

Substantial cooling effect from aerosol-induced increase in tropical marine cloud cover

In the format provided by the
authors and unedited

1

2 **Table-of-Contents:**

3

4 Supplementary Discussion Section S1-S2

5 Supplementary Table S1

6

7

8

9

10

Supplementary Discussion Section S1

Fidelity of machine-learning approach in discerning signals of aerosol-cloud interactions

The fidelity of this machine-learning (ML) approach to discern aerosol fingerprints on clouds has been well demonstrated in a recent study⁶. We further demonstrate it in this study with two steps: 1) using the “leave-one-year-out” cross validation to demonstrate the excellent ability of ML surrogates to reproduce cloud properties under normal conditions (Extended Data Fig. 1), and 2) using a statistically robust Monte Carlo-bootstrapping approach to demonstrate that the discerned aerosol-cloud interactions (ACI) signals from ML are statistically significant (Fig. 2). More details can be found in the Methods section.

Extended Data Fig. 1 shows comparisons of ML surrogates against observations of cloud properties (N_d , τ_{eff} , LWP and CF), radiation fluxes and precipitation, with the left panels showing validation against observations under normal (non-volcanic) conditions. ML surrogates show excellent agreement with the observations for all variables under normal conditions, with regressions (pink lines) very close to the 1:1 black lines and 80% of the data (within black dash circle) around the 1:1 lines. The middle and right panels of Extended Data Fig. 1 represent different eruption months, and show the impacts of volcanic aerosol on clouds, radiation fluxes and precipitation with regression (pink lines) clearly shifting away from the 1:1 lines.

To robustly validate and quantify the impacts of ACI (aerosol fingerprints), we adopt an advanced Monte Carlo-bootstrapping statistical approach. The statistical uncertainties of ML surrogates are estimated (Fig. 2), following the uncertainty guideline of IPCC to use the 90% probabilities of representing “very likely”⁸⁰. One can clearly screen out the impacts of volcanic aerosols by comparing the boxplots of red (volcanic eruptions) against black ones (normal conditions). The differences between red and black boxplots are statistically significant at the 95% confidence level (p -value < 0.05) of the Mann-Whitney-Wilcoxon test, with median and average values of black boxplots close to the 1:1 line. This again demonstrates the excellent fidelity of the ML approach in discerning robust ACI signals.

To further demonstrate the fidelity of the ML approach in discerning the ACI signals. We have carried out extra analyses for the volcanic plume upstream over the oceanic region, i.e., east to Hawaii ($12^{\circ}\text{N} \sim 25^{\circ}\text{N}$, $155^{\circ}\text{W} \sim 150^{\circ}\text{W}$), see Fig. S1. Hawaii is located in the trade wind region, and this upstream region is barely (or even not at all) influenced by the volcanic plumes but still close to the studied region, therefore experiencing similar cloud regimes and meteorological conditions. The hypothesis is that ACI signals should be much weaker (or even no signal) compared to the downstream region, thus increasing our confidence in the fidelity of the ML approach. We carried out the upstream region study as a separate case, instead of combining it with the downstream region as a whole to perform the ML training, so as to avoid the noise of the Hawaii Islands orographic effect in the head-wind and tail-wind areas. We examined the meteorological conditions over the upstream region for the years with volcanic events, and find that 1) there are many cloud-relevant meteorological conditions lying outside of the training dataset range for June 2008, such as RH at 850 hPa, specific humidity in many layers, instant moisture flux, etc.; 2) the extreme dry area in July 2008 (Fig. 3a) developed further to the north and covered a very large part of the upstream region and hence introduced outliers in the meteorological conditions; 3) the meteorological conditions in both June and July 2018 lay well within the range of the training datasets. ML generally has limited ability in dealing with extrapolation outside training data range⁸¹; therefore, we carried out the Monte Carlo bootstrapping ML analysis (see Method) for June and July 2018. All the cloud properties in the upstream region for the reference years (normal years without volcano) are reproduced with a similar uncertainty range compared to that of the downstream region, despite that the precipitation in the reference years was not being well reproduced by the ML approach for the upstream region (the uncertainty range increased by at least a factor of three compared with the downstream region). We therefore only discuss the cloud properties and the relevant climate impact (reflected shortwave flux at the top-of-the-atmosphere (TOA)), both of which are the most relevant factors for marine cloud brightening.

The ML surrogates are able to reproduce the non-eruption year conditions of cloud properties (black boxplots in Fig. S2a-S2b) for the upstream regions, with median and average

lying on the 1:1 line and an uncertainty range of $\pm 20\%$ (similar to downstream regions). The aerosol has negligible influences on all cloud properties and the TOA shortwave flux in June 2018, with median (and average) changes less than 5% (red boxplots in Fig. S2a). This is consistent with the hypothesis, because trade winds in June 2018 constantly advect the volcanic plume over the ocean west off Hawaii, whereas the ocean east off Hawaii was virtually untouched by the volcanic plume (Fig. S1a, aerosol optical depth (AOD) anomaly < 0.01), and therefore showed no ACI signal. In July 2018, a small portion of the volcanic plume was transported to the ocean east off Hawaii where a weak positive AOD anomaly of 0.05 observed (Fig. S1b). This led to consistent, but much weaker ACI signals (red boxplots in Fig. S2b) compared with the downstream region (red boxplots in Fig. 2c). In the upstream (downstream) region in July 2018, we observe an increase in N_d of 16% (26%), a decrease in r_{eff} of 4% (6%), a decrease in LWP of 2% (6%), and an increase in CF of 9% (36%), leading to 8% (18%) greater sunlight reflection back to space. The negligible ACI signals in the near pristine upstream testbed region in June 2018 and the weak (but consistent) ACI signals in the slighted polluted upstream testbed region in July 2018 demonstrate satisfactorily the fidelity of the ML approach in disentangling aerosol fingerprints on clouds.

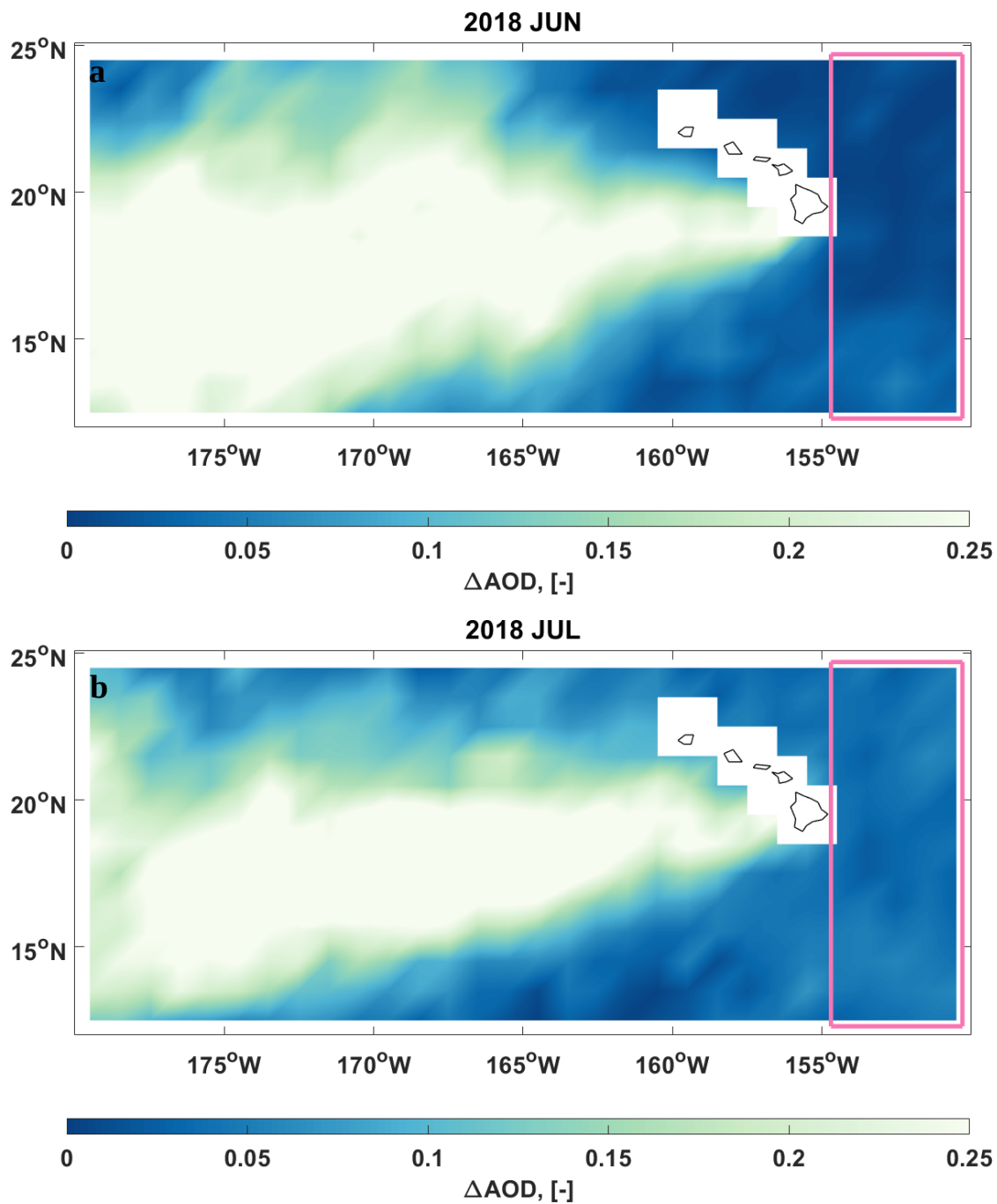


Fig. S1 | AOD climatological anomaly. Panel (a) June 2018, (b) July 2018. Climatological AOD anomaly is shown against the 2001-2020 (excluding 2008 and 2018). The volcanic plume upstream the testbed region is indicated by the pink box. Negligible AOD anomalies (<0.01 on average) are found in June 2018 (still pristine), but a weak positive AOD anomaly (average = 0.05) is found for July 2018 (slightly polluted).

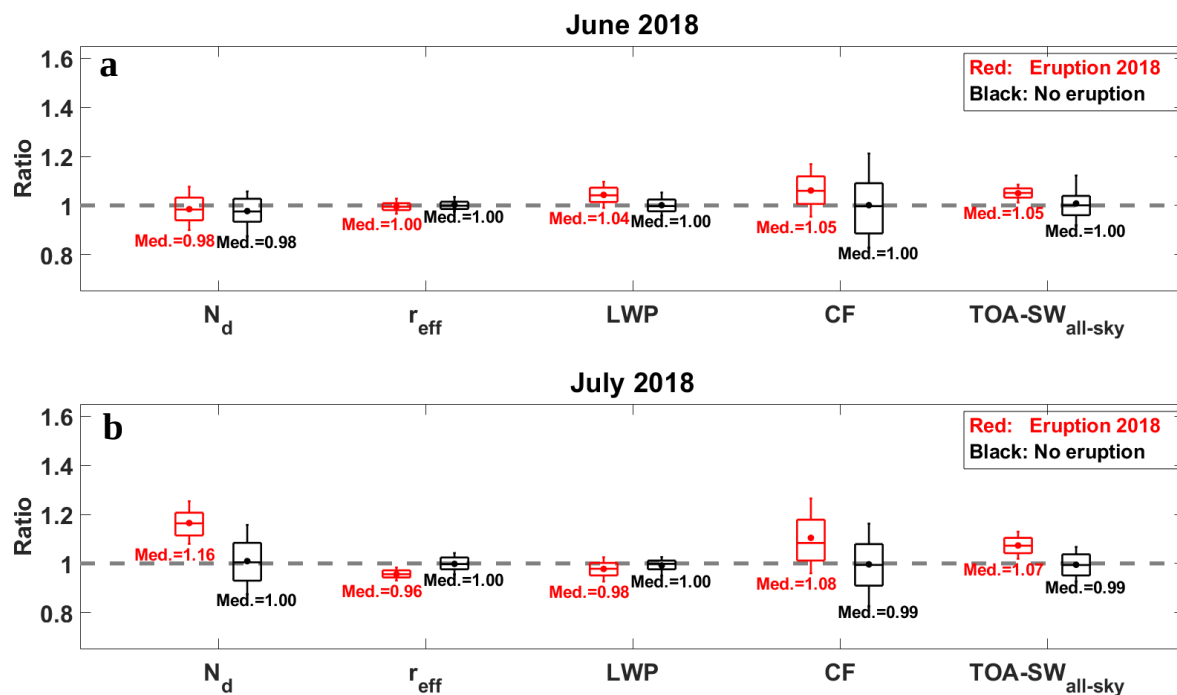


Fig. S2 | Testbed of ML approach for volcanic upstream plume. Panel (a) June 2018, (b) July 2018. The clouds in volcanic event periods are shown in red, as ratios between observations and machine-learning surrogates; non-perturbed baselines of non-volcano periods (normal conditions) are shown in black. The boxplots show 10th, 25th, median (Med.), 75th and 90th percentiles with the average indicated by a dot. The uncertainty is estimated via a bootstrapping Monte Carlo method (see Methods), with black boxes showing the validation of machine-learning surrogates against observations under normal conditions. Area (in units of km²) weighted averaging is used to calculate average cloud properties over the geographical region (the pink box in Fig. S1), in order to estimate an unbiased large-scale response.

90

91

92

93

94

95

96

97

Supplementary Discussion Section S2

Uncertainty of aerosol fingerprint analysis associated with the anomalously greater frequency of high-level clouds in July 2018.

The southern part of the studied domain ($< 15^{\circ}\text{N}$) experienced anomalously more high-level clouds in July 2018. To understand the influence of this high-level cloud anomaly, we analysed the signal of aerosol fingerprints on clouds by excluding this southern part. We find similar ACI signals compared to the original domain, indicating that our analysis and conclusion are not significantly impacted by this uncertainty. By excluding the southern part of the domain, we find that the increase in N_d changes from +26% to +28%. This slightly enhances the Twomey effect with a decrease in τ_{eff} from -6% to -7%. The ΔLWP changes from -6% to -8%. We also find a stronger increase in cloud cover (from +36% to +40%), and a slightly weaker increase in reflected shortwave flux at TOA (from +18% to +17%). This result is expected, because the anomalously greater frequency of high-level clouds partly obscure the low-level liquid clouds and would likely mask the observation of an increase in low-level clouds, meanwhile high-level clouds also contribute to the increase in reflected shortwave flux at TOA.

Although there is some uncertainty associated with the southern part of the domain with the high-level clouds anomaly, we still include it in the main text analysis. This is because we choose to keep the extremely dry scenario in July 2008 (Fig. 3a-3b), which also coincidentally occurred in the southern part of the domain ($< 15^{\circ}\text{N}$). Including this strongly contrasting RH condition provides a great opportunity to interpret the different cloud cover responses in accordance with atmospheric humidity levels.

Table S1 | List of variables in random forest based machine-learning training*. Sourced from Chen et al., (2022).

1	Temperature at 1000 hpa	Temperature at 950 hpa	Temperature at 900 hpa	Temperature at 850 hpa
2	Temperature at 800 hpa	Temperature at 750 hpa	Temperature at 700 hpa	Temperature at 650 hpa
3	Temperature at 600 hpa	Temperature at 550 hpa	Relative Humidity at 1000 hpa	Relative Humidity at 950 hpa
4	Relative Humidity at 900 hpa	Relative Humidity at 850 hpa	Relative Humidity at 800 hpa	Relative Humidity at 750 hpa
5	Relative Humidity at 700 hpa	Relative Humidity at 650 hpa	Relative Humidity at 600 hpa	Relative Humidity at 550 hpa
6	Potential Vorticity at 1000 hpa	Potential Vorticity at 950 hpa	Potential Vorticity at 900 hpa	Potential Vorticity at 850 hpa
7	Potential Vorticity at 800 hpa	Potential Vorticity at 750 hpa	Potential Vorticity at 700 hpa	Potential Vorticity at 650 hpa
8	Potential Vorticity at 600 hpa	Potential Vorticity at 550 hpa	Wind-U at 1000 hpa	Wind-U at 950 hpa
9	Wind-U at 900 hpa	Wind-U at 850 hpa	Wind-U at 800 hpa	Wind-U at 750 hpa
10	Wind-U at 700 hpa	Wind-U at 650 hpa	Wind-U at 600 hpa	Wind-U at 550 hpa
11	Wind-V at 1000 hpa	Wind-V at 950 hpa	Wind-V at 900 hpa	Wind-V at 850 hpa
12	Wind-V at 800 hpa	Wind-V at 750 hpa	Wind-V at 700 hpa	Wind-V at 650 hpa
13	Wind-V at 600 hpa	Wind-V at 550 hpa	Wind: updraft at 1000 hpa	Wind: updraft at 950 hpa
14	Wind: updraft at 900 hpa	Wind: updraft at 850 hpa	Wind: updraft at 800 hpa	Wind: updraft at 750 hpa
15	Wind: updraft	Wind: updraft	Wind: updraft	Wind: updraft

	at 700 hpa	at 650 hpa	at 600 hpa	at 550 hpa
16	Vorticity at 1000 hpa	Vorticity at 950 hpa	Vorticity at 900 hpa	Vorticity at 850 hpa
17	Vorticity at 800 hpa	Vorticity at 750 hpa	Vorticity at 700 hpa	Vorticity at 650 hpa
18	Vorticity at 600 hpa	Vorticity at 550 hpa	Specific Humidity at 1000 hpa	Specific Humidity at 950 hpa
19	Specific Humidity at 900 hpa	Specific Humidity at 850 hpa	Specific Humidity at 800 hpa	Specific Humidity at 750 hpa
20	Specific Humidity at 700 hpa	Specific Humidity at 650 hpa	Specific Humidity at 600 hpa	Specific Humidity at 550 hpa
21	Geopotential at 1000 hpa	Geopotential at 950 hpa	Geopotential at 900 hpa	Geopotential at 850 hpa
22	Geopotential at 800 hpa	Geopotential at 750 hpa	Geopotential at 700 hpa	Geopotential at 650 hpa
23	Geopotential at 600 hpa	Geopotential at 550 hpa	Longitude	Latitude
24	Dew point at 2 meter	K-index	Wind gust at 10 meter	Instant moisture flux
25	Large-scale precipitation fraction	Large-scale precipitation	Precipitation type	Friction velocity
26	Wind speed at 10 meter	Wind-U at 100 meter	Wind-V at 100 meter	Sea-ice area fraction
27	Skin temperature	Total column water vapor	Convective available potential energy	Sea surface temperature
28	Mean sea level pressure	Large-scale rain rate	Total column rain water	Total precipitation
29	Boundary layer height	Trapping layer base height		