

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

## Journal Article

### Author(s):

Chen, Ying; Haywood, Jim; Wang, Yu; Malavelle, Florent; Jordan, George; Partridge, Daniel; Fieldsend, Jonathan; De Leeuw, Johannes; Schmidt, Anja; Cho, Nayeong; Oreopoulos, Lazaros; Platnick, Steven; Grosvenor, Daniel; Field, Paul; Lohmann, Ulrike



### Publication date:

2022-08

### Permanent link:

<https://doi.org/10.3929/ethz-b-000561341>

### Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

### Originally published in:

Nature Geoscience 15, <https://doi.org/10.1038/s41561-022-00991-6>

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

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

7     <sup>1</sup>College of Engineering, Mathematics, and Physical Sciences, University of Exeter, UK

8     <sup>2</sup>Met Office Hadley Centre, Exeter, UK

9     <sup>3</sup>Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

10    <sup>4</sup>Met Office, Exeter, UK

11    <sup>5</sup>Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of  
12    Cambridge, UK

13    <sup>6</sup>Department of Geography, University of Cambridge, UK

14    <sup>7</sup>Earth Sciences Division, NASA GSFC, Greenbelt, Maryland, USA

15    <sup>8</sup>National Centre for Atmospheric Sciences, University of Leeds, Leeds, UK

16    <sup>9</sup>School of Earth and Environment, University of Leeds, Leeds, UK

17    \*Correspondence to: Ying Chen ([y.chen6@exeter.ac.uk](mailto:y.chen6@exeter.ac.uk); [ying.chen@psi.ch](mailto:ying.chen@psi.ch))

18    #Now at Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, Villigen, Switzerland

19    †Now at Institute of Atmospheric Physics (IPA), German Aerospace Center (DLR),  
20    Oberpfaffenhofen, Germany and Meteorological Institute, Ludwig Maximilian University of  
21    Munich, Munich, Germany

22

23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

**Abstract:**

**Aerosol-cloud interactions have a potentially large impact on climate, but are poorly quantified and thus contribute a significant and long-standing uncertainty in climate projections. The impacts derived from climate models are poorly constrained by observations, because retrieving robust large-scale signals of aerosol-cloud interactions are frequently hampered by the considerable noise associated with meteorological co-variability. The Iceland-Holuhraun effusive eruption in 2014 resulted in a massive aerosol plume in an otherwise near-pristine environment and thus provided an ideal natural experiment to quantify cloud responses to aerosol perturbations. Here we disentangle significant signals from the noise of meteorological co-variability using a satellite-based machine-learning approach. Our analysis shows that aerosols from the eruption increased cloud cover by approximately 10%, and this appears to be the leading cause of climate forcing, rather than cloud brightening as previously thought. We find that volcanic aerosols do brighten clouds by reducing droplet size, but this has a significantly smaller radiative impact than changes in cloud fraction. These results add substantial observational constraints on the cooling impact of aerosols. Such constraints are critical for improving climate models, which still inadequately represent the complex macro-physical and micro-physical impacts of aerosol-cloud interactions.**

46 Marine low-level liquid clouds have a profound impact on the energy balance of the Earth  
47 system, exerting a net cooling effect by reflecting sunlight<sup>1,2</sup>. It has been previously estimated  
48 that only a 6% increase of their albedo could offset the warming from a doubling of CO<sub>2</sub><sup>3,4</sup>.  
49 Aerosol-cloud interactions (ACI) are postulated to enhance albedo and prolong the lifetime of  
50 liquid clouds<sup>5,6</sup>, and therefore counterbalance a substantial, yet poorly constrained, portion of  
51 greenhouse gas warming<sup>7-10</sup>, leading to only a small net positive overall forcing. As the Earth  
52 has warmed by around 1.2 °C since pre-industrial times<sup>10,11</sup>, this would imply that the Earth  
53 system is highly sensitive, and therefore vulnerable, to anthropogenic climate forcing<sup>12</sup>. Such  
54 a high sensitivity would suggest a very limited remaining carbon budget if the +1.5 °C target  
55 of the 21<sup>st</sup> Conference of the Parties at Paris (COP21) is to be met<sup>11</sup>.

56 Despite decades of effort, ACI still contribute significantly to uncertainties in climate  
57 projections<sup>1,7,9-11</sup>. A primary reason for the large uncertainty in ACI is the lack of suitable large-  
58 scale constraints to challenge General Circulation Models (GCMs)<sup>13-15</sup>. ACI operates through  
59 processes whereby cloud droplets form on aerosol particles. For a fixed cloud liquid water path  
60 (LWP), high concentrations of aerosol lead to more droplets with smaller effective radius ( $r_{\text{eff}}$ ,  
61 Twomey  $r_{\text{eff}}$  effect<sup>5</sup>) which increases cloud albedo. Smaller cloud droplets may inhibit  
62 precipitation due to weakened collision-coalescence<sup>6</sup> and suppressed precipitation implies  
63 clouds retain more water leading to an increased LWP (LWP adjustment), and prolong their  
64 lifetime and areal extent which manifests as increased cloud fraction (CF, CF adjustment)<sup>6</sup>.  
65 There is clear evidence of the Twomey  $r_{\text{eff}}$  effect from numerous comprehensive satellite  
66 observations (e.g., ref. <sup>8,15-19</sup>), but continuous debate surrounds the LWP adjustment with  
67 different magnitudes and signs reported<sup>8,9,15,20,21</sup>, possibly due to confounding adjustments such  
68 as effects of entrainment and droplet evaporation processes<sup>22-26</sup>. The CF adjustment is even  
69 more difficult to constrain owing to the large-scale impacts of meteorological co-variability<sup>27</sup>,  
70 leading to long-standing and ongoing disputes in the scientific literature<sup>16,19,28-32</sup>. Satellite

71 observational constraints of ACI tend to be limited to either small-scale observations or large-  
72 scale climatological analyses<sup>33</sup>. A typical example of a small-scale observation is “ship-tracks”,  
73 manifesting as brighter lines in stratocumulus cloud decks caused by ship emissions. Such  
74 small-tracks are generally able to rule out confounding meteorology<sup>8,19</sup>, but with a scale far  
75 below the resolution of GCMs and a short temporal signature; they are therefore not ideal  
76 constraints for these models<sup>33,34</sup>. Climatological analyses examine the correlations between  
77 cloud properties and aerosol on a large spatiotemporal scale, but such correlations can be  
78 confounded by meteorological co-variability and therefore may not confirm the causal  
79 processes of ACI<sup>29,33,35-38</sup>.

80 Here, we overcome these limitations by developing a meteorological reanalysis and satellite-  
81 based machine-learning approach that predicts cloud properties in a near-pristine environment,  
82 and compare the results with observations of clouds perturbed by the large-scale effusive  
83 Icelandic eruption of Holuhraun. The machine-learning approach is enabled by an almost  
84 threefold expansion of satellite data from Moderate Resolution Imaging Spectroradiometer  
85 (MODIS) compared to the earlier work<sup>15</sup>, offering thus a robust training dataset. The machine-  
86 learning approach allows us to quantify ACI-induced cloud responses and show an  
87 unmistakable increase in cloud cover. It also allows us to infer the relative contributions to  
88 ACI radiative effect from the Twomey effect, and the LWP and CF adjustments. Our results  
89 improve current understanding of cloud-induced climate change, and provide robust large-  
90 scale constraints for climate models.

91

## 92 **Volcanic aerosol perturbation**

93 The effusive volcanic eruption at Holuhraun in Iceland, emitted about 40,000 tonnes of SO<sub>2</sub>  
94 per day on average during its eruptive phase in September-October 2014 and 120,000 tonnes

95 per day at the peak of eruption<sup>15,39</sup>. The sulphate aerosol formed from volcanic SO<sub>2</sub> interacts  
96 with liquid-water clouds creating an invaluable natural experiment for testing ACI hypotheses  
97 at a large-scale<sup>15</sup>. Detecting CF changes above meteorological noise requires a larger data  
98 volume and was left unexplored in the previous study<sup>15</sup>, which uses the MODIS Aqua 2002-  
99 2014 dataset. Here, by extending the satellite data to both MODIS Aqua and Terra and the  
100 length of the analysis period to 2001-2020, we have sufficient training data to develop a robust  
101 machine-learning approach for quantitatively disentangling Holuhraun eruption ACI signals  
102 from the noise of meteorological co-variability (see Methods). We focus primarily on October  
103 2014, because in this second eruption month the volcanic plume dispersed sufficiently across  
104 the entire region of about 3000 km × 6000 km (45°N ~ 75°N; 60°W ~ 30°E, see Supplementary  
105 Figure S6.2 in Malavelle et al.<sup>15</sup>). This region is an otherwise near-pristine environment and  
106 encompasses the whole spectrum of liquid-dominated cloud regimes, with their frequencies of  
107 occurrence being comparable to those observed globally (Extended Data Fig. 1)<sup>15,40</sup>.

108 To disentangle the ACI signal from the noise of meteorological co-variability, we train a  
109 machine-learning surrogate MODIS (ML-MODIS) using historical meteorology and MODIS  
110 observations during 2001-2020 but excluding the year of the volcanic perturbation (2014, see  
111 Methods). ML-MODIS is designed to predict cloud properties for given meteorological  
112 conditions when unperturbed by volcanic aerosol. Our “leave-one-year-out” cross validation  
113 (see Methods) shows that the surrogate ML-MODIS can reproduce the MODIS observations  
114 of cloud droplet number concentration ( $N_d$ ),  $r_{eff}$ , LWP and CF when no volcanic aerosol-  
115 perturbation exists, as shown in the left column of Fig. 1. However, significant differences  
116 between the ML-MODIS predictions and MODIS observations are observed in the presence of  
117 the volcanic perturbation in October 2014 (right column of Fig. 1). Similar results are found  
118 for September 2014 (Supplementary Discussion section S1).

119 We examine the ACI corresponding to the increase in  $N_d$  instead of aerosol optical depth,  
120 because MODIS aerosol products are hampered by the overcast nature of the geographical  
121 region and using  $N_d$  has several advantages as a mediating variable<sup>29</sup>. We first quantify the  
122 increase in  $N_d$  and then estimate the susceptibility of other cloud properties, i.e.,  $d\ln r_{\text{eff}}/d\ln N_d$ ,  
123  $d\ln \text{LWP}/d\ln N_d$ , and  $d\ln \text{CF}/d\ln N_d$ . The volcano-induced increase in  $N_d$  is observed across  
124 nearly the entire region with a positive signal across the zonal means (Fig. 2a). We also observe  
125 a clear shift of the  $N_d$  probability distribution towards larger values due to the volcanic  
126 perturbation with an average increase of  $20 \text{ cm}^{-3}$ .

127 We perform Monte Carlo analyses (see Methods) to estimate the uncertainty of ML-MODIS  
128 and to quantify the impact of ACI on relevant cloud properties. In assessing the statistical  
129 uncertainties, we follow the Intergovernmental Panel on Climate Change (IPCC) uncertainty  
130 guideline<sup>41</sup> and use the 90% probabilities (that are assigned “very likely”). A validation of ML-  
131 MODIS by MODIS for conditions unperturbed by Holuhraun is further achieved by these  
132 results, with median and average values close to the 1:1 line (Fig. 3) and with a 90% probability  
133 of the Pearson correlation coefficients ( $R$ ) exceeding 0.6 for  $N_d$ ,  $r_{\text{eff}}$  and CF (Extended Data  
134 Fig. 2, higher than 0.5 for LWP). In contrast, the 90% probability of  $R$  being below 0.6 for all  
135 cloud properties in volcano-perturbed conditions, indicates large influences of the volcanic  
136 aerosol on cloud properties. We estimate a volcanic aerosol-induced increase in  $N_d$  of 28%  
137 over the region (Fig. 3, showing that the ratio between ML-MODIS and MODIS is 1.27 with  
138 against 0.99 without volcano), which is clearly statistically significant because the perturbation  
139 lies outside the range of uncertainty of the machine-learning method. This increase is similar  
140 to the ~32% increase in  $N_d$  from pre-industrial to present day according to multi-model  
141 estimates<sup>14</sup>, suggesting that the results from our analysis may be a reasonable proxy for  
142 anthropogenic aerosols in terms of the strength in perturbing clouds since pre-industrial times.

143

#### 144 **Twomey effect and liquid water path adjustment**

145 We first use our machine-learning approach to examine the Twomey  $r_{\text{eff}}$  effect and LWP  
146 adjustment. We observe a consistent spatial pattern of volcano-induced increase in  $N_d$  and an  
147 average reduction in  $r_{\text{eff}}$  (Figs. 2a and 2b) from 15.2  $\mu\text{m}$  to 13.9  $\mu\text{m}$ . The spatial pattern is also  
148 consistent with the climatological MODIS anomaly analysis<sup>15</sup> (Extended Data Fig. 3), but with  
149 some difference in the strength of ACI signal. This further demonstrates the viability of our  
150 machine-learning approach in identifying changes in cloud created by volcanic aerosols above  
151 those expected due to meteorological variability. Climatological anomalies may identify  
152 regions influenced by the Holuhraun plume<sup>15</sup> but may not be robust in quantifying ACI signals  
153 arising from Holuhraun, because the ACI signal is confounded by meteorology where 2014  
154 conditions are not necessarily equal to climatological average. Indeed, while Malavelle et al.<sup>15</sup>  
155 developed a robust method for removing the meteorological variability in the modelled  
156 response, they also cautioned that meteorological differences from the long-term mean could  
157 cause some of the observed response (their Figures S6.1 and S6.2). Our machine-learning  
158 approach overcomes these issues (Methods, see also Supplementary Discussion section S1 and  
159 S2). We estimate an 8% decrease in  $r_{\text{eff}}$  as a response to a 28% increase in  $N_d$  on average (and  
160 median) over the geographical region (Fig. 3). In line with previous studies<sup>8,17,31</sup>, no significant  
161 LWP response is found when examining the region as a whole (Fig. 3 and Extended Data Fig.  
162 4). This may be due to the cancellation of the LWP adjustment-induced increase<sup>6</sup> by  
163 entrainment-induced decrease of LWP<sup>22</sup>, as suggested by Toll et al.<sup>8</sup> who examined over  
164 10,000 globally representative aerosol-perturbation tracks of small-scale in liquid clouds.

165

#### 166 **Cloud fraction response**



167 So far results from our large-scale machine-learning approach agree with previous analyses: a  
168 distinct and robust Twomey  $r_{\text{eff}}$  effect but a weak LWP adjustment (e.g., ref. <sup>8,17,31</sup>). We now  
169 examine the adjustment of liquid phase CF, which is a macro-property of cloud and difficult to  
170 examine using small-scale aerosol-induced tracks<sup>8</sup>. Our results of volcanic aerosol-perturbed  
171 conditions show an overall increase of zonal CF at all latitudes of our domain, and a clear shift  
172 of probability distribution from a median value of 0.36 to 0.39 (Fig. 2c). The CF increase  
173 exhibits a spatial pattern that is consistent with the Twomey  $r_{\text{eff}}$  effect (Fig. 2b and 2c). This  
174 strongly suggests that it is the aerosol perturbation that leads to increased cloud cover, since  
175 the Twomey  $r_{\text{eff}}$  effect has been well documented as an ACI indicator<sup>8,9,15,18</sup>.

176 We present the response of CF and other cloud properties over the geographical region using  
177 the Monte Carlo method (Fig. 3). For all non-perturbed cloud properties, the validation shows  
178 the median and average values on the 1:1 line. For volcano-induced changes in  $N_d$  and  $r_{\text{eff}}$ , we  
179 confirm the expected increase and decrease respectively, but see little LWP response. For CF,  
180 we observe a statistically significant median (and average) relative increase of 11% with the  
181 signal variability range lying outside the uncertainty. We estimate  $d\ln CF/d\ln N_d = 0.41$  [0.05 ~  
182 1.53, 90% confidence interval], indicating a strong susceptibility of CF to aerosol-induced  
183 perturbation in  $N_d$ . Rosenfeld et al.<sup>30</sup> found a similar strong susceptibility using a climatological  
184 approach, but for the convective cores of southern ocean liquid clouds. This strong  
185 susceptibility is also consistent with other studies (e.g. ref. <sup>16,29,31,36</sup>), although, unlike the  
186 present study, their results are likely either influenced by the confounding meteorology  
187 associated with the climatological correlation approach<sup>33,36</sup> or limited by relatively small-scale  
188 Lagrangian trajectories<sup>33</sup>. For example, Ghan et al.<sup>14</sup> showed that climatological correlation  
189 analysis differs greatly from perturbation analysis across multiple GCMs, despite efforts to  
190 classify and isolate different meteorological regimes.

191 To back up our finding of CF increase, we perform a traditional climatological anomaly  
192 analysis which shows a similar spatial pattern for the CF response (Extended Data Fig. 3c).  
193 Additionally, we investigate the impact of the unusually low sea-surface temperature that  
194 developed to the south of the region (Extended Data Fig. 5a) owing to factors that appear to be  
195 independent from the eruption<sup>42</sup>. While this could affect CF, it cannot be accounted for in the  
196 climatological anomaly analysis using only MODIS data. Our machine-learning approach,  
197 however, accounts for this variability (Extended Data Fig. 6 and Supplementary Discussion  
198 section S2). We are therefore in position to better quantify a significantly weaker CF increase  
199 over the corresponding region (45°N ~ 60°N, 20°W ~ 45°W; compare Fig. 2c against Extended  
200 Data Fig. 3c). We also find 14% fewer cloud-free high-resolution (1-km<sup>2</sup>) MODIS pixels  
201 during October 2014 compared to the long-term October mean. Again, this implies CF  
202 increases in response to the volcanic aerosol. Any conceivable increase in cloud cover from  
203 ice-clouds is also investigated and cannot be discerned (Extended Data Fig. 5b); this suggests  
204 that any potential confounding effect from ice-cloud or transition to ice-cloud is small, and that  
205 our results regarding ACI of liquid clouds are robust.

206

### 207 **Cloud fraction adjustment dominates radiative forcing**

208 We revisit the relative contributions to ACI-induced radiative forcing from the Twomey effect,  
209 LWP and CF adjustments, see Methods section “Radiative Forcing”. In line with previous  
210 studies<sup>8,31</sup>, we find a weak contribution ( $2 \pm 17\%$ ) from the LWP adjustment. However, in  
211 contrast to recent studies reporting that the Twomey  $r_{\text{eff}}$  effect dominates ( $> 70\%$ ) the ACI  
212 radiative forcing<sup>8,19,31</sup>; we show that, for this large-scale study across a wide range of  
213 meteorological and cloud regimes, the CF adjustment ( $61 \pm 23\%$ ) surpasses the Twomey  $r_{\text{eff}}$   
214 effect ( $37 \pm 18\%$ ) in terms of ACI cooling (Fig. 3). This new finding may be due to the much

215 larger spatiotemporal scales of our investigation, which extends up to tens of thousands of km  
216 with perturbation lasting for months. Given the large range of meteorological conditions and  
217 cloud regimes included (Extended Data Fig. 1), our study appears arguably more suitable for  
218 constraining large-scale climate models and ACI associated with anthropogenic emissions,  
219 which themselves persist across many geographical areas and are associated with a wide variety  
220 of cloud regimes.

221 Our results suggest that cooling caused by a CF increase is substantially underestimated in  
222 current climate projections<sup>10</sup>. A recent multi-model assessment of the susceptibility of  
223  $d\ln CF/d\ln N_d$  versus  $-d\ln r_{\text{eff}}/d\ln N_d$  (Ghan et al.<sup>14</sup>; their Figure 1) suggests ratios of  
224 approximately 1:3. Our results suggest that the CF adjustment is possibly larger than the  
225 Twomey  $r_{\text{eff}}$  effect, since the ratio of their susceptibilities is around 5:4. It is possible that GCMs  
226 compensate for the lack of CF response with overly strong LWP adjustment<sup>8,10,15,19,34</sup> – i.e.  
227 estimate the “right” cooling but for manifestly the wrong reasons.

228 This work sheds light into certain aspects of ACI which conventionally thought to follow the  
229 following route: an increase in aerosols gives rise to i) an increase in  $N_d$  leading to ii) a larger  
230 number of smaller cloud droplets leading to iii) a decrease in the collision-coalescence growth  
231 rate of cloud droplets, leading to iv) a reduction in precipitation leading to v) an increase in  
232 LWP leading to vi) an increase in cloud lifetime leading to vii) an increase in CF. Malavelle et  
233 al.<sup>15</sup> suggested that iv) and v) do not operate as expected, while, this new study provides strong  
234 evidence for vi) and vii). This conundrum needs to be addressed in further research.  
235 Suggestions for how to approach this in future work includes performing large eddy model  
236 simulation of the Holuhraun event to identify difference in the ACI causal chain between the  
237 heavily parameterized GCMs representation and the more explicit cloud-resolving models.  
238 Identifying any changes in cloud regimes (e.g., ref. <sup>31,40</sup>) might also provide further clues in

239 solving this puzzle. We maintain that because clouds are such a fundamentally important  
240 component of the Earth's hydrological cycle and energy flows that the underlying reasons of  
241 deficient model performance need to be urgently addressed. Our findings appear to provide  
242 robust new constraints for climate models, despite the uncertainties associated with machine-  
243 learning and MODIS retrievals. We acknowledge that the cold SST anomaly in October 2014  
244 could potentially introduce more uncertainty in the machine-learning representation of cloud  
245 conditions, but this influence appears insignificant in this study (Supplementary Discussion  
246 Section S2). ACI signals are statistically significant, lying outside the uncertainty range of the  
247 machine-learning approach (Fig. 3). Uncertainty in the MODIS retrievals can be decomposed  
248 into systematic errors and random errors. Random errors are greatly suppressed by averaging  
249 over a geographical region of thousands of kilometres<sup>43</sup>, while systematic errors are largely  
250 cancelled when taking differences between MODIS and ML-MODIS<sup>8</sup>.

251 The quantified constraints from our machine-learning study pave the way to advance our  
252 current understanding of physical ACI processes, and point to new directions and challenges  
253 towards future improvement of climate models. With advances in both areas, we expect that  
254 our large-scale constraints on ACI will lead to reduced uncertainty in climate projections and  
255 future estimates of climate sensitivity.

256

257

258

259

260

261

262 **Acknowledgments**

263 We would like to acknowledge the support of the UK Natural Environment Research Council  
264 (NERC) funded ADVANCE project (NE/T006897/1) which funded JH, YC, DP, AS, DG and PF. JH, GJ  
265 and FM were also part funded under funding provided by the EU's Horizon 2020 research and innovation  
266 programme under the CONSTRAIN grant agreement 820829. JH, PF, GJ, AJ and FM are supported by the  
267 Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). YW would like to thank  
268 the support from Mr. Philippe Sarasin and the ETH Zurich Foundation (ETH Fellowship project: 2021-HS-  
269 332). DG is funded by the National Centre for Atmospheric Science (NCAS), one of the UK NERC's  
270 research centres. JdL acknowledges funding from the NERC funded V-PLUS grant NE/S00436X/1. NC, LO  
271 and SEP are funded by USA NASA programs. The machine-learning training is performed using the  
272 "Statistics and Machine Learning Toolbox" in MATLAB (version R2019b, MathWorks Inc., Natick, MA,  
273 USA). We would like to thank Ken Carslaw (University of Leeds) for co-developing and co-leading the  
274 Leeds aspect of ADVANCE project, and Andy Jones (UK Met Office) for helpful discussions.

275 **Author contributions Statement**

276 YC and JH conceived the study. YW and YC designed and developed the machine-learning  
277 approach used in this study with help from JH and JF. JH led the ADVANCE project funded by UK-NERC.  
278 YC, FM, JG and JH performed the analysis of MODIS data with help from DG, NC, LO and SEP. NC, LO  
279 and SEP performed the cloud regime analysis. YC, JH, YW, DG, UL, PF, LO, SEP, JdL, AS, DP and JF  
280 contributed to the uncertainty discussion. YC and JH performed the analyses and interpreted the results with  
281 inputs from all co-authors. YC and JH led the manuscript writing with specific inputs and edits from DG,  
282 LO and UL. All co-authors discussed the results and commented on the manuscript.

283 **Competing Interests Statement:** The authors declare no competing interests.

284

285

286

287

288

289 **Figure Legends/Captions:**

290 **Fig. 1 | Comparison between machine-learning predictions (ML-MODIS) and MODIS**

291 **observations.** Left panels (a-d): validation against non-perturbed observations (excluding 2014) of  
292 cloud properties, from top to bottom they show cloud droplet number concentration ( $N_d$ ), cloud  
293 droplet effective radius ( $r_{eff}$ ), cloud liquid water path (LWP) and cloud fraction (CF). Right panels  
294 (e-h): volcanic perturbation signals in October 2014, indicated by the difference between the machine-  
295 learning predictions and the observations. October MODIS observations from Aqua (2002-2020) and  
296 Terra (2001-2020) are analyzed. Colour indicates the normalized data density function with a  
297 maximum value of one, with 80% of the data being contained within the black dashed area.

298 **Fig. 2 | Changes in cloud properties caused by the volcanic perturbation** estimated using  
299 machine-learning predictions and MODIS observations for October 2014. The spatial distribution and  
300 zonal means of the changes in  $N_d$ ,  $r_{eff}$  and CF are shown in the left panels of **a-c** while right panels  
301 show probability density functions (so that the areas under the curves are equivalent) for MODIS and  
302 ML-MODIS.

303 **Fig. 3 | Responses of cloud properties to the volcanic aerosol-perturbation in October 2014.** The  
304 aerosol-cloud interactions (ACI) signals of responses are indicated as the ratios between MODIS  
305 (Aqua and Terra) observations and machine-learning predictions, i.e., Ratio = MODIS divided by  
306 ML-MODIS. Uncertainties of non-perturbed baseline references are estimated using a Monte Carlo  
307 method and are shown in black (see Methods, based on non-volcanic October datasets spanning 2001-  
308 2020). The variability of the cloud responses to the Holuhraun volcanic aerosol perturbation are  
309 shown in pink. The boxplots show 10<sup>th</sup>, 25<sup>th</sup>, median (Med.), 75<sup>th</sup> and 90<sup>th</sup> percentiles with the mean  
310 value indicated by a dot. The susceptibilities of  $r_{eff}$ , LWP and CF to changes in  $N_d$  are given in a green  
311 colour, median [90% confidence interval]. Area (in units of km<sup>2</sup>) weighted averaging is used to  
312 calculate average cloud properties over the geographical region (Fig. 2), in order to estimate an  
313 unbiased large-scale response signal. Therefore, the ratios shown here are slightly different from the  
314 slopes shown in Fig. 1, in which area-weighted averaging is not applied.

315 **References:**

- 316 1 Lohmann, U. & Feichter, J. Global indirect aerosol effects: a review. *Atmos. Chem. Phys.* **5**,  
 317 715-737, doi:10.5194/acp-5-715-2005 (2005).
- 318 2 L’Ecuyer, T. S., Hang, Y., Matus, A. V. & Wang, Z. Reassessing the Effect of Cloud Type on  
 319 Earth’s Energy Balance in the Age of Active Spaceborne Observations. Part I: Top of  
 320 Atmosphere and Surface. *Journal of Climate* **32**, 6197-6217, doi:10.1175/jcli-d-18-0753.1  
 321 (2019).
- 322 3 Latham, J. *et al.* Global temperature stabilization via controlled albedo enhancement of low-  
 323 level maritime clouds. *Philosophical Transactions of the Royal Society A: Mathematical,*  
 324 *Physical and Engineering Sciences* **366**, 3969-3987, doi:doi:10.1098/rsta.2008.0137 (2008).
- 325 4 Chen, Y.-C., Christensen, M. W., Stephens, G. L. & Seinfeld, J. H. Satellite-based estimate of  
 326 global aerosol–cloud radiative forcing by marine warm clouds. *Nature Geoscience* **7**, 643-  
 327 646, doi:10.1038/ngeo2214 (2014).
- 328 5 Twomey, S. Pollution and the planetary albedo. *Atmospheric Environment* **8**, 1251-1256,  
 329 doi:https://doi.org/10.1016/0004-6981(74)90004-3 (1974).
- 330 6 Albrecht, B. A. Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science* **245**, 1227-  
 331 1230, doi:10.1126/science.245.4923.1227 (1989).
- 332 7 Boucher, O. *et al.* in *Climate Change 2013: The Physical Science Basis. Contribution of*  
 333 *Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate*  
 334 *Change* (eds T. F. Stocker *et al.*) 571-657 (Cambridge University Press, 2013).
- 335 8 Toll, V., Christensen, M., Quaas, J. & Bellouin, N. Weak average liquid-cloud-water response  
 336 to anthropogenic aerosols. *Nature* **572**, 51-55, doi:10.1038/s41586-019-1423-9 (2019).
- 337 9 Bellouin, N. *et al.* Bounding Global Aerosol Radiative Forcing of Climate Change. *Reviews*  
 338 *of Geophysics* **58**, e2019RG000660, doi:10.1029/2019rg000660 (2020).
- 339 10 IPCC. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I  
 340 to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-  
 341 Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L.  
 342 Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T.  
 343 Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. *Cambridge University Press In Press*  
 344 (2021).
- 345 11 Rogelj, J., Forster, P. M., Kriegler, E., Smith, C. J. & Séférian, R. Estimating and tracking the  
 346 remaining carbon budget for stringent climate targets. *Nature* **571**, 335-342,  
 347 doi:10.1038/s41586-019-1368-z (2019).
- 348 12 Andreae, M. O., Jones, C. D. & Cox, P. M. Strong present-day aerosol cooling implies a hot  
 349 future. *Nature* **435**, 1187-1190 (2005).
- 350 13 Seinfeld, J. H. *et al.* Improving our fundamental understanding of the role of aerosol–cloud  
 351 interactions in the climate system. *Proceedings of the National Academy of Sciences* **113**,  
 352 5781-5790, doi:10.1073/pnas.1514043113 (2016).
- 353 14 Ghan, S. *et al.* Challenges in constraining anthropogenic aerosol effects on cloud radiative  
 354 forcing using present-day spatiotemporal variability. *Proceedings of the National Academy of*  
 355 *Sciences* **113**, 5804-5811, doi:10.1073/pnas.1514036113 (2016).
- 356 15 Malavelle, F. F. *et al.* Strong constraints on aerosol–cloud interactions from volcanic  
 357 eruptions. *Nature* **546**, 485-491, doi:10.1038/nature22974 (2017).
- 358 16 Kaufman, Y. J., Koren, I., Remer, L. A., Rosenfeld, D. & Rudich, Y. The effect of smoke,  
 359 dust, and pollution aerosol on shallow cloud development over the Atlantic Ocean.  
 360 *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11207-  
 361 11212, doi:10.1073/pnas.0505191102 (2005).
- 362 17 McCoy, D. T. & Hartmann, D. L. Observations of a substantial cloud-aerosol indirect effect  
 363 during the 2014–2015 Bárðarbunga-Veiðivötn fissure eruption in Iceland. *Geophysical*  
 364 *Research Letters* **42**, 10,409–410,414, doi:https://doi.org/10.1002/2015GL067070 (2015).
- 365 18 Toll, V., Christensen, M., Gassó, S. & Bellouin, N. Volcano and Ship Tracks Indicate  
 366 Excessive Aerosol-Induced Cloud Water Increases in a Climate Model. *Geophysical*  
 367 *Research Letters* **44**, 12,492–412,500, doi:10.1002/2017gl075280 (2017).

- 368 19 Diamond, M. S., Director, H. M., Eastman, R., Possner, A. & Wood, R. Substantial Cloud  
369 Brightening From Shipping in Subtropical Low Clouds. *AGU Advances* **1**, e2019AV000111,  
370 doi:<https://doi.org/10.1029/2019AV000111> (2020).
- 371 20 Gryspeerdt, E. *et al.* Constraining the aerosol influence on cloud liquid water path. *Atmos.*  
372 *Chem. Phys.* **19**, 5331-5347, doi:10.5194/acp-19-5331-2019 (2019).
- 373 21 Possner, A., Eastman, R., Bender, F. & Glassmeier, F. Deconvolution of boundary layer depth  
374 and aerosol constraints on cloud water path in subtropical stratocumulus decks. *Atmos. Chem.*  
375 *Phys.* **20**, 3609-3621, doi:10.5194/acp-20-3609-2020 (2020).
- 376 22 Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E. & Toon, O. B. The impact of humidity  
377 above stratiform clouds on indirect aerosol climate forcing. *Nature* **432**, 1014-1017,  
378 doi:10.1038/nature03174 (2004).
- 379 23 Stevens, B. & Feingold, G. Untangling aerosol effects on clouds and precipitation in a  
380 buffered system. *Nature* **461**, 607-613, doi:10.1038/nature08281 (2009).
- 381 24 Lebo, Z. J. & Feingold, G. On the relationship between responses in cloud water and  
382 precipitation to changes in aerosol. *Atmos. Chem. Phys.* **14**, 11817-11831, doi:10.5194/acp-  
383 14-11817-2014 (2014).
- 384 25 Koren, I., Dagan, G. & Altaratz, O. From aerosol-limited to invigoration of warm convective  
385 clouds. *Science* **344**, 1143-1146, doi:10.1126/science.1252595 (2014).
- 386 26 Seifert, A., Heus, T., Pincus, R. & Stevens, B. Large-eddy simulation of the transient and  
387 near-equilibrium behavior of precipitating shallow convection. *Journal of Advances in*  
388 *Modeling Earth Systems* **7**, 1918-1937, doi:<https://doi.org/10.1002/2015MS000489> (2015).
- 389 27 Mauger, G. S. & Norris, J. R. Meteorological bias in satellite estimates of aerosol-cloud  
390 relationships. *Geophysical Research Letters* **34**, doi:<https://doi.org/10.1029/2007GL029952>  
391 (2007).
- 392 28 Kaufman, Y. J. & Koren, I. Smoke and Pollution Aerosol Effect on Cloud Cover. *Science*  
393 **313**, 655-658, doi:10.1126/science.1126232 (2006).
- 394 29 Gryspeerdt, E., Quaas, J. & Bellouin, N. Constraining the aerosol influence on cloud fraction.  
395 *Journal of Geophysical Research: Atmospheres* **121**, 3566-3583,  
396 doi:<https://doi.org/10.1002/2015JD023744> (2016).
- 397 30 Rosenfeld, D. *et al.* Aerosol-driven droplet concentrations dominate coverage and water of  
398 oceanic low-level clouds. *Science* **363**, eaav0566, doi:10.1126/science.aav0566 (2019).
- 399 31 Christensen, M. W., Jones, W. K. & Stier, P. Aerosols enhance cloud lifetime and brightness  
400 along the stratus-to-cumulus transition. *Proceedings of the National Academy of Sciences*  
401 **117**, 17591-17598, doi:10.1073/pnas.1921231117 (2020).
- 402 32 Breen, K. H., Barahona, D., Yuan, T., Bian, H. & James, S. C. Effect of volcanic emissions  
403 on clouds during the 2008 and 2018 Kilauea degassing events. *Atmos. Chem. Phys.* **21**, 7749-  
404 7771, doi:10.5194/acp-21-7749-2021 (2021).
- 405 33 Glassmeier, F. *et al.* Aerosol-cloud-climate cooling overestimated by ship-track data. *Science*  
406 **371**, 485-489, doi:10.1126/science.abd3980 (2021).
- 407 34 Christensen, M. W. *et al.* Opportunistic experiments to constrain aerosol effective radiative  
408 forcing. *Atmos. Chem. Phys.* **22**, 641-674, doi:10.5194/acp-22-641-2022 (2022).
- 409 35 Bender, F. A. M., Frey, L., McCoy, D. T., Grosvenor, D. P. & Mohrmann, J. K. Assessment  
410 of aerosol–cloud–radiation correlations in satellite observations, climate models and  
411 reanalysis. *Climate Dynamics* **52**, 4371-4392, doi:10.1007/s00382-018-4384-z (2019).
- 412 36 Fuchs, J., Cermak, J. & Andersen, H. Building a cloud in the southeast Atlantic:  
413 understanding low-cloud controls based on satellite observations with machine learning.  
414 *Atmos. Chem. Phys.* **18**, 16537-16552, doi:10.5194/acp-18-16537-2018 (2018).
- 415 37 Oreopoulos, L., Cho, N. & Lee, D. A Global Survey of Apparent Aerosol-Cloud Interaction  
416 Signals. *Journal of Geophysical Research: Atmospheres* **125**, e2019JD031287,  
417 doi:<https://doi.org/10.1029/2019JD031287> (2020).
- 418 38 Fan, J., Wang, Y., Rosenfeld, D. & Liu, X. Review of Aerosol–Cloud Interactions:  
419 Mechanisms, Significance, and Challenges. *Journal of the Atmospheric Sciences* **73**, 4221-  
420 4252, doi:10.1175/jas-d-16-0037.1 (2016).
- 421 39 Gettelman, A., Schmidt, A. & Egill Kristjánsson, J. Icelandic volcanic emissions and climate.  
422 *Nature Geoscience* **8**, 243-243, doi:10.1038/ngeo2376 (2015).



423 40 Oreopoulos, L., Cho, N., Lee, D. & Kato, S. Radiative effects of global MODIS cloud  
424 regimes. *Journal of Geophysical Research: Atmospheres* **121**, 2299-2317,  
425 doi:<https://doi.org/10.1002/2015JD024502> (2016).  
426 41 IPCC. Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent  
427 Treatment of Uncertainties.  
428 [https://www.ipcc.ch/site/assets/uploads/2017/2008/AR2015\\_Uncertainty\\_Guidance\\_Note.pdf](https://www.ipcc.ch/site/assets/uploads/2017/2008/AR2015_Uncertainty_Guidance_Note.pdf)  
429 (2010).  
430 42 Grist, J. P. *et al.* Extreme air–sea interaction over the North Atlantic subpolar gyre during the  
431 winter of 2013–2014 and its sub-surface legacy. *Climate Dynamics* **46**, 4027-4045,  
432 doi:[10.1007/s00382-015-2819-3](https://doi.org/10.1007/s00382-015-2819-3) (2016).  
433 43 Grosvenor, D. P. *et al.* Remote Sensing of Droplet Number Concentration in Warm Clouds: A  
434 Review of the Current State of Knowledge and Perspectives. *Reviews of Geophysics* **56**, 409-  
435 453, doi:<https://doi.org/10.1029/2017RG000593> (2018).  
436

437

438

## 439 **Methods**

### 440 **MODIS observations**

441 We used the Level-3 products of MODIS Collection 6.1, i.e., MYD08 for 2002-2020 from  
442 Aqua and MOD08 for 2001-2020 from Terra. The reported retrieval bias due to instrument  
443 degradation in Terra-MODIS Collection 5.1 datasets<sup>15</sup> has been rectified in Collection 6.1. An  
444 inadvertent artifact in the calculations of cloud fraction (derived from cloud optical property)  
445 in Collection 5.1 has also been removed in Collection 6.1<sup>44</sup> and both Terra-MODIS and Aqua-  
446 MODIS now show consistent results<sup>45,46</sup>. Cloud droplet effective radius ( $r_{\text{eff}}$ ), in-cloud liquid  
447 water path (LWP), cloud optical thickness and cloud phase are retrieved from observed  
448 radiances using a radiative transfer model at 1-km nadir resolution in Level-2 products and  
449 aggregated to the  $1^\circ \times 1^\circ$  Level-3 products<sup>46</sup>. The Level-3 Cloud Optical Property Cloud  
450 Fraction product for the liquid phase (dataset name Cloud Retrieval Fraction Liquid)<sup>46</sup> is used  
451 in the cloud fraction (CF) analysis, because this CF product can distinguish between clouds of  
452 liquid and ice phase and is consistent with the other microphysical retrievals of cloud properties  
453 used in this study. Note that ref. <sup>17</sup> used MODIS Collection 6 data, and used cloud fraction  
454 derived from the cloud mask<sup>47</sup> multiplied by the fractional liquid cloud and found a more  
455 modest increase in cloud fraction of  $\sim 1.7\%$  in October. Differences between our findings and  
456 those from ref. <sup>17</sup> in the climatological analysis likely arise from a combination of the use of  
457 different CF products, the extension of the MODIS data to include data from 2015-2020, and  
458 differences in the areas of investigation. Monthly-mean products are used in this study, with  
459 differences being negligible when aggregating Level-3 daily products into monthly means<sup>15,32</sup>.  
460 An exception is liquid cloud droplet number concentration ( $N_d$ ) which is derived from  $r_{\text{eff}}$  and  
461 cloud optical thickness assuming adiabatic conditions<sup>8,20,31,48</sup>, and because of non-linear  
462 dependences,  $N_d$  is first obtained daily and then averaged to monthly means<sup>48,49</sup>. Only pixels  
463 with  $r_{\text{eff}}$  between  $4 \mu\text{m}$  to  $30 \mu\text{m}$  and cloud optical thickness between 4 to 70 are used for the

464 most reliable  $N_d$  retrievals<sup>49</sup>. The uncertainty of the derived  $N_d$  is discussed in detail in  
465 Grosvenor et al.<sup>43</sup> who estimated that the uncertainty can be largely reduced to about 50% when  
466 averaged over  $1^\circ \times 1^\circ$  regions. The uncertainty is expected to be even smaller in our study, since  
467 we average across a geographical region of about  $3000 \text{ km} \times 6000 \text{ km}$ .  
468 To further back up our finding of increased CF, we also analysed the frequency of cloud-free  
469 conditions in arguably the most stringent MODIS product, namely pixels with retrieved aerosol  
470 optical depth (AOD) at 550 nm which are used as a proxy of cloud-free pixels. This pixels are  
471 most stringent because any thin or sub-grid scale cloud is screened out to prevent contamination  
472 of AOD retrievals. Level-3 monthly MODIS AOD products record the number of validated 1-  
473  $\text{km}^2$  pixels used in the Level-2 products when performing aggregation. These statistics are used  
474 to calculate the relative reduction of cloud-free pixels in our region in October 2014 relative to  
475 the long-term 2001-2020 October mean excluding 2014. While the number of pixels with AOD  
476 retrievals do not have a one-to-one correspondence to the number of cloud-free pixels because  
477 factors such as sun-glint in cloud-free pixels can reduce the number of AOD pixels, it is still a  
478 good relative (rather than absolute) proxy for cloud-free pixels.

479

## 480 **Meteorological reanalyses**

481 Meteorological reanalyses represent the best estimate of global atmospheric conditions<sup>50</sup>, and  
482 are available from the European Centre for Medium-Range Weather Forecasts ERA5 products  
483 (<https://cds.climate.copernicus.eu/>). To train the machine-learning surrogate MODIS (ML-  
484 MODIS), we use the monthly averaged ERA5 reanalysis from the surface up to 550 hPa level  
485 at  $0.25^\circ \times 0.25^\circ$  horizontal resolution and 50 hPa vertical resolution. This vertical range covers  
486 most of the low-level liquid clouds. In total, 114 meteorological parameters are re-gridded to  
487 MODIS grid cells and used in the training, details of which are provided in Supplementary  
488 Table S1. The ERA5 monthly reanalysis products at 11:00 and 13:00 Icelandic time (same as

489 UTC) are closest to the daytime Terra and Aqua overpass times and are paired with the  
490 respective MODIS products from these satellites for the training.

491

## 492 **Machine-learning**

493 Previous studies that use machine-learning to investigate the statistical correlation between  
494 cloud properties and aerosol (e.g., ref. <sup>36,51</sup>) can possibly be affected by confounding  
495 meteorological co-variability that would prevent confirmation of the causal processes of  
496 aerosol-cloud interactions (ACI)<sup>33,36</sup>. Here, we use a random forest algorithm<sup>52</sup> to train a ML-  
497 MODIS that diagnoses cloud properties for given meteorological conditions but unperturbed  
498 by volcanic aerosol. This allows comparisons of cloud properties between conditions with and  
499 without volcanic aerosol-perturbation but otherwise alike, therefore quantifying cloud  
500 responses only to volcanic aerosol, i.e. signals of ACI. Note that this machine-learning  
501 approach is not designed to calculate the temporal evolution of cloud properties and cannot  
502 predict the development of meteorological systems. The latter is obtained from the ERA5  
503 reanalysis, which provides the best estimate of atmospheric state<sup>50</sup>.

504 The random forest algorithm is chosen because of its excellent performance in dealing with  
505 relatively small sample sizes and high-dimensional feature spaces and in avoiding over-  
506 fitting<sup>52,53</sup>. Random forest based machine-learning has been successfully applied to isolate the  
507 confounding meteorological variability in air quality assessments and has been shown to  
508 perform much better than multinomial regression models<sup>54-56</sup>. A regression mode forest of one  
509 hundred trees is trained independently for each cloud property ( $N_d$ ,  $r_{eff}$ , LWP and CF) and for  
510 each month (October and September), with a minimal leaf size of seven for each tree without  
511 merge leaves. Each tree samples ~60% of the input data with replacement for the training data  
512 and the remaining data is used as out-of-bag observations. With larger forests, we find a  
513 negligible reduction in out-of-bag mean squared error and a negligible increase in out-of-bag

514 coefficient of determination (a more informative estimate of performance than mean squared  
515 error<sup>57</sup>) of up to 0.87 for CF prediction. This indicates a good stability and avoidance of over-  
516 fitting<sup>58</sup>. The number of randomly selected predictors is 38 (one third of the total number of  
517 features) and the interactive-curvature method is used to select split predictors. The ERA5  
518 meteorological reanalysis is independent of the MODIS datasets, which are not assimilated in  
519 the reanalysis<sup>50</sup>, and provides the explanatory variables in the ML-MODIS training. The  
520 dependent variables are the corresponding cloud properties observed by MODIS with no  
521 volcanic eruption. The successful training of ML-MODIS is enabled by the large MODIS  
522 dataset from continuous observations over the past 20 years on two satellite platforms. We  
523 employ the “out-of-bag permuted predictor delta error” method<sup>52,59</sup> to measure the importance  
524 of each explanatory feature in predicting cloud properties. The results for CF shown in  
525 Extended Data Fig. 6.

526 The performance of ML-MODIS as a surrogate of the MODIS observations under conditions  
527 without the volcanic perturbation is evaluated using “leave-one-year-out” cross validation<sup>60</sup> for  
528 each cloud property, as shown in the left panels of Fig. 1 and Extended Data Fig. 7. This  
529 involves training ML-MODIS using randomly selected sets of 18 years of ERA5-MODIS  
530 dataset pairs and then evaluating ML-MODIS against the remaining 19<sup>th</sup> year of MODIS  
531 observations. This evaluation is carried out for each non-eruption year during 2001-2020. The  
532 uncertainty of ML-MODIS is further estimated using a Monte Carlo method, and the variability  
533 of the reference baselines are shown as black boxplots in Fig. 3 and Extended Data Fig. 8a. For  
534 the Monte Carlo uncertainty estimate, we randomly perform “leave-one-year-out” validation  
535 500 times for each cloud property, by excluding both Terra and Aqua datasets of the randomly  
536 selected year over the entire region from machine-learning training but use them for validation.  
537 A test for  $N_d$  using the validation of a 700-member Monte Carlo ensemble showed negligible  
538 differences. The ratios of cloud properties between the ML-MODIS prediction (without

539 volcano-perturbation) and MODIS observations in 2014 (with volcano-perturbation) are in  
 540 pink in Fig. 3 for October and in Extended Data Fig. 8a for September, with the pink boxplots  
 541 showing the variability of all decision-trees within the random forest Monte Carlo ensembles,  
 542 i.e., the variability of the ACI signals.

543

## 544 Radiative forcing

545 We estimate the relative contributions from the Twomey  $r_{\text{eff}}$ , LWP adjustment, and CF  
 546 adjustment to ACI-induced radiative forcing using the susceptibilities of  $r_{\text{eff}}$ , LWP and CF to  
 547  $N_d$  perturbations. The radiative forcing arising from cloud albedo brightening can be described  
 548 as Eq. (1) at a constant CF<sup>8,9,61</sup>, and the forcing arising from CF enhancement can be described  
 549 as Eq. (2) at a constant cloud albedo  $\alpha_{\text{cld}}$ .

$$\left. \frac{dSW_{\text{TOA}}}{d \ln AOD} \right|_{CF} = -SW_{\text{down}} \times CF \times \alpha_{\text{cld}} \times (1 - \alpha_{\text{cld}}) \times \frac{d \ln N_d}{d \ln AOD} \times \left( \frac{1}{3} + \frac{5}{6} \frac{d \ln LWP}{d \ln N_d} \right) \quad (1)$$

550

$$\left. \frac{dSW_{\text{TOA}}}{d \ln AOD} \right|_{\text{albedo}} = -SW_{\text{down}} \times (\alpha_{\text{cld}} - \alpha_{\text{cs}}) \times \frac{dCF}{d \ln AOD} = -SW_{\text{down}} \times (\alpha_{\text{cld}} - \alpha_{\text{cs}}) \times CF \times \frac{d \ln CF}{d \ln N_d} \times \frac{d \ln N_d}{d \ln AOD} \quad (2)$$

551 where  $dSW_{\text{TOA}}$  is the short-wave radiative forcing at the top of atmosphere,  $SW_{\text{down}}$  is the  
 552 incoming short-wave solar radiation at the top of the atmosphere,  $\alpha_{\text{cld}}$  is the average broadband  
 553 short-wave cloud albedo with a global mean of 0.38 for liquid clouds<sup>9</sup>, and  $\alpha_{\text{cs}}$  is clear-sky  
 554 broadband ocean surface albedo which is about 0.07 for representative of global average (solar  
 555 zenith angle of 60 degrees)<sup>62</sup>. The total ACI-induced short-wave radiative forcing is the sum  
 556 of Eq. (1) and Eq. (2), as shown in Eq. (3).

$$\begin{aligned} \frac{dSW_{\text{TOA}}}{d \ln AOD} &= \left. \frac{dSW_{\text{TOA}}}{d \ln AOD} \right|_{\text{albedo}} + \left. \frac{dSW_{\text{TOA}}}{d \ln AOD} \right|_{CF} \\ &= -SW_{\text{down}} \times CF \times \frac{d \ln N_d}{d \ln AOD} \times \left[ \frac{1}{3} \alpha_{\text{cld}} (1 - \alpha_{\text{cld}}) + \alpha_{\text{cld}} (1 - \alpha_{\text{cld}}) \times \frac{5}{6} \frac{d \ln LWP}{d \ln N_d} + (\alpha_{\text{cld}} - \alpha_{\text{cs}}) \frac{d \ln CF}{d \ln N_d} \right] \end{aligned} \quad (3)$$

557 The radiative forcing contributions from the Twomey  $r_{\text{eff}}$  effect, LWP adjustment and CF  
558 adjustment are described as the three terms in the square bracket from left to right, respectively.

559

560 **Data availability:** The MODIS cloud and aerosol products from Aqua (MYD08\_L3) and Terra  
561 (MOD08\_L3) used in this study are available from the Atmosphere Archive and Distribution System  
562 Distributed Active Archive Center of National Aeronautics and Space Administration (LAADS-DAAC,  
563 NASA), <https://ladsweb.modaps.eosdis.nasa.gov>. ERA5 datasets are available from the European Centre for  
564 Medium-range Weather Forecast (ECMWF) archive, <https://cds.climate.copernicus.eu>. The full datasets  
565 shown in the figures are provided in source data files.

566 **Code availability:** Code is available from the corresponding author on reasonable request.

567

#### 568 **References for Methods and Supplementary Information:**

- 569 44 Platnick, A. S. *et al.* MODIS Cloud Optical Properties: User Guide for the Collection 6/6.1  
570 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets. [https://atmosphere-](https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/MODISCloudOpticalPropertyUserGuideFinal_v1.1_1.pdf)  
571 [imager.gsfc.nasa.gov/sites/default/files/ModAtmo/MODISCloudOpticalPropertyUserGuideFi-](https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/MODISCloudOpticalPropertyUserGuideFinal_v1.1_1.pdf)  
572 [nal\\_v1.1\\_1.pdf](https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/MODISCloudOpticalPropertyUserGuideFinal_v1.1_1.pdf) (2018).
- 573 45 Platnick, S. *et al.* The MODIS Cloud Optical and Microphysical Products: Collection 6  
574 Updates and Examples From Terra and Aqua. *IEEE Transactions on Geoscience and Remote*  
575 *Sensing* **55**, 502-525, doi:10.1109/TGRS.2016.2610522 (2017).
- 576 46 Hubanks, P., Platnick, A. S., King, M. & Ridgway, B. MODIS Atmosphere L3 Gridded  
577 Product Algorithm Theoretical Basis Document (ATBD) & Users Guide. *available from:*  
578 [https://icdc.cen.uni-](https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/icdc_Dokumente/MODIS/MODIS_Collection6_AtmosphereL3_GriddedProduct_ATBDandUsersGuide_v4.1_Sep22_2015.pdf)  
579 [hamburg.de/fileadmin/user\\_upload/icdc\\_Dokumente/MODIS/MODIS\\_Collection6\\_Atmosphe-](https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/icdc_Dokumente/MODIS/MODIS_Collection6_AtmosphereL3_GriddedProduct_ATBDandUsersGuide_v4.1_Sep22_2015.pdf)  
580 [reL3\\_GriddedProduct\\_ATBDandUsersGuide\\_v4.1\\_Sep22\\_2015.pdf](https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/icdc_Dokumente/MODIS/MODIS_Collection6_AtmosphereL3_GriddedProduct_ATBDandUsersGuide_v4.1_Sep22_2015.pdf) (2019).
- 581 47 Maddux, B. C., Ackerman, S. A. & Platnick, S. Viewing Geometry Dependencies in MODIS  
582 Cloud Products. *Journal of Atmospheric and Oceanic Technology* **27**, 1519-1528,  
583 doi:10.1175/2010jtecha1432.1 (2010).
- 584 48 Quaas, J., Boucher, O., Bellouin, N. & Kinne, S. Satellite-based estimate of the direct and  
585 indirect aerosol climate forcing. *Journal of Geophysical Research: Atmospheres* **113**,  
586 doi:<https://doi.org/10.1029/2007JD008962> (2008).
- 587 49 Quaas, J., Boucher, O. & Lohmann, U. Constraining the total aerosol indirect effect in the  
588 LMDZ and ECHAM4 GCMs using MODIS satellite data. *Atmos. Chem. Phys.* **6**, 947-955,  
589 doi:10.5194/acp-6-947-2006 (2006).
- 590 50 Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal*  
591 *Meteorological Society* **146**, 1999-2049, doi:<https://doi.org/10.1002/qj.3803> (2020).
- 592 51 Andersen, H., Cermak, J., Fuchs, J., Knutti, R. & Lohmann, U. Understanding the drivers of  
593 marine liquid-water cloud occurrence and properties with global observations using neural  
594 networks. *Atmos. Chem. Phys.* **17**, 9535-9546, doi:10.5194/acp-17-9535-2017 (2017).
- 595 52 Breiman, L. Random Forests. *Machine Learning* **45**, 5-32, doi:10.1023/A:1010933404324  
596 (2001).
- 597 53 Biau, G. & Scornet, E. A random forest guided tour. *TEST* **25**, 197-227, doi:10.1007/s11749-  
598 016-0481-7 (2016).

- 599 54 Grange, S. K., Carslaw, D. C., Lewis, A. C., Boleti, E. & Hueglin, C. Random forest  
600 meteorological normalisation models for Swiss PM10 trend analysis. *Atmos. Chem. Phys.* **18**,  
601 6223-6239, doi:10.5194/acp-18-6223-2018 (2018).
- 602 55 Shi, Z. *et al.* Abrupt but smaller than expected changes in surface air quality attributable to  
603 COVID-19 lockdowns. *Science Advances* **7**, eabd6696, doi:10.1126/sciadv.abd6696 (2021).
- 604 56 Yang, J. *et al.* From COVID-19 to future electrification: Assessing traffic impacts on air  
605 quality by a machine-learning model. *Proceedings of the National Academy of Sciences* **118**,  
606 e2102705118, doi:10.1073/pnas.2102705118 (2021).
- 607 57 Chicco, D., Warrens, M. J. & Jurman, G. The coefficient of determination R-squared is more  
608 informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation.  
609 *PeerJ Computer Science* **7**, e623, doi:10.7717/peerj-cs.623 (2021).
- 610 58 Cutler, A., Cutler, D. R. & Stevens, J. R. Random Forests. In: Zhang C., Ma Y. (eds)  
611 Ensemble Machine Learning. . Springer, Boston, MA. , doi:https://doi.org/10.1007/978-1-  
612 4419-9326-7\_5 (2012).
- 613 59 Bonte, S., Goethals, I. & Holen, R. V. in *2017 IEEE Nuclear Science Symposium and Medical  
614 Imaging Conference (NSS/MIC)*. 1-3.
- 615 60 Bastos, L. S. & O'Hagan, A. Diagnostics for Gaussian Process Emulators. *Technometrics* **51**,  
616 425-438, doi:10.1198/TECH.2009.08019 (2009).
- 617 61 Ackerman, A. S. *et al.* Effects of Aerosols on Cloud Albedo: Evaluation of Twomey's  
618 Parameterization of Cloud Susceptibility Using Measurements of Ship Tracks. *Journal of the  
619 Atmospheric Sciences* **57**, 2684-2695, doi:10.1175/1520-  
620 0469(2000)057<2684:eoaoa>2.0.co;2 (2000).
- 621 62 Jin, Z., Charlock, T. P., Smith Jr., W. L. & Rutledge, K. A parameterization of ocean surface  
622 albedo. *Geophysical Research Letters* **31**, doi:https://doi.org/10.1029/2004GL021180 (2004).
- 623 63 Twigg, M. M. *et al.* Impacts of the 2014–2015 Holuhraun eruption on the UK atmosphere.  
624 *Atmos. Chem. Phys.* **16**, 11415-11431, doi:10.5194/acp-16-11415-2016 (2016).
- 625 64 Myers, T. A. & Norris, J. R. On the Relationships between Subtropical Clouds and  
626 Meteorology in Observations and CMIP3 and CMIP5 Models. *Journal of Climate* **28**, 2945-  
627 2967, doi:10.1175/jcli-d-14-00475.1 (2015).
- 628 65 Höjgård-Olsen, E. *Observations of the tropical atmospheric water cycle and its variations  
629 with sea surface temperature using a constellation of satellites Observations du cycle de l'eau  
630 atmosphérique tropicale et de ses variations avec la température de surface de la mer, à  
631 l'aide d'une constellation de satellites*, Université Paris-Saclay, (2020).
- 632 66 Costa, A. A., Cotton, W. R., Walko, R. L., Pielke, R. A. & Jiang, H. SST Sensitivities in  
633 Multiday TOGA COARE Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences*  
634 **58**, 253-268, doi:10.1175/1520-0469(2001)058<0253:Ssimtc>2.0.Co;2 (2001).
- 635 67 Sandu, I., Stevens, B. & Pincus, R. On the transitions in marine boundary layer cloudiness.  
636 *Atmos. Chem. Phys.* **10**, 2377-2391, doi:10.5194/acp-10-2377-2010 (2010).

637

638

639

640

641

642

643

644

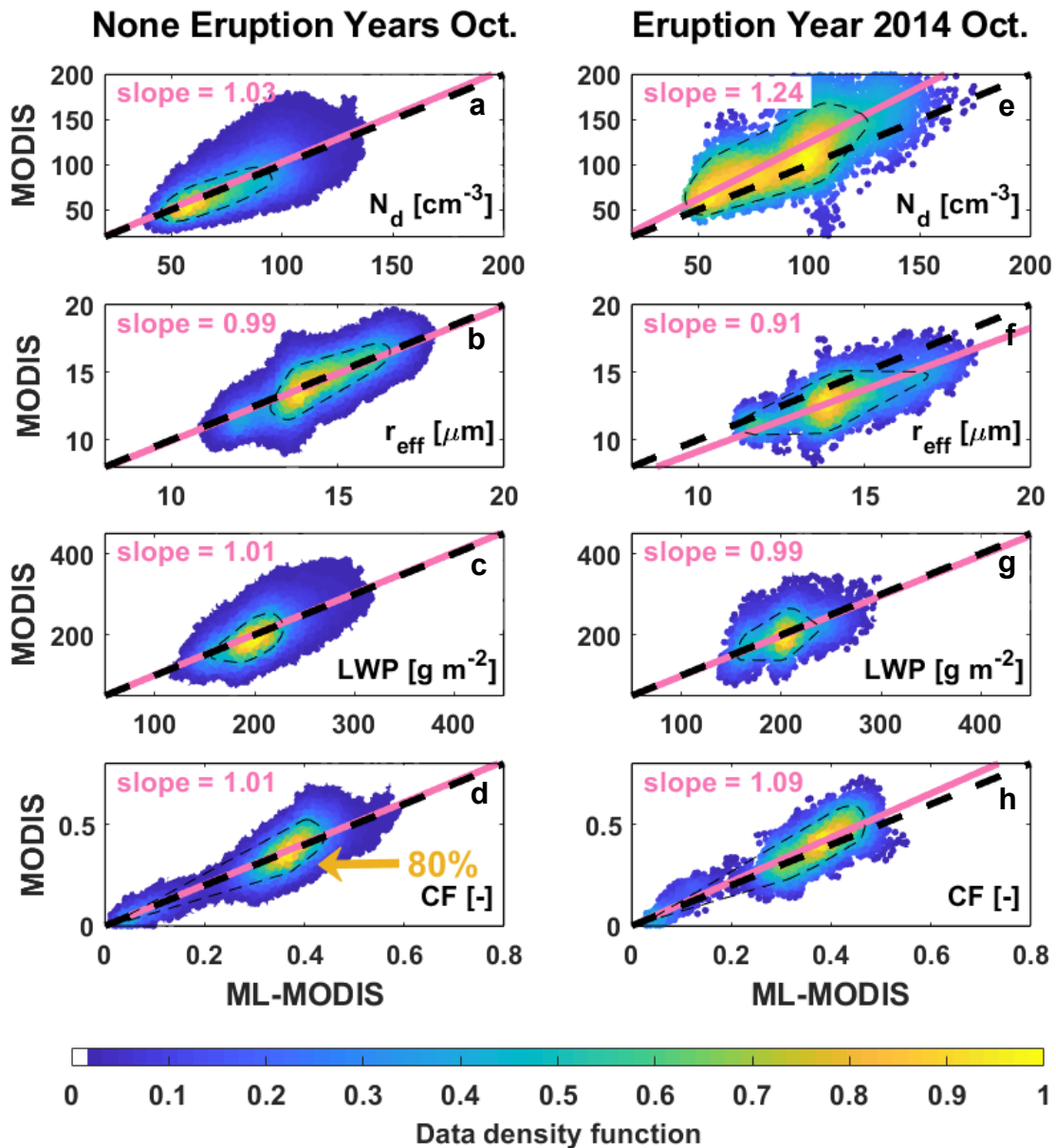
645

646



647

648



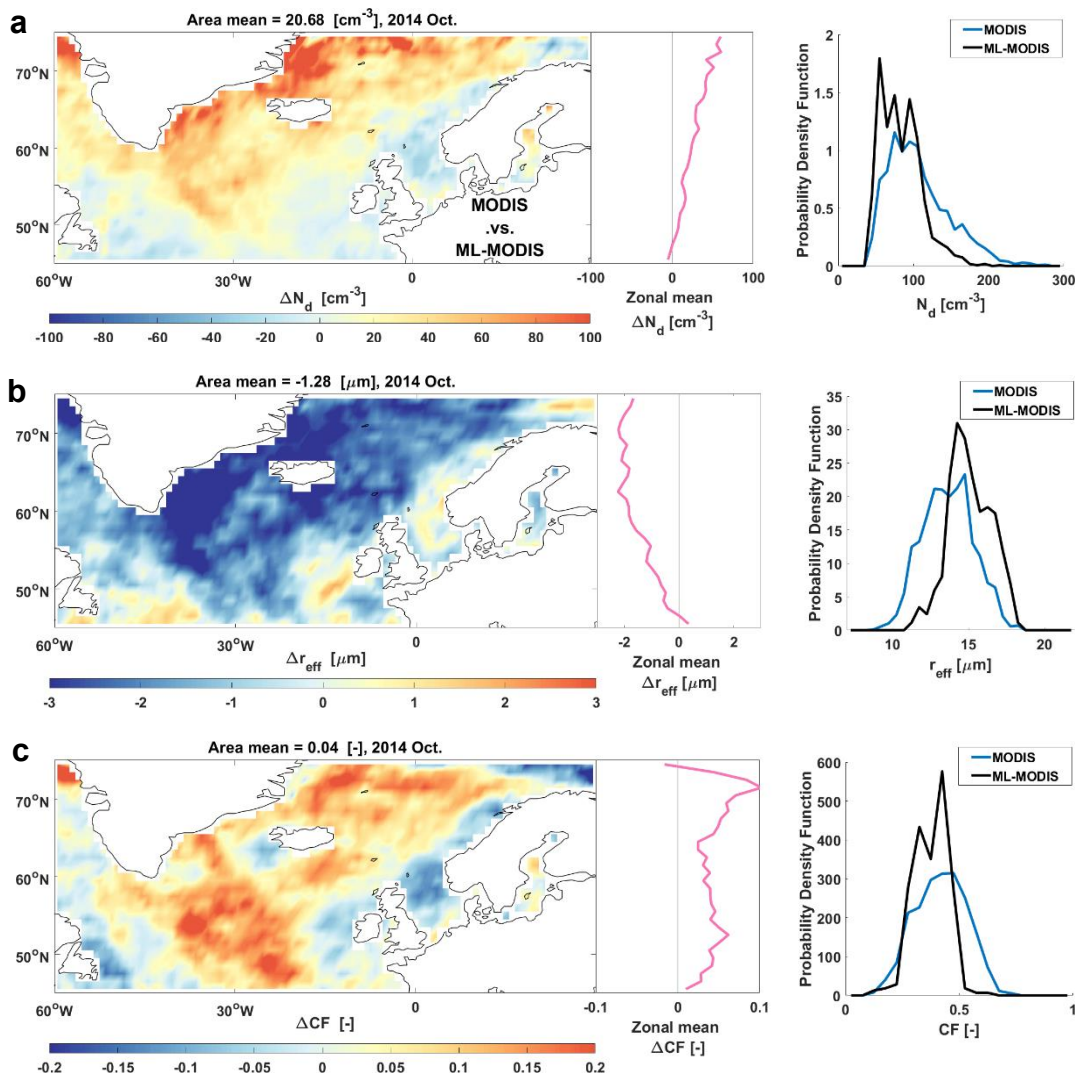
**Fig. 1 | Comparison between machine-learning predictions (ML-MODIS) and MODIS observations.** Left panels (a-d): validation against non-perturbed observations (excluding 2014) of cloud properties, from top to bottom they show cloud droplet number concentration ( $N_d$ ), cloud droplet effective radius ( $r_{eff}$ ), cloud liquid water path (LWP) and cloud fraction (CF). Right panels (e-h): volcanic perturbation signals in October 2014, indicated by the difference between the machine-learning predictions and the observations. October MODIS observations from Aqua (2002-2020) and Terra (2001-2020) are analyzed. Colour indicates the normalized data density function with a maximum value of one, with 80% of the data being contained within the black dashed area.

649

650

651

652



**Fig. 2 | Changes in cloud properties caused by the volcanic perturbation** estimated using machine-learning predictions and MODIS observations for October 2014. The spatial distribution and zonal means of the changes in  $N_d$ ,  $r_{eff}$  and CF are shown in the left panels of **a-c** while right panels show probability density functions (so that the areas under the curves are equivalent) for MODIS and ML-MODIS.

654

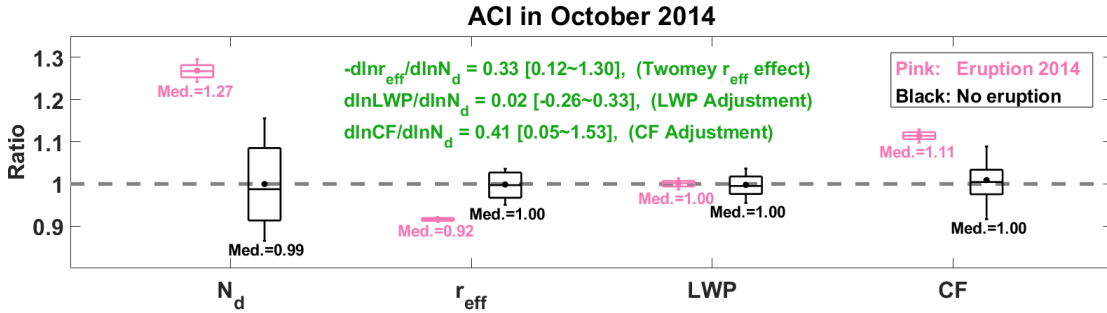
655

656

657

658

659



**Fig. 3 | Responses of cloud properties to the volcanic aerosol-perturbation in October 2014.** The aerosol-cloud interactions (ACI) signals of responses are indicated as the ratios between MODIS (Aqua and Terra) observations and machine-learning predictions, i.e., Ratio = MODIS divided by ML-MODIS. Uncertainties of non-perturbed baseline references are estimated using a Monte Carlo method and are shown in black (see Methods, based on non-volcanic October datasets spanning 2001-2020). The variability of the cloud responses to the Holuhraun volcanic aerosol perturbation are shown in pink. The boxplots show 10<sup>th</sup>, 25<sup>th</sup>, median (Med.), 75<sup>th</sup> and 90<sup>th</sup> percentiles with the mean value indicated by a dot. The susceptibilities of  $r_{eff}$ , LWP and CF to changes in  $N_d$  are given in a green colour, median [90% confidence interval]. Area (in units of km<sup>2</sup>) weighted averaging is used to calculate average cloud properties over the geographical region (Fig. 2), in order to estimate an unbiased large-scale response signal. Therefore, the ratios shown here are slightly different from the slopes shown in Fig. 1, in which area-weighted averaging is not applied.

660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678

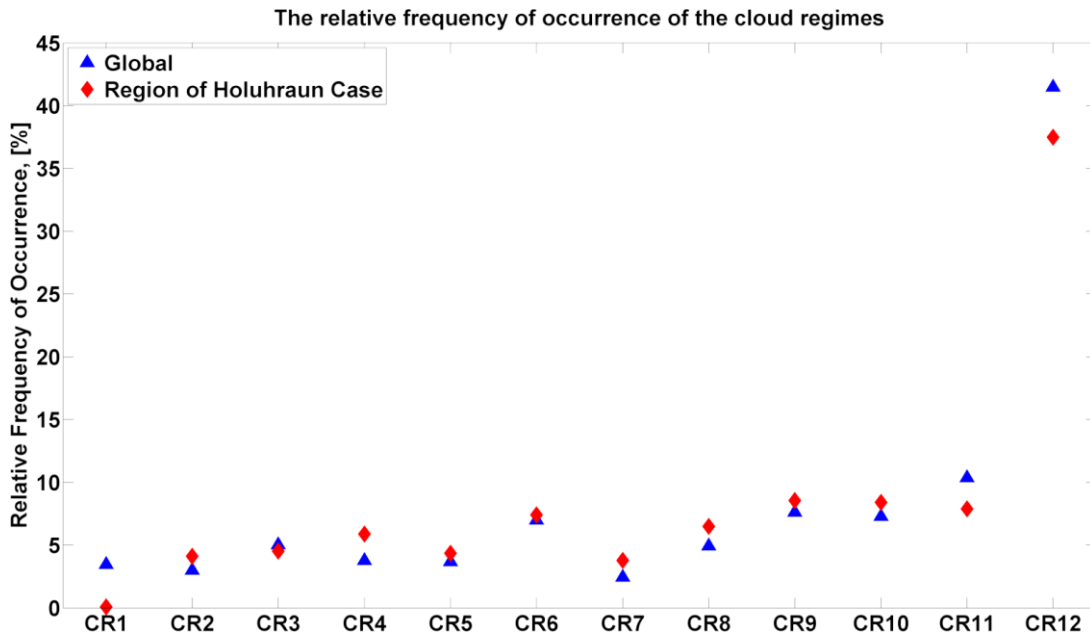
679

## 680 Extended Data Figures.

681

682

683

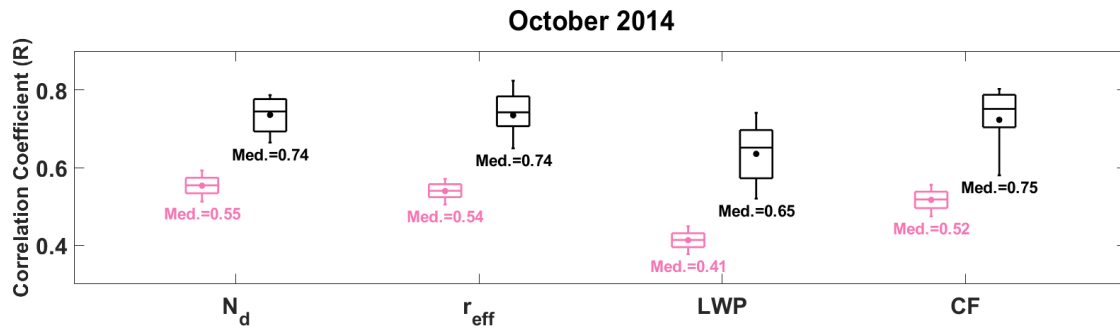


**Extended Data Fig. 1 | Relative frequency of occurrence (RFO) of cloud regimes.** The RFO values of the region studied here in September-October 2014 are given in red diamonds, data sourced from Malavelle et al.<sup>15</sup>. The RFO values during 2002-2014 globally are given in blue triangles, data sourced from Oreopoulos et al.<sup>40</sup>. CR6-CR11 are liquid-dominated cloud regimes, and the others are ice-dominated cloud regimes. The details of each cloud regime are given in the above references accordingly.

684

685

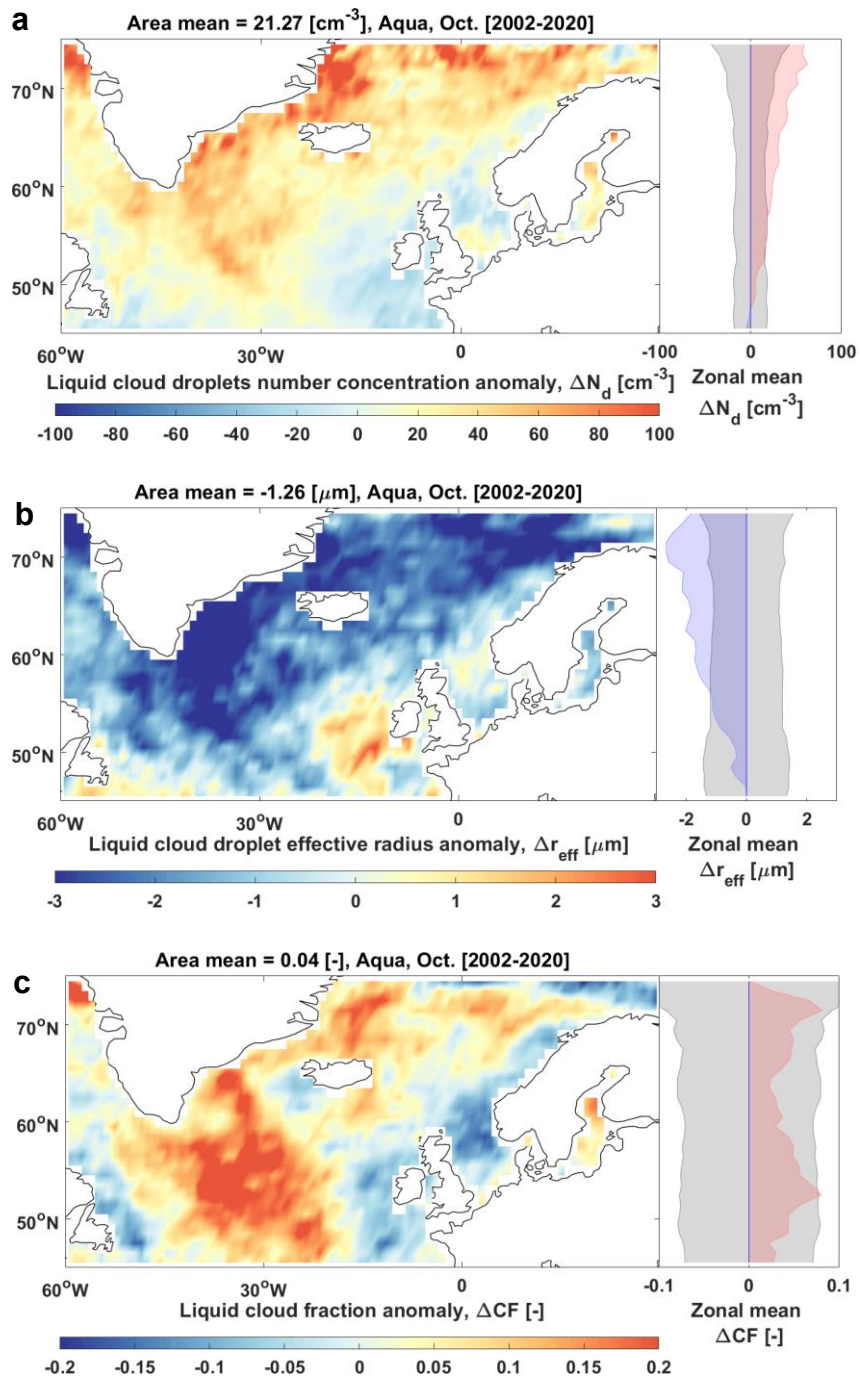
686



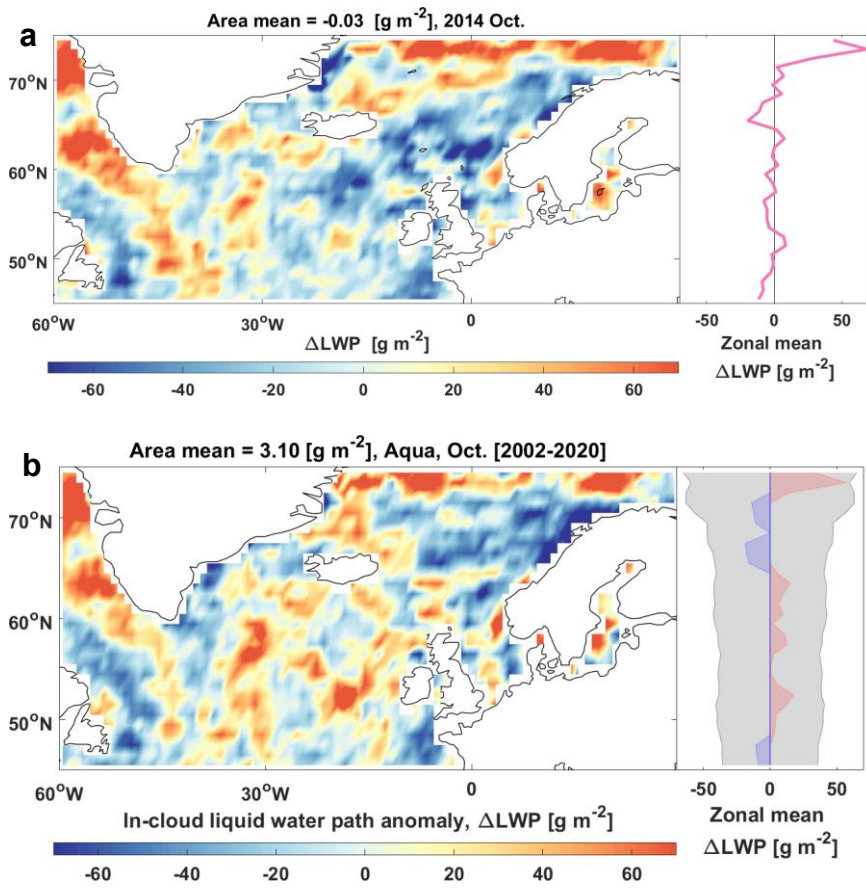
**Extended Data Fig. 2 | Correlation coefficient between machine-learning predictions and MODIS observations** of cloud properties, including liquid cloud droplet number concentration ( $N_d$ ), liquid droplet effective radius ( $r_{eff}$ ), liquid water path (LWP) and liquid cloud fraction (CF). The Monte Carlo results of ML-MODIS validation against MODIS observations without volcanic aerosol-perturbation are given in black. The variations of comparisons with volcanic aerosol-perturbation in October 2014 are given in pink. The boxplot shows 10<sup>th</sup>, 25<sup>th</sup>, median (Med.), 75<sup>th</sup> and 90<sup>th</sup> percentiles with the mean value indicated by a dot.

687

688

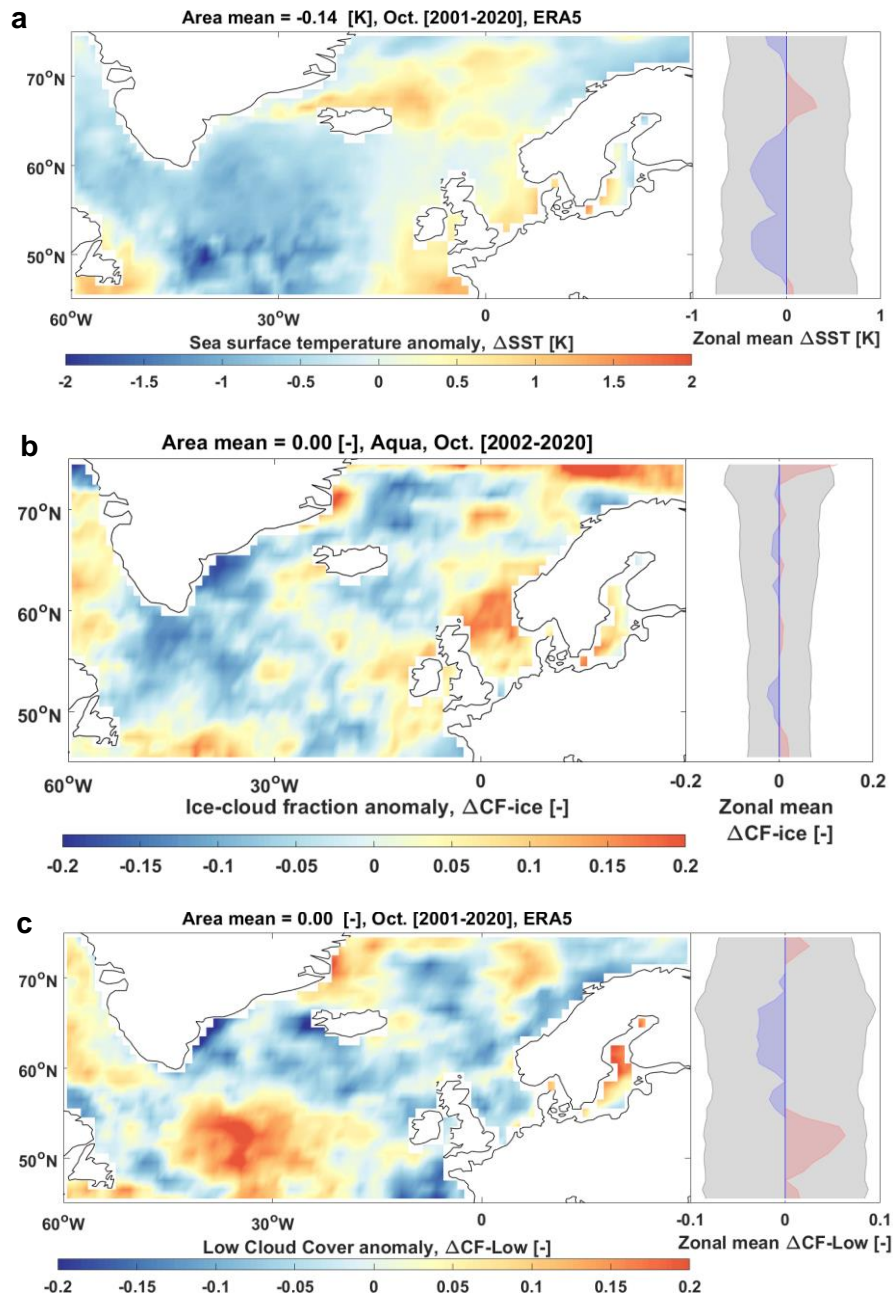


**Extended Data Fig. 3 | Anomalies in MODIS-Aqua cloud properties for October 2014.** The spatial distributions and zonal means of anomalies in  $N_d$ ,  $r_{eff}$  and CF are shown in the panels a-c. Anomalies correspond to the deviation from the 2002-2020 climatology (excluding the 2014 eruption year). The positive anomalies are shown in red and negative ones in blue. The standard deviation is shown by the grey shading.



**Extended Data Fig. 4 | Change (a) and anomaly (b) in liquid water path (LWP).** a) Similar to Fig. 2, changes are detected using machine-learning; b) similar to Extended Data Fig. 3, anomaly corresponds to the deviation from 2002-2020 climatology.

692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703

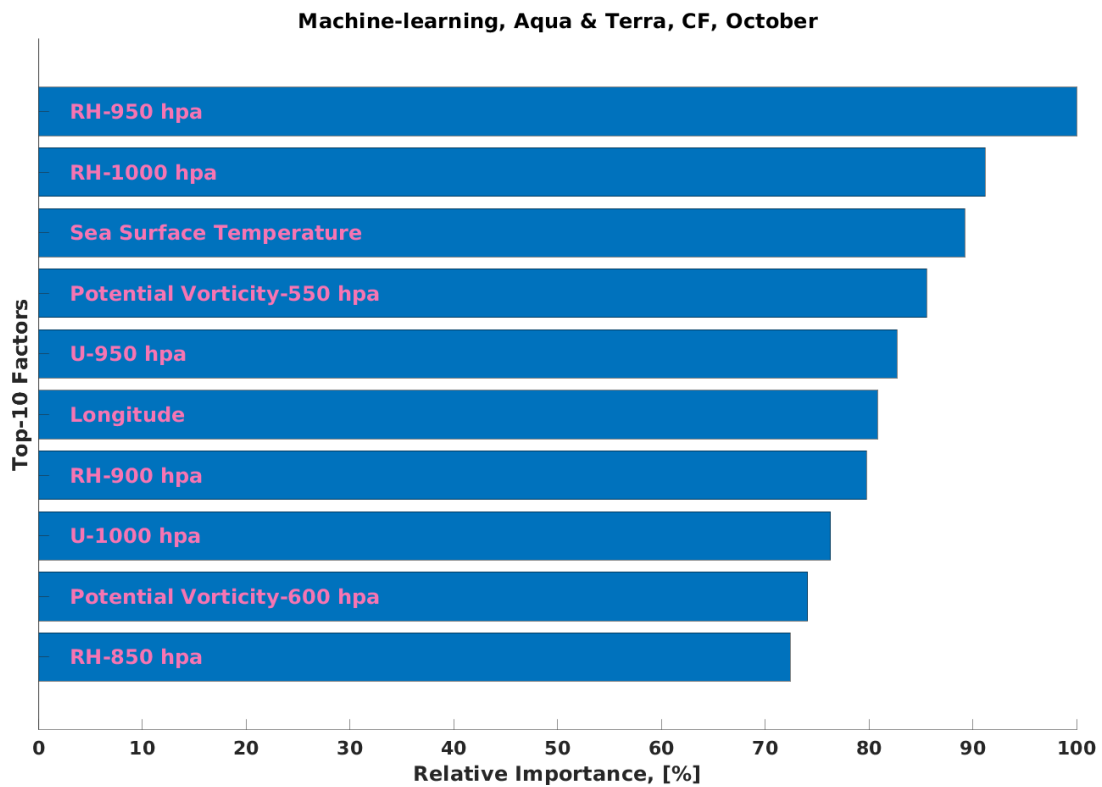


**Extended Data Fig. 5** | Similar to Extended Data Fig. 3, but show anomaly in sea-surface temperature (a), anomaly in ice-cloud fraction in October 2014 (b), and climatological anomaly of low-level cloud cover in October 2014 using ERA5 reanalysis (c).

704  
705  
706  
707  
708  
709  
710

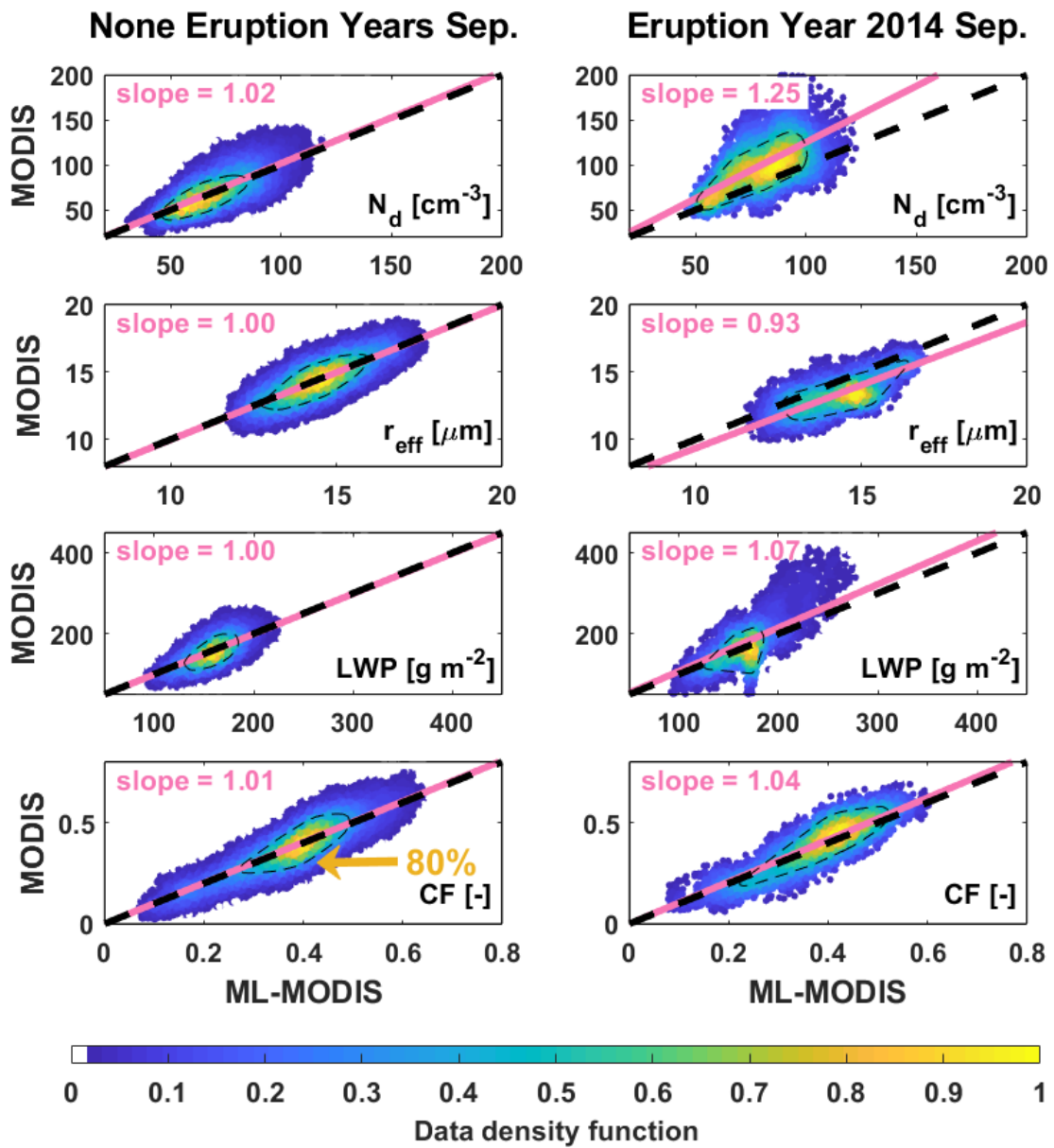


711  
712  
713  
714  
715  
716



**Extended Data Fig. 6 | The top-10 most important features for machine-learning to predict unperturbed liquid cloud fraction in October.** The feature importance is normalized with the maximum as 100%. The value of these features in 2014 are entirely within the variation range of machine-learning training dataset, see Extended Data Fig. 10.

717  
718  
719  
720



Extended Data Fig. 7 | Similar to Fig. 1, but show results in September 2014.

722

723

724

725

726

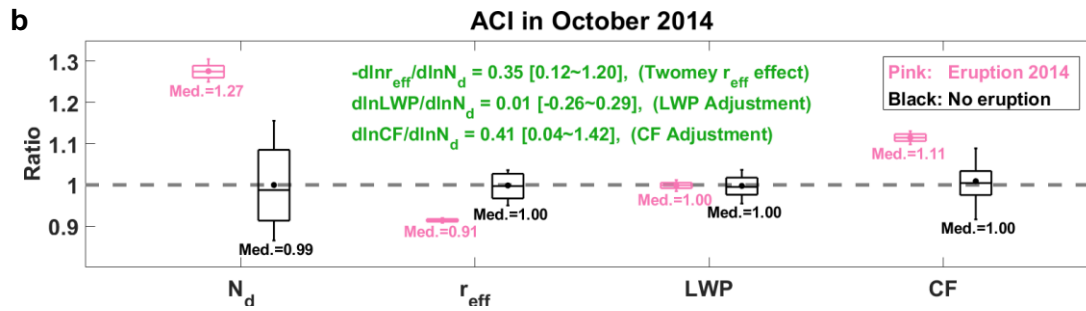
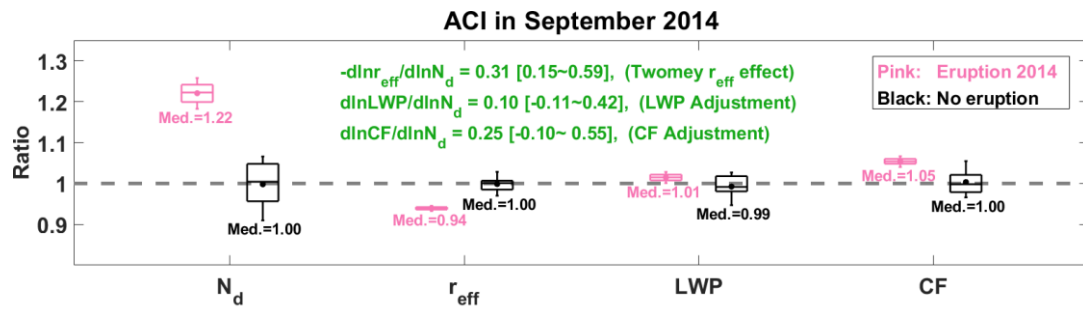
727

728

729

730

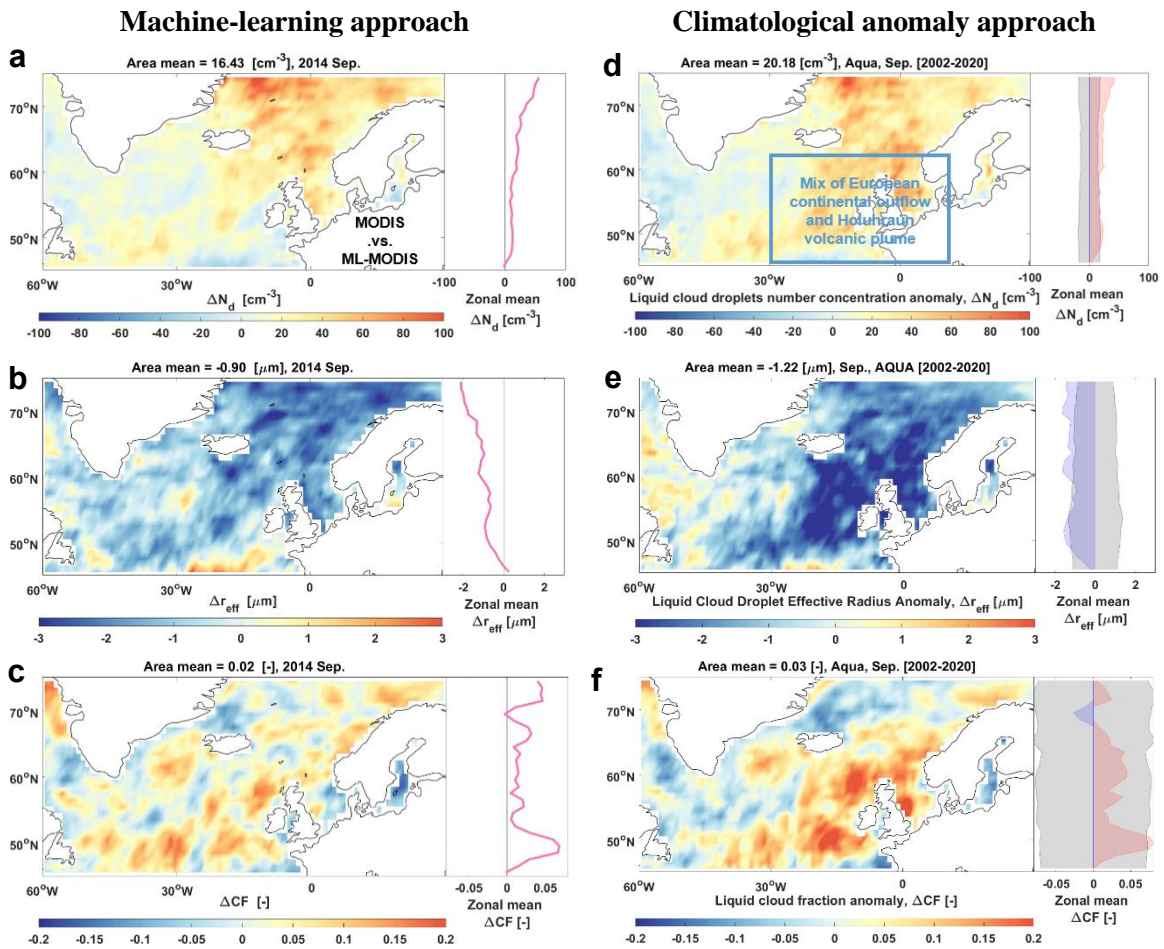
a



**Extended Data Fig. 8 | Similar to Fig. 3. Panel a** shows results in September 2014. **Panel b** shows results in October 2014 but excluding the regions where the cold anomalous SSTs were outside the variation range at the same location.

731

732



**Extended Data Fig. 9 | Cloud responses to Holuhraun volcanic aerosol in September 2014.** Left panels a-c (similar to Fig. 2 but for September 2014) show cloud responses to volcanic aerosol using machine-learning (ML) approach. Right panels d-f (similar to Extended Data Fig. 3 but for September 2014) show anomalies in cloud properties.

734

735

736

737

738

739

740

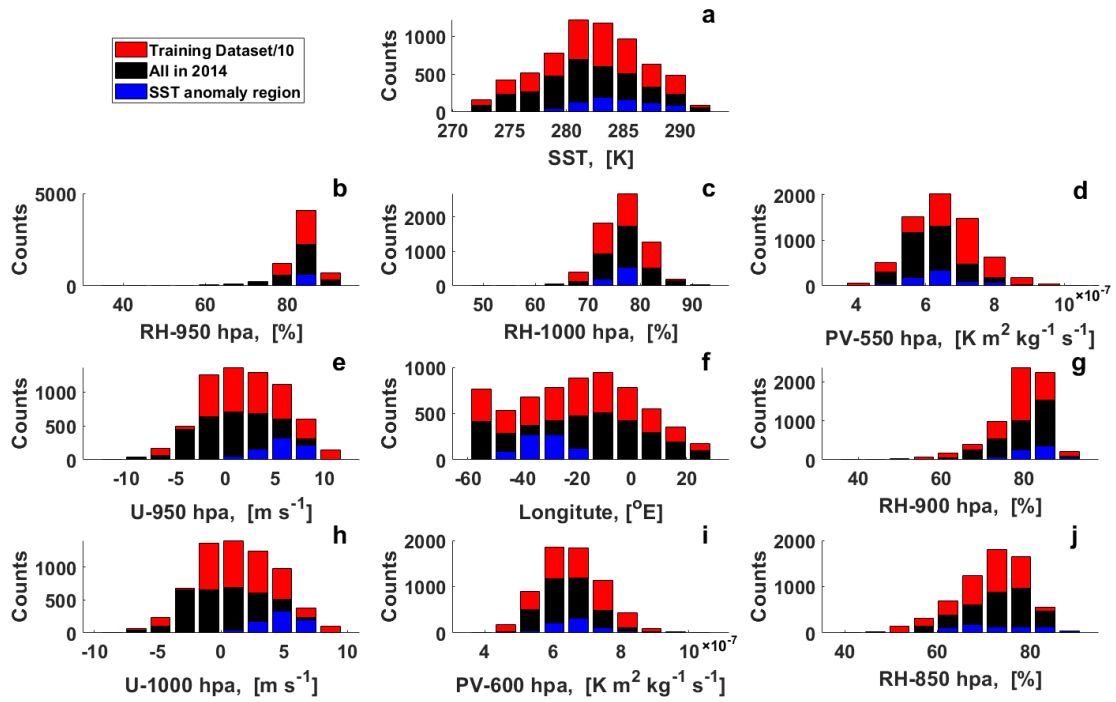
741

742

743

744

745



**Extended Data Fig. 10 | Probability distribution of the top-10 most important features**, as shown in Extended Data Fig. 6. Red bars indicate the counts (scaled by 0.1 to fit the display range) of the training data in each bin, which covers the entire variability range of black and blue bars; black bars indicate the data counts from the entire studied region in October 2014; and blue bars indicate the counts from the SST anomaly region only. Note that the counts per longitude are different, because we only consider data over the oceans.

746

747

748