Supplementary Information:

Machine-learning reveals climate forcing from aerosols is dominated by increased cloud cover

Ying Chen[1*,#], Jim Haywood[1,2], Yu Wang[3], Florent Malavelle[4], George Jordan[2], Daniel Partridge[1], Jonathan Fieldsend[1], Johannes De Leeuw[5], Anja Schmidt[5,6,†], Nayeong Cho[7], Lazaros Oreopoulos[7], Steven Platnick[7], Daniel Grosvenor[8], Paul Field[4,9], Ulrike Lohmann[3]

[1]College of Engineering, Mathematics, and Physical Sciences, University of Exeter, UK
[2]Met Office Hadley Centre, Exeter, UK
[3]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[4]Met Office, Exeter, UK
[5]Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge, UK
[6]Department of Geography, University of Cambridge, UK
[7]Earth Sciences Division, NASA GSFC, Greenbelt, Maryland, USA
[8]National Centre for Atmospheric Sciences, University of Leeds, Leeds, UK
[9]School of Earth and Environment, University of Leeds, Leeds, UK
[*]Correspondence to: Ying Chen (y.chen6@exeter.ac.uk; ying.chen@psi.ch)
[#]Now at Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, Villigen, Switzerland
[†]Now at Institute of Atmospheric Physics (IPA), German Aerospace Center (DLR), Oberpfaffenhofen, Germany and Meteorological Institute, Ludwig Maximilian University of Munich, Munich, Germany

This PDF file includes:

Supplementary Discussion

Section S1: Discussion of results in September 2014

The Holuhraun effusive eruption also resulted in a massive aerosol plume in the lower troposphere in September 2014. Unlike October 2014, the unusual easterly wind in September 2014 brought European outflow with anthropogenic aerosol to the southeast part of the geographical region (latitude < 63 ^{o}N, longitude > 30 ^{o}W)[15] in addition to the volcanic plume[15,63].

The predictions of cloud properties using the machine-learning surrogate MODIS (ML-MODIS) also validate well with the MODIS observations in non-eruption Septembers during 2001-2020 (left panels in Extended Data Fig. S7). In line with October, a clear difference between ML-MODIS prediction and MODIS observation is observed in September 2014 due to the volcanic eruption (right panels in Extended Data Fig. S7). The volcanic aerosol-perturbation led to a clear increase in N_d and a decrease in r_{eff}, as expected, especially over the northeast quarter of the geographical region (Extended Data Fig. S9a and S9b), which is dominated by volcanic aerosol plume[15]. The enhanced N_d and the resultant Twomey r_{eff} effect are clearly discerned over all latitude bands, and lead to a higher cloud fraction (Extended Data Fig. S9c). The spatial patterns of volcanic aerosol-perturbation induced changes in cloud properties are similar to the spatial patterns of climatological anomalies (right panels in Extended Data Fig. S9). However, the unusual easterly European outflow increases aerosol loading in the southeast part of the region and leads to stronger cloud responses than that induced solely by Holuhraun plume. This noise cannot be ruled out using climatological anomaly analysis[15]. Our machine-learning approach is able to account for meteorology variability, and rules out the noise driven by the unusual easterly flow. We therefore quantify significantly weaker responses in N_d, r_{eff} and CF using the machine-learning approach over the southeast part of the region (compared left and right panels in Fig. S9). This again further demonstrates the viability of our machine-learning approach in identifying changes in cloud created by volcanic aerosols above and beyond the expected meteorological variability.

Our Monte Carlo analysis also shows statistically significant changes in N_d, r_{eff} and CF due to the volcanic eruption. The variability of N_d and r_{eff} response signals lie outside the uncertainty range (Extended Data Fig. S8a). Most parts of CF response signal lies outside the uncertainty range, with a significant shift to higher cloud fraction (Extended Data Fig. S8a). The N_d is increased by 22% on average, leading to a 6% decrease in r_{eff} and a 5% relative increase in CF on median (and average) over the domain. We observe a weak LWP increase of 2% (1.01 minus 0.99) on average in September 2014, but it is not significant because the signal variability lies in the uncertainty range (Extended Data Fig. S8a). The susceptibilities of r_{eff}, LWP and CF to N_d are estimated as 0.31 [CI90: 0.15 ~ 0.59], 0.10 [CI90: -0.11 ~ 0.42] and 0.25 [CI90: -0.10 ~ 0.55], respectively, where CI90 stands for 90% confidence interval. Therefore, according to Eq. (3), the relative contributions to ACI-induced radiative forcing are 46 ± 29% (CF adjustment), 42 ± 16% (Twomey r_{eff} effect) and 12 ± 20% (LWP adjustment), respectively. It is worth noting that these values in September are potentially less representative of the global distribution of cloud regimes than in October, because of the more limited areal extent of the plume[15]. However, all of these responses are in line with the findings in October, providing further evidence of that our findings are robust.

Section S2: Discussion on the influence of the sea-surface temperature (SST) anomaly in 2014 in disentangling the ACI-induced CF increase

In October 2014, a cold SST anomaly developed to the south (45°N ~ 60°N, 20°W ~ 45°W) of the study domain, owing to factors that appear to be independent from the Holuhraun eruption[42]. These colder SSTs could favor a higher low-level liquid CF even without volcanic aerosol perturbation, due to enhanced static stability and thinner boundary layers[64-67]. While such a confounding factor induced by SST anomaly is not accounted for in the climatological analysis using only MODIS observations, our machine-learning (ML) approach however accounts for it, because CF results from ML-MODIS predictions experience the same SST conditions as the MODIS observations.

The cold SST anomaly does not undermine the ML representation of SST variability. The SST in 2014 lies entirely in the variability range of the ML training dataset (see Extended Data Fig. S10a). The prediction of CF over the anomaly region is based on the ML trained by the large dataset over the entire study domain spanning the years 2001-2020, which consists of 64,713 pairs of training data. The cold SST anomaly in 2014 actually shifts the SST probability distribution towards the center of the SST distribution of the training dataset, instead of shifting it outside the range of variability (blue bars in Extended Data Fig. S10a). However, there remains the possibility that the co-variation of meteorological variables associated with the cold SST anomaly may result in multi-variate conditions that are not well captured by the range found in the training conditions. We therefore perform a new Monte Carlo ML analysis which excludes the regions where for the October 2014 SSTs lie outside of the climatological range for the same place. This extreme cold SST anomalies occurred in fewer than 5% of the pixels in the domain. We find a negligible difference between this new ML analysis (Extended

Data Fig. S8b) and the initial ML analysis (Fig. 3), indicating that the ACI signals derived using our ML approach are not significantly impacted or contaminated by the SST anomaly.

The strong cold SST anomaly is limited to a region south of approximately 60°N (Extended Data Fig. S5a), while the impact of Twomey effect (a well-documented indicator of ACI[8,9,15,18]) is clearly seen Atlantic-wide (Fig. 2b), and coincide with an Atlantic-wide CF response (Fig. 2c). Compared with climatological analysis (Extended Data Fig. S3c), one of the impacts of the ML approach is to reduce the CF response in the south where SSTs are below average (most clearly seen in the zonal mean at ~52°N, Fig. 2c .vs. Extended Data Fig. S3c). These indicate that ML is able to distinguish the extra CF increase due to aerosol on top of the likely CF increase due to SST-covariant factors.

To further demonstrate the fidelity of the ML approach in disentangling ACI signals, the total impact of the SST anomaly on CF should be indicated by anomalous low-level cloud cover (LCC) in ERA5 reanalysis, where the volcanic aerosol is not included. Despite the higher LCC in October 2014 to the south due to cold SST anomaly, this cloud fraction increase in the ERA5 anomaly analysis (which accounts for meteorology only) is significantly less than the perturbation that is derived from either ML-MODIS (which accounts for aerosols) or MODIS alone (which accounts for aerosols and meteorology). In addition, many regions in the ERA5 anomaly analysis show reduction in cloud fraction (Extended Data Fig. S5c), while ML-MODIS (Fig. 2c) and MODIS (Extended Data Fig. S3c) show Atlantic-wide increases in cloud fraction. We recognize that the CF from MODIS and LCC from ERA5 are derived from satellite and model reanalysis, respectively, and are not entirely equivalent or directly comparable; but again, this result suggests that aerosols are the primary driver of the increase in CF over the study domain. Negligible LCC anomaly is found in September 2014 in ERA5 (not shown here), because SST anomaly is much weaker in September; however, a consistent increase of CF in all latitude zonal means with an average of 0.02 is found in our ML-MODIS results, which should be solely due to ACI (see Extended Data Fig. S9c).

Table S1 | The explanatory variables in random forest based machine-learning training[*].

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 at 700 hpa	Wind: updraft at 650 hpa	Wind: updraft at 600 hpa	Wind: updraft 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		

**All datasets are available from ECMWF ERA5 reanalysis.*