Geológico Minero (SEGEMAR) and Comisión Nacional de Actividades Espaciales (CONAE).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions. T.G. conceived the idea and led the data analysis and writing of the manuscript. V.R. and P.L. collaborated in writing the manuscript and designing figures. T.G., V.R., and P.L. discussed the new algorithm, model, results, and conclusions.

Author Information

- Reprints and permissions information is available at www.nature.com/reprints.
- The authors declare that they have no competing interests.

- Correspondence and requests for materials should be addressed to tarsilo.girona@gmail.com.

FIGURES AND TABLES

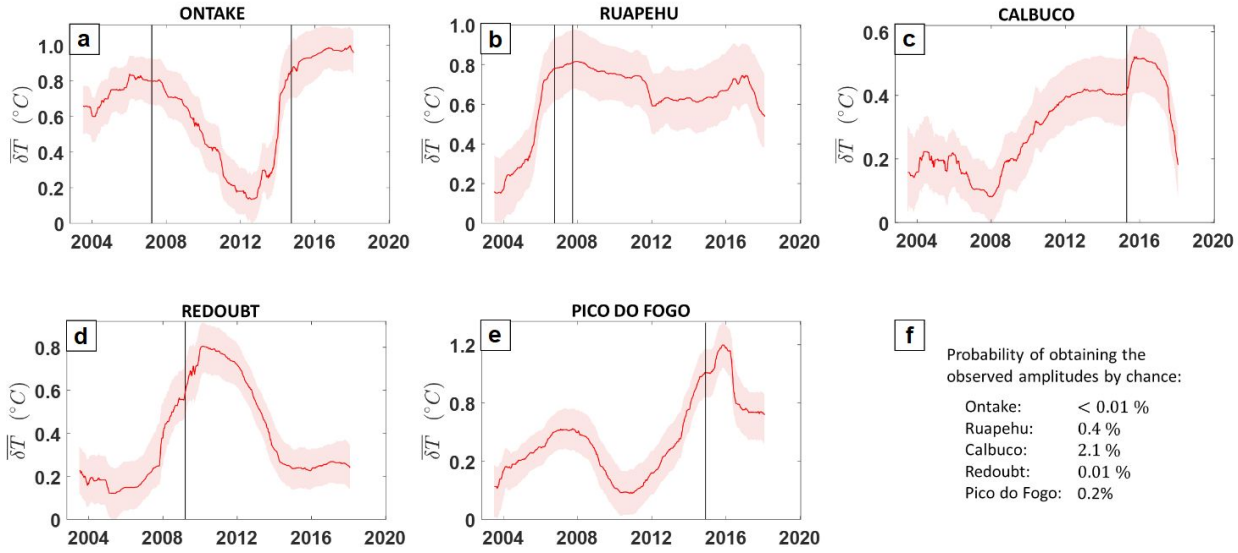

Figure 1. Temporal variations of the median anomaly $\overline{\delta T}$ at the target volcanoes. **(a-e)** Results for Ontake, Ruapehu, Calbuco, Redoubt, and Pico do Fogo. The solid red line is the median anomaly; the shaded bands represent its uncertainty (95% confidence interval); and the black vertical lines represent the onset of magmatic or phreatic eruptions. **(f)** Probability of obtaining the observed amplitudes by chance.

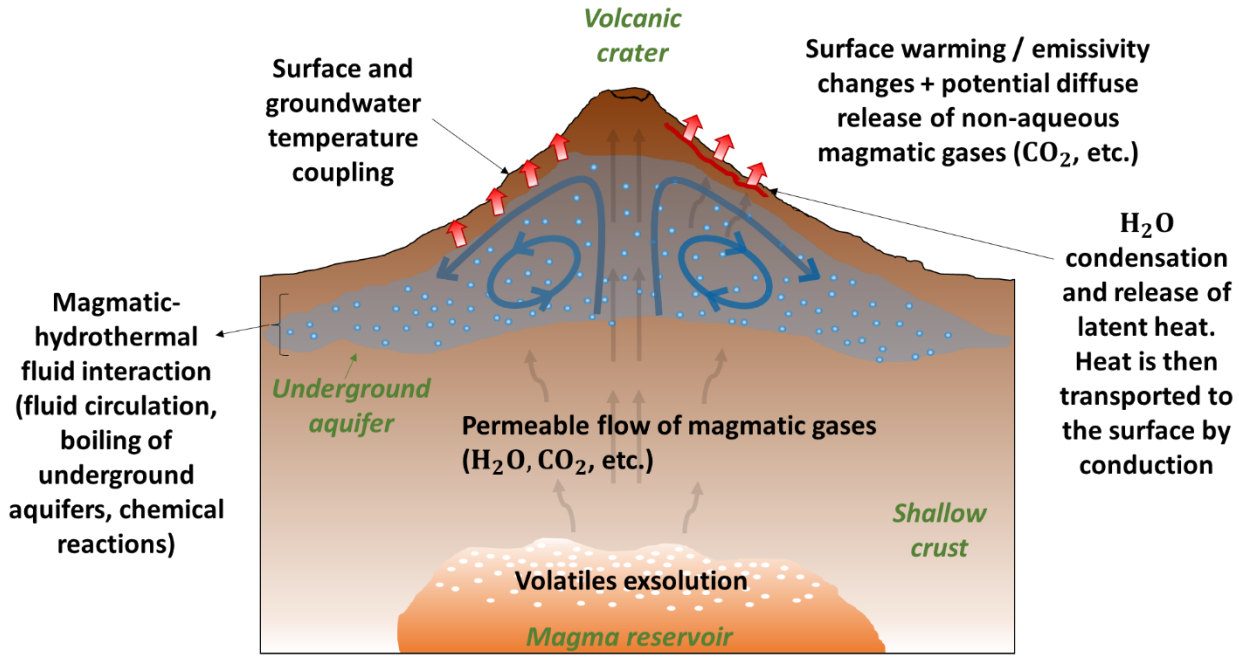

Figure 2. Conceptual model of the large-scale (10s of km^2) thermal unrest detected at the target volcanoes. We propose that the long-term (~years) pre-eruptive variations of the median anomaly reflect subsurface magmatic-hydrothermal fluid interactions, which in turn may be enhanced by diffuse degassing. In particular, magmatic gases escape from shallow magma reservoirs through the crust via permeable flow (grey arrows) and enhance circulation and/or boiling of underground aquifers (blue). The coupling between surface and groundwater temperature, the condensation of steam at shallow levels beneath the surface and consequent release of latent heat (red line), and/or variations in soil properties and emissivity changes could lead to the pre-eruptive increase of radiance (or radiant temperature) detected.

Table 1. Pre-eruptive changes of radiant temperature calculated from the median anomaly.

Volcano	Eruption start date	Type of eruption	$\overline{\delta T}(t) - \overline{\delta T}(t_0)$ (℃) [*1]
Ontake (Japan)	2007 March 24 (± 7 days)	Gas explosion (VEI=0)	0.20 (0.02 − 0.38)
	2014 September 27	Phreatic (VEI=3)	0.72 (0.51 − 0.93)
Ruapehu (New Zealand)	2006 October 4 – 2007 September 25 [*2]	Gas explosion –phreatomagmatic (VEI=1)	0.65 (0.42 − 0.88) [*3]
Calbuco (Chile)	2015 April 22	Magmatic (VEI=4)	0.32 (0.19 − 0.45)
Redoubt (Alaska, USA)	2009 March 15	Magmatic (VEI=3)	0.47 (0.30 − 0.64)
Pico do Fogo (Cape Verde)	2014 November 23	Magmatic (VEI=3)	0.82 (0.59 − 1.05)

[*1]Pre-eruptive warming, i.e., median anomaly difference between the onset of an eruption ($\overline{\delta T}(t)$) and the onset of the pre-eruptive warming phase ($\overline{\delta T}(t_0)$). [*2]Our method does not allow us to distinguish different events occurring over such a short time. [*3]These are minimum estimates because the eruptions did occur at the beginning of the time series and thus part of the pre-eruptive increase may have been missed.