Radiative forcing from the 2020 shipping fuel regulation is large but hard to detect

Check for updates

Jianhao Zhang [1,2] ✉, Yao-Sheng Chen [1,2], Edward Gryspeerdt [3], Takanobu Yamaguchi [1,2] & Graham Feingold [2]

Reduction in aerosol cooling unmasks greenhouse gas warming, exacerbating the rate of future warming. The strict sulfur regulation on shipping fuel implemented in 2020 (IMO2020) presents an opportunity to assess the potential impacts of such emission regulations and the detectability of deliberate aerosol perturbations for climate intervention. Here we employ machine learning to capture cloud natural variability and estimate a radiative forcing of $+0.074 \pm 0.005$ W m^{-2} related to IMO2020 associated with changes in shortwave cloud radiative effect over three low-cloud regions where shipping routes prevail. We find low detectability of the cloud radiative effect of this event, attributed to strong natural variability in cloud albedo and cloud cover. Regionally, detectability is higher for the southeastern Atlantic stratocumulus deck. These results raise concerns that future reductions in aerosol emissions will accelerate warming and that proposed deliberate aerosol perturbations such as marine cloud brightening will need to be substantial in order to overcome the low detectability.

Marine low-level, warm clouds exert a strong net cooling effect on the climate[1]. As a result of their responses to increases in aerosol over the twentieth century, these climate-cooling clouds have been a key player in mitigating the rate at which the Earth warms in response to increasing greenhouse gases (GHG) through aerosol-cloud interactions (ACI)[2]. Although future aerosol mitigation policies will have important health benefits, they may exacerbate global warming as a result of the unmasking of GHG warming due to reduced aerosol radiative forcing[3–5]. An increase in aerosol loading leads to an increase in the population of smaller cloud droplets, all else equal, by ramping up available cloud condensation nuclei (CCN). This increases the total surface area of the droplets that scatters sunlight, resulting in more reflective clouds[6]. Knock-on effects such as precipitation suppression, enhanced entrainment-evaporation, or short-wave heating might augment or offset this brightening[7–11]. Ship tracks—bright, linear features found over subtropical shipping corridors where marine low clouds prevail[12,13]—are a visible manifestation of these aerosol effects on clouds.

At the beginning of 2020, the world's maritime shipping sector implemented the International Maritime Organization(IMO) Marine Environment Protection Committee (MEPC) fuel regulation that limits the maximum fuel sulfur mass concentration in ships from 3.5% to 0.5% or requires an equivalent sulfur oxide reduction in the exhaust using gas cleaning systems. This regulation is hereafter abbreviated as IMO2020[14]. The combustion of sulfur-containing fuel produces sulfate aerosol—an

effective source of CCN for marine boundary layer clouds—as a byproduct or via oxidation of the emitted gaseous sulfur dioxide. This regulation leads to a reduction in sulfate aerosol production over global oceans[15], resulting in lower cloud droplet number concentrations (N_d) that have been detected regionally[16–18]. Since IMO2020, a pronounced reduction in ship track occurrence has been shown globally[15,16], adding to documented impacts from earlier fuel regulations designated for specific emission zones near the Californian and European coasts[19]. Furthermore, in the southeast Atlantic shipping corridor, recent studies present robust evidence for a reduction in cloud microphysical perturbations inside the shipping lane, along with a weakened although less robust cloud brightening[17].

In spite of the routine nature of shipping traffic, the occurrence of visible ship tracks is only a few percent at the global scale[15,16], in part because of variable meteorological conditions and in part because of changes in the background aerosol. Therefore, whether IMO2020 exerts a detectable perturbation to the cloud radiative effect after 3 years, taking into account the co-varying meteorological conditions that also modify the cloud field, is an open question and remains to be assessed. Previous studies suggest that (i) at least 5 to 6 years of observations are needed to derive significant signals in both N_d and cloud albedo changes from the shipping lane over the southeast Atlantic[20], and (ii) in a hypothetical geoengineering setting, substantial, sustained increases in shortwave (SW) cloud albedo are required to surpass the detection limit (or noise) set by the natural variability in cloud albedo[21]. While these studies establish foundations for the topic of ACI detectability in

[1]Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, CO, USA. [2]Chemical Sciences Laboratory, National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, USA. [3]Grantham Institute - Climate Change and the Environment, Imperial College London, London, UK. ✉e-mail: jianhao.zhang@noaa.gov

the climate system (a signal-to-noise ratio problem), they each focus on one side of the problem, i.e., signal and noise, respectively. When the detectability of a real-world event, such as IMO2020, is assessed, one needs to connect both sides.

In this study, we aim to assess (i) signal significance—the strength of the IMO2020 signal, due to aerosol changes alone, all else equal, by comparing the observed cloud field against a counterfactual cloud field constructed by an ensemble of neural networks, and (ii) signal detection relative to natural variability—whether the post-perturbation observational record stands out above the variability in the cloud radiative effect. The latter evaluation includes all drivers of cloud variability, not only anthropogenic but also natural, by comparing the observational records before and after IMO2020. The combination of the two assessments informs us about the general detectability of the IMO2020 event and has implications for the detectability of marine cloud brightening (MCB), which if adopted, would deliberately inject aerosol into clouds[22–24].

Here, we focus on the relative cloud radiative effect (rCRE) of low clouds, a quantity of high relevance to radiative forcing, that incorporates both cloud brightness (cloud albedo) and areal coverage (cloud fraction; see Methods). We use a data-driven approach[25] to quantify the radiative forcing of IMO2020 due to changes in the SW cloud radiative effect from the three main marine low-cloud regions where shipping prevails, at a spatiotemporal scale beyond that of individual ship tracks. This approach utilizes an ensemble of neural networks, trained to capture the climatological relationship between large-scale meteorological conditions and cloud properties, to construct counterfactual cloud fields as if IMO2020 were not implemented (i.e., business-as-usual). The ensemble approach not only boosts the robustness of the forcing estimate, but also enables uncertainty and significance evaluations. Despite an estimated radiative forcing of $+0.074 \pm 0.005$ W m^{-2}, we find low detectability of the cloud radiative effect of this event, attributed to strong cloud natural variability.

Results

A data-driven approach to perturbation assessment
It has been shown that cloud properties and their radiative effect can be robustly predicted by combinations of large-scale meteorological variables (or cloud controlling factors) using data-driven approaches[25–30]. In particular, the ability of data-driven approaches to capture the climatological, non-linear relationship among meteorology, aerosol conditions, and cloud macrophysical properties is central to this work[27,28], enabling a counterfactual continuation of this climatological, non-linear relationship after 2020 under the assumption that it can be used to represent the business-as-usual emission scenario. We train Neural Network Ensembles (NNEs) to map 8 large-scale meteorological conditions to low-cloud properties, including rCRE and N_d, on a gridded (1°) and monthly scale using satellite observations and reanalysis datasets (see Methods). In this work, the three major stratocumulus decks, namely the Northeastern Pacific (NEP), the Southeastern Pacific (SEP), and the Southeastern Atlantic (SEA) stratocumulus decks that coincide with major shipping corridors are investigated for 3 years since IMO2020 (2020–2022, inclusive).

The task we assign to the Neural Network (NN) is to learn and capture the complex, non-linear co-variability among large-scale meteorological conditions, aerosol conditions, and low-cloud properties from the climatological data record (2003–2018) for the stratocumulus regime. Applying such climatological co-variability to the IMO2020 impacted years (2020–2022) is equivalent to intentionally ignoring the implementation of IMO2020—an abrupt change in the aerosol field, while keeping everything else the same, especially the large-scale meteorological conditions. This creates a counterfactual scenario from which the impact of IMO2020 and its detectability can be assessed.

Inevitably for any machine-learning (ML) approach, constraining uncertainty within the ML-prediction is a challenge—in our case to provide a robust radiative forcing estimate. The skill of the NN and the robustness of the predictions are central. To this end, we train an ensemble of NN instead of a single NN to be able to assess the uncertainty associated with our

particular NN framework. The performance of these NNEs for the 4 target variables, namely rCRE, N_d, low-cloud fraction (LCF), and low-cloud albedo (A_c), is well above community standards (Supplementary Fig. 1), and is not subject to overfitting (Supplementary Fig. 2). Furthermore, a linear correction to the NN model is applied to account for cases where high spatial correlations in cloud responses to the aerosol perturbation from ships are evident, e.g., a shipping lane (see Methods). This step is particularly important for robustly predicting the N_d field over the SEA, where meteorological factors alone are less skilled at capturing the sharp gradient in the climatological N_d distribution across a spatially narrow lane because the main shipping route is aligned with the prevailing boundary layer winds[20]. Another issue that NN-predictions are known to be prone to is extrapolation. We examine the ranges of the 8 meteorological conditions during 2020-2022 relative to the ranges of the training dataset (2003–2018) and conclude that the probability of extrapolation in our case is minimal, based on the overlapping distributions of the predictors between the training and the predicting datasets (Supplementary Fig. 3). Known expected responses, if they do exist, can serve as a sanity check for the behavior of the NN. Among all the shipping corridors, the one in the southeastern Atlantic presents such an opportunity, given the aligned wind direction and the shipping route. This makes cloud micro- and macro-properties within the shipping lane easily distinguishable from the surrounding areas[17,20]. A lane-shaped decrease in N_d due to IMO2020 is indeed evident in the SEA region of this study, aligning well with the main shipping lane (Fig. 1a), providing assurance of the skill of our approach to assess the IMO2020 perturbation.

Substantial radiative forcing from changes in cloudiness and cloud brightness
Changes in N_d (ΔN_d, the difference between observed and counterfactual values, see Methods) during 2020–2022 are negative for most parts of the three stratocumulus decks (Fig. 1); these are more pronounced near the coasts, consistent with the strong reduction in fuel sulfur content since IMO2020 and the dense shipping traffic in coastal regions[15,16]. Most of the decreases in N_d are statistically significant at the grid level against the NNE-predicted counterfactual values (black dots in Fig. 1; based on the Wilcoxon signed-rank test at 95% confidence interval, see Methods). Over the SEA, a band of stronger negative ΔN_d that is easily discernible from the neighboring grids aligns well with the main shipping lane therein. The fact that N_d decreases both inside and adjacent to the shipping lane are statistically significant suggests a dilution of the shipping emission at the monthly timescale. Regionally, the NEP (SEP) has the most (least) substantial decrease in N_d (Fig. 1a, white labels), consistent with the fact that the NEP (SEP) region has the highest (lowest) sulfur dioxide emission from ships and the most frequent occurrence of ship tracks[16]. The SEA presents the most discernible spatial pattern in N_d changes, including not only the linear feature of a shipping lane but also the significant increases in N_d closer to the equator, which can be attributed, in part, to the 2019–2020 Australian bushfires (positive ΔN_d near the equator diminishes when 2020–2021 are excluded from the analysis; not shown). Although, in the mean (spatiotemporally), regional decreases in N_d are less than 5 cm^{-3} (Fig. 2b), the percentage decrease can be as large as ~6% (e.g., over the NEP; Supplementary Fig. 4) and local decreases can reach as high as ~30 cm^{-3} near the coast (Figs. 1a, 2b).

Decreases in rCRE, corresponding to more incoming solar radiation reaching the surface (i.e., a warming effect), are observed for most parts of the stratocumulus regime (Fig. 1b). Substantial and significant warming is concentrated in the northwestern part of the NEP stratocumulus deck. Over the SEP, significant warming is mostly concentrated downwind of the main stratocumulus deck, despite little change in N_d, suggesting a strong cloud susceptibility to aerosol perturbations. This occurs in the broken cumulus regime and is likely related to the precipitation-suppression mechanism[7]. The lane-shaped feature observed in the spatial pattern of ΔN_d is evident in ΔrCRE over the SEA, although to a lesser extent, but nevertheless indicates a persistent warming along the main shipping lane since IMO2020. The
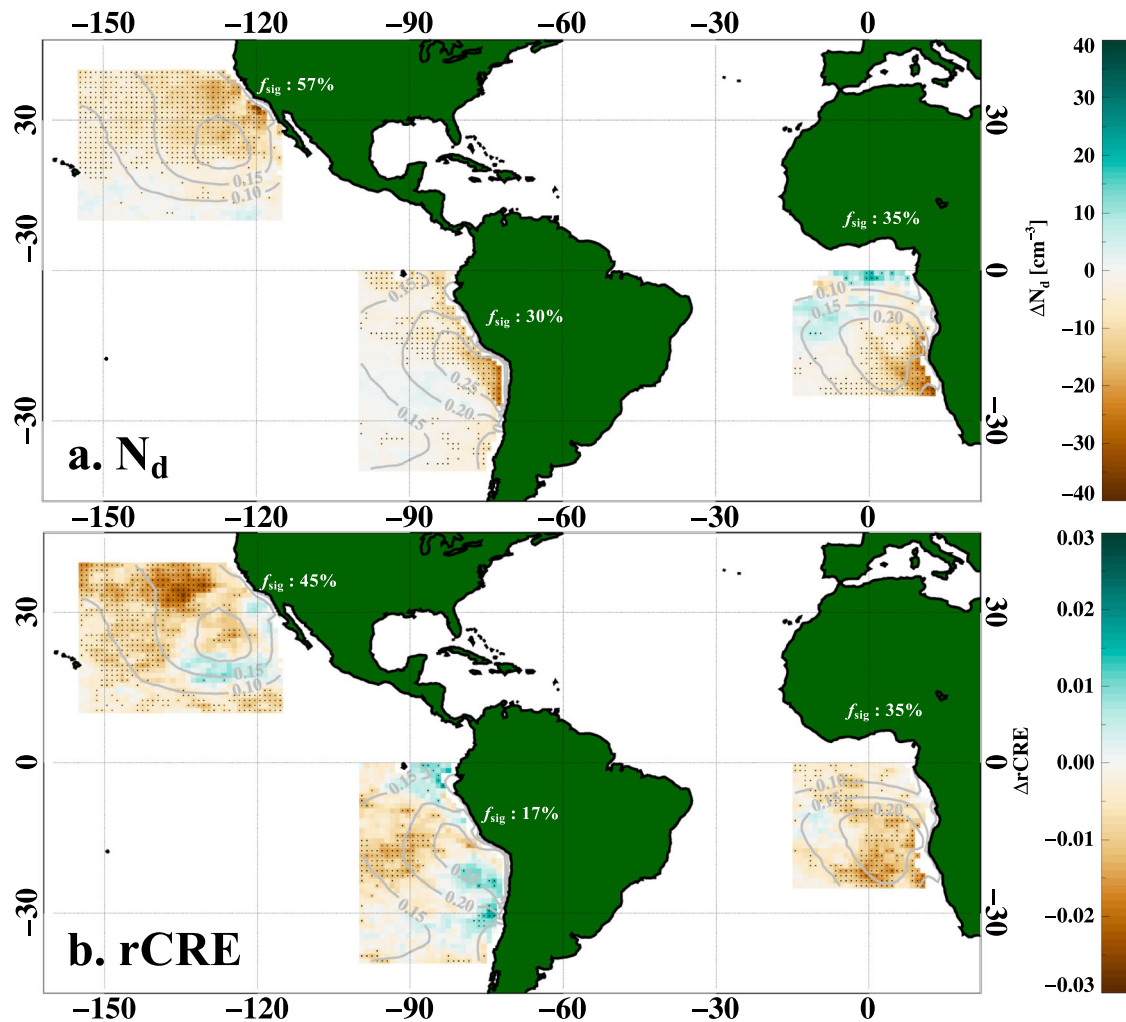
Fig. 1 | IMO2020 perturbations of cloud properties. 3-year mean change in (**a**) cloud droplet number concentration (N_d) and (**b**) relative cloud radiative effect (rCRE) during 2020–2022 for the three stratocumulus decks. rCRE climatology (2003-2018 mean) is indicated by gray contours. Black dots denote grid points for which the IMO2020 aerosol-induced (all else equal) perturbation (i.e., OBS_{IMO} versus NNE_{IMO}) is statistically significant according to the Wilcoxon signed-rank test at 95% confidence interval. The fraction of significant (f_{sig}) grids is indicated in white labels.
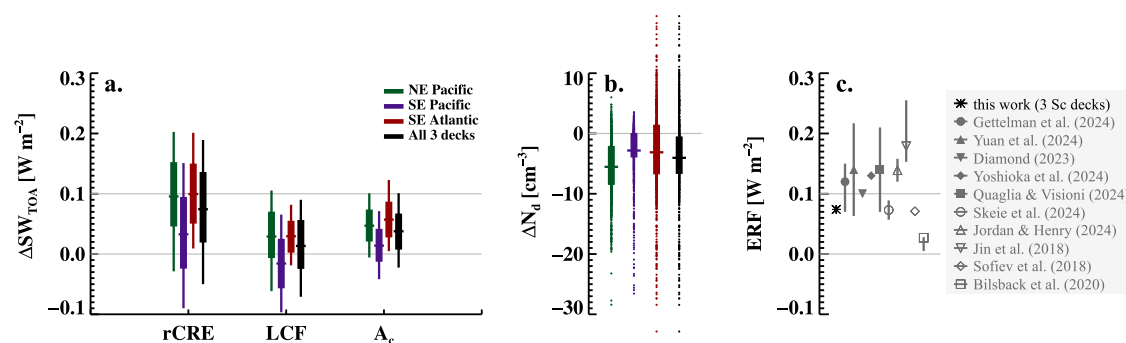

Fig. 2 | IMO2020 radiative forcings. a Global shortwave radiative forcings based on the 3 stratocumulus decks due to changes in low-cloud fraction (LCF), low-cloud albedo (A_c), and low-cloud rCRE (including LCF and A_c adjustments). Positive values indicate gain in incoming solar radiation, i.e., warming. **b** As in (**a**), but for changes in N_d. Box-whiskers indicate the 10th, 25th, 75th, 90th percentiles, and the means of the spatial distribution of the 3-year mean. ΔN_d values outside the 10th and 90th percentiles are indicated by small dots. **c** Effective radiative forcing estimates of the IMO2020 event from literature are compared to this study which reports the forcing due to changes in the SW cloud radiative effect from the 3 stratocumulus decks. Mean values are shown in symbols with uncertainty ranges indicated by vertical bars.

general marine boundary layer flow over these low-cloud regions, i.e., from cooler to warmer waters, dictates a typical transition from stratocumulus clouds (more cloudiness) to cumulus clouds (less cloudiness), which would otherwise be delayed under elevated aerosol loading (through drizzle suppression)[31], consistent with the pronounced changes in the cloud radiative effect, especially via changes in LCF downwind of this transition in each deck (Fig. 1b and Supplementary Fig. 5). The fact that the spatial pattern of ΔrCRE does not match exactly that of ΔN_d also suggests (i) a

spatially varying cloud susceptibility to aerosol perturbations[32] and (ii) a strong dependence of ACI-induced radiative forcing on the temporal co-variation between background aerosol and meteorological conditions that is regionally distinct[33]. On average, our method suggests a positive susceptibility of cloud radiative effect to N_d perturbations of ~0.25 for the stratocumulus regime (see Methods and Supplementary Fig. 4), mostly attributed to an albedo susceptibility of similar magnitude, which is in line with a previous satellite-based assessment[33]. In contrast, the regional LCF susceptibility is weak (medians near zero) due to an offsetting between grids with positive versus negative susceptibilities.

We estimate a global radiative forcing (RF) of +0.074 ±0.005 W m^{-2} (in shortwave radiation at the top-of-atmosphere, ΔSW_{TOA}; see Methods) from the changes in low-cloud shortwave rCRE over the three stratocumulus regions during 2020–2022, scaled by their areal coverage of the globe (Fig. 2). This is equivalent to the effective RF due to ACI (ERF$_{ACI}$)[2,34], including instantaneous RF and cloud adjustments. Although within the range provided by existing model-facilitated assessments of the RF of this event (between +0.027 and +0.179 W m^{-2} in the mean; Fig. 2c)[17,35–43], we note that our estimate only includes contributions from the three main marine low-cloud regions (where shipping routes are prevalent) and does not account for the longwave cloud radiative effect or the direct aerosol radiative effect (which tend to be negligible when low-cloud fraction is high). The small uncertainty range reported here represents the spread of the global mean RF estimates (i.e., after spatial averaging) among the NN ensemble members; this uncertainty is specific to the particular NN framework employed in this work. There are, actually, pronounced spatial variations in our RF estimates, ranging from ~+0.1 W m^{-2} over the NEP and the SEA to +0.03 W m^{-2} over the SEP. The +0.1 W m^{-2} warming over the SEA is in line with the estimate from a recent study that focuses on the shipping corridor in that region[17], consistent with the fact that shipping emissions are strongly constrained within a narrow band in that basin. Variations in ΔN_d alone between basins are not enough to explain the

variations in the basin-specific radiative forcing (e.g., the difference between SEP and SEA; Fig. 2), underscoring the importance of accounting for spatiotemporal co-variations between ΔN_d and cloud susceptibilities, particularly those corresponding to cloud adjustments. On average, changes in A_c contribute two-thirds to our RF estimate, and changes in LCF contribute about one-third (Fig. 2a).

Natural variability masks detectability of the IMO2020-perturbation

A key question in this study is whether the significant IMO2020-perturbations to cloud microphysics (ΔN_d) and radiative effect (ΔrCRE) due to only aerosol are detectable when compared to an a priori observational record, taking into account all natural drivers of cloud variability. Now that we have established the significance of the aerosol-induced signal, we evaluate it against the noise, set by the natural variability in cloud properties, to assess the detectability of this event. Note that the Δ in this study always denotes changes that are attributed to IMO2020 emission changes, all else equal (i.e., the difference between observed and counterfactual values during 2020–2022, see Methods).

Using the Wilcoxon signed-rank test, we evaluate (i) if the observational records before and after IMO2020 are significantly different from each other—whether a change in the observational record is detected (or signal detection against natural variability), and (ii) if the changes in cloud properties due to aerosol emission changes are statistically significant, based on the counterfactual cloud fields constructed by the NNE—significance of the signal (see Methods and Fig. 3). In addition to the detection of significant IMO2020 perturbations, this framework allows us to identify cases where an insignificant IMO2020 perturbation due to emission changes is detected post-2020 in the observational record owing to meteorologically driven changes in the cloud field (i.e., false detection). On the other hand, significant IMO2020 perturbations can be masked by overwhelmingly large natural variability in cloud properties, post-2020, in the observational record (i.e., missed detection). Both these scenarios underscore an important

Fig. 3 | Definition of IMO2020 detectability. A table showing how detectability of the IMO2020 event is characterized in this study. OBS$_{IMO}$ denotes the observational record post-IMO2020, OBS$_{pre-IMO}$ denotes the observational record pre-IMO2020, and NNE$_{IMO}$ denotes the NNE-predicted counterfactual cloud fields based on the 2020–2022 meteorological conditions, representing the business-as-usual scenario, with all else equal except for the aerosol. Color scheme mimics that in Fig. 4.
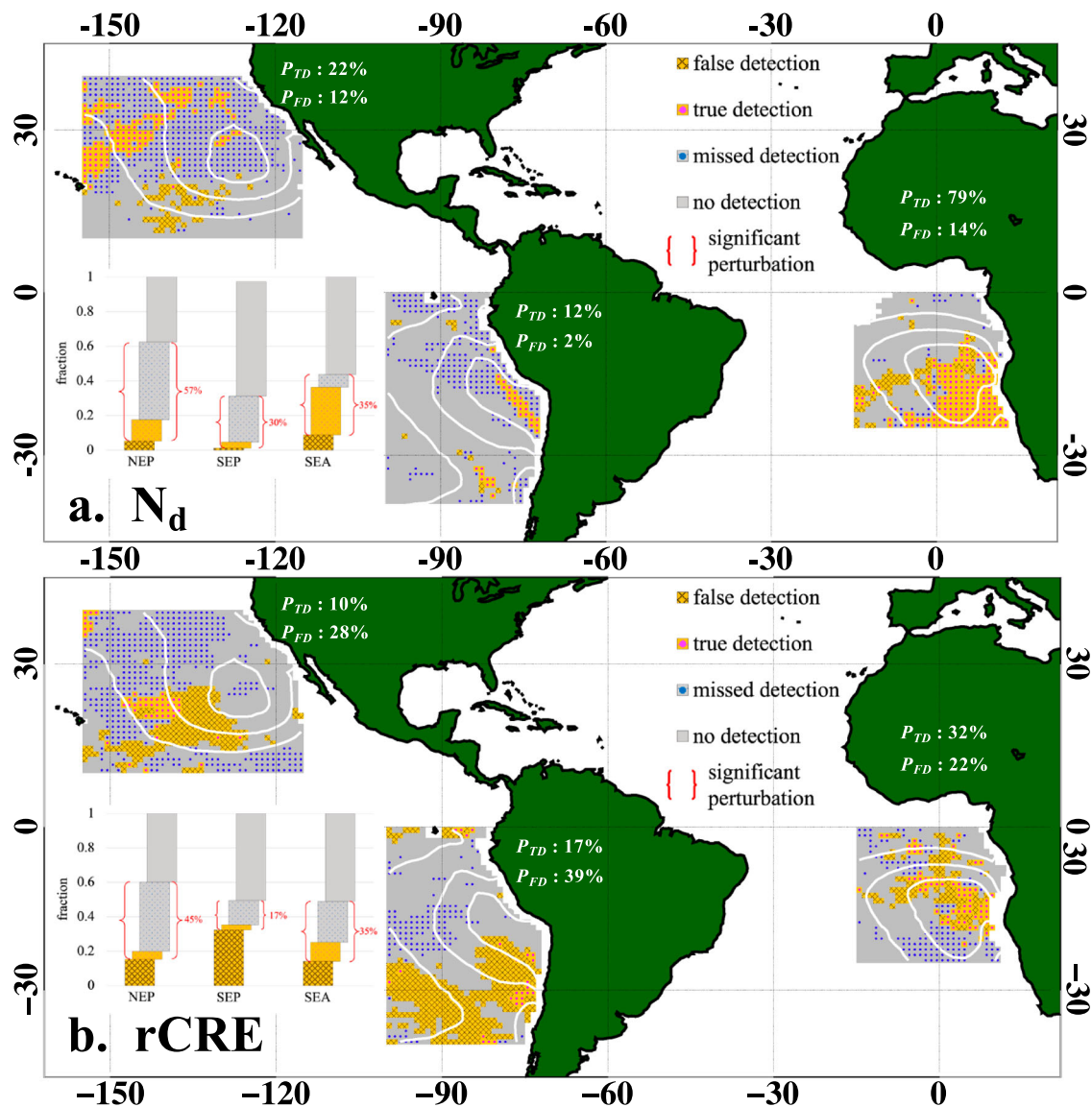
Fig. 4 | Maps of IMO2020 detectability. Detectability of (**a**) N_d and (**b**) rCRE for the IMO2020 event, when using the 2017–2019 observational record as $OBS_{pre-IMO}$, based on the Wilcoxon signed-rank test at 95% confidence interval (see Methods and Fig. 3). Grids where observational records before ($OBS_{pre-IMO}$) and after (OBS_{IMO}) IMO2020 are statistically different are shaded in yellow, including *true detection* (magenta dots on yellow) and *false detection* (crossings on yellow). Grids where OBS_{IMO} is indifferent from $OBS_{pre-IMO}$ are shaded in gray, including *missed*

detection (blue dots on gray) and *no detection* (gray). Inserts indicate the areal fraction of the four detection-scenarios for each region. The fraction of significant IMO2020 (aerosol-induced) perturbation being detected (P_{TD}) and the fraction of insignificant IMO2020 (aerosol-induced) perturbation being detected (P_{FD}) are labeled in white. rCRE climatology (2003–2018 mean) are overlaid with white contours.

confounding effect of natural variability on the detectability of natural or deliberate perturbations to the aerosol field, with the latter being a critical checkpoint for the viability of marine cloud brightening[44,45].

Using the 3-year immediately proceeding 2020 as the reference observational record, we find the true detection rate (P_{TD}, the percentage of significant IMO2020 perturbations being detected in the observational record) for N_d ranges from 12% over the SEP to 79% over the SEA (Fig. 4a). While only 12% of the significant IMO2020 N_d-perturbation is detected over the SEP, mostly in the shipping corridor, the false detection rate (P_{FD}, the percentage of insignificant IMO2020 perturbations being detected) is the lowest (2%) therein. The high P_{TD} over the SEA is consistent with the fact that shipping aerosol emissions are narrowly constrained in space by the general boundary layer flow of the region. Over the NEP and SEP regions, more than half of the significant ΔN_d are undetected (missed detection, gray shading overlaid with blue dots) owing to the large natural variability at these grids. The overall false detection rate for N_d-perturbation

is less than 15% for all three decks, indicating that the climatological meteorology-N_d relationship is less likely to cause a significant departure in the observational record of N_d.

The detectability of ΔrCRE is much lower than that of ΔN_d (Fig. 4b). This is evident in the halving of P_{TD} and the doubling of P_{FD} for all decks, except for the SEP where P_{TD} slightly increases while P_{FD} increases 20-fold (a degrading of detectability, see discussion of Fig. 5). This comes as no surprise given the detection limit of shortwave scene-albedo changes is about 0.04–0.05 (at the grid level) for the NEP and the SEP regions[21], whereas our assessment indicates that ΔrCRE due to IMO2020 does not exceed 0.03 for these two regions (Fig. 1). The lack of a spatially coherent distribution of true detections (yellow shading overlaid with magenta dots), compared to that of N_d, reflects the complex, non-linear nature of aerosol-cloud interactions, particularly the meteorology-dependent, spatiotemporally varying cloud susceptibility to aerosol perturbations[33,46]. This suggests that detecting a shift in the spatial pattern of rCRE is similarly difficult. The
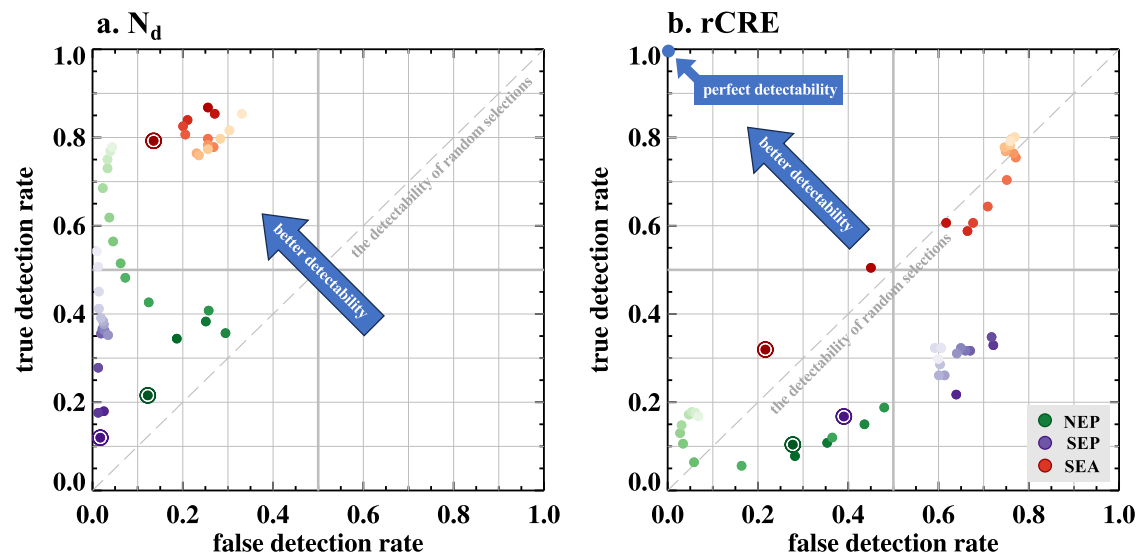
Fig. 5 | Detectability of each deck as ROC curves. Regional (Sc deck-level) detectability illustration for (**a**) N_d and (**b**) rCRE in terms of true and false detection rates. Shades of colors indicate the number of years that goes into the construction of $OBS_{pre-IMO}$ with light colors denoting more years (see Methods). Open circles highlight the results shown in Fig. 4 where the 3-year period immediately proceeding 2020 is used as $OBS_{pre-IMO}$. Perfect detectability lies at the upper-left corner (0,1) of the diagram, and a random detectability lies on the 1-to-1 line. Detectability increases from bottom-right to upper-left.

20-fold increase in the false detection rate of rCRE from that of N_d over the SEP indicates a dominating role of the large-scale meteorological conditions in driving the natural variability in cloud radiative effect, especially over higher latitudes where synoptic variability overwhelms anthropogenic perturbations (Fig. 4b)[47].

On a regional level, moving up-scale from grid level detection, we find that the rCRE-detectability is in general worse than that of a random binary-classifier (Fig. 5b), with the SEA region being the best among the three decks. When more years prior to 2020 are used to construct the pre-IMO2020 observational record (shades of the colors in Fig. 5, see Methods), we do not see an improvement in the rCRE-detectability, as points are moving along the bottom-left to upper-right diagonal, meaning that the true detection rate and false detection rate increase at the same time. Similarly low detectabilities are found in LCF and A_c as well, except that the A_c-detectability is higher over the SEA (Supplementary Fig. 6). In contrast, N_d-detectabilities for all three decks are better than a random binary-classifier (Fig. 5a), and they improve with increasing the number of years that go into the pre-IMO2020 observational record. Again, the SEA region has the highest N_d-detectability, which does not improve further when increasing the number of years in the pre-IMO2020 observational record, representing a saturated detectability. This is expected given the climatologically persistent, narrowly constrained shipping emissions in the SEA shipping lane.

Using the IMO2020 event as a test, we find low detectability in cloud radiative effect at both grid and regional levels, attributed to the strong masking effect of its natural variability, driven by the variability in large-scale meteorological conditions. In general, cloud microphysical changes (e.g., N_d) in response to aerosol perturbations have a higher chance of being detected, due to the relatively weak variability in the climatological relationship between meteorology and N_d and without the added layer of complexity from cloud adjustments. Among the three marine low-cloud regions, the SEA presents the highest detectability in cloud micro- and macro-properties. This suggests that deliberate aerosol seeding proposals for marine cloud brightening may need to be ambitious in terms of the increase in aerosol loading, compared to that of IMO2020, which is already substantial.

Discussion

Variations in the cloud radiative effect derive from natural variability, long term, and transient perturbations. Physical process understanding, existing modeling, and observational evidence suggest that the first is predictable from given meteorological conditions, the second is somewhat predictable given a long enough observational record, while third is rather unpredictable due to the unpredictable nature of socioeconomic activities (e.g. policy changes, global pandemics, etc.) and of transient natural events (e.g., volcanic eruptions and wildfires). In this study, we take advantage of the predictable nature of natural variability and task a data-driven computer-learning program to represent it. This method, by design, helps to remove natural variability from the observed variations, attributing the remaining variability to transient events such as the IMO2020.

Complications of this type of attribution arise from, first, any long-term trend in the aerosol field that co-exists with its natural variability, such as the change in the long-term anthropogenic aerosol trend due to global clean-air efforts[48] and previous IMO regulations on shipping emission of different magnitudes at different spatial scales[19]. These non-meteorological changes and trends in the background aerosol field complicate the relationship between meteorological fields and cloud properties that the NN is tasked to learn, but given their magnitudes, abruptness, and scales, compared to the IMO2020 event, their effect is small. Moreover, dissecting our training datasets to address the impact of individual trends and all previous IMO regulations would greatly limit our sample size for training and degrade the skill of the NN model. Therefore, we choose the longest observational record available to train the NN model, acknowledging the existence of weak, non-meteorological influences in the climatological relationship between meteorological conditions and low-cloud properties. This causes our NN model to essentially extract the mean relationship when given a long-term trend.

Second, transient events that have occurred during approximately the same time period, but possibly of different duration, also complicate the attribution to IMO2020. In our case, this includes the COVID-19 pandemic and the 2019-20 Australian bushfires. Although the most severe fires peaked between December 2019 to January 2020 during the 2019-20 Australian bushfires, Earth System Model simulations have shown that the emitted aerosol traveled around the southern Hemisphere, carried by the atmospheric circulation, and caused radiative perturbations through direct and indirect (via clouds) effects for at least three months at the beginning of 2020[49]. We tested the impact of the 2019–2020 Australian bushfires, to the best of our ability, by repeating the analysis excluding the first four months of 2020. This did not change any of our qualitative conclusions or the spatial patterns in N_d and rCRE perturbations. It only led to a slight

difference of -3 mW m^{-2} in the global radiative forcing estimate (more warming when the first four months are left out of the analysis). This result suggests that the IMO2020 perturbation overwhelmed perturbations from the shorter-duration wildfires. Large-scale simulations are warranted for further assessment of the degree of complexity added by the 2019–2020 Australian bushfires to the impact of IMO2020. Unfortunately, for an observation-based study, it is not possible to disentangle the impact of the COVID-19 pandemic from that of IMO2020 on the observed changes in N_d and rCRE because of the similar spatial scales of the two events. However, a number of modeling and observational studies suggest that the impact of COVID-19 on global and regional cloud properties, surface temperature and precipitation is limited and undetectable (smaller than natural variability), in part due to the offsetting between the weakening in both aerosol cooling and GHG warming[50-53]. Nonetheless, if COVID-19 had not coincided with IMO2020, our estimate might be seen as an upper limit on forcing.

Conclusions

In this work, we devise a data-driven approach to assess the radiative forcing and its detectability of the IMO2020 event, by training an ensemble of neural networks to capture the natural variability in the cloud radiative effect, which is subsequently removed from the 2020-2022 observational record of the cloud field to reveal a substantial warming of 0.074 ±0.005 W m^{-2} as a result of the implementation of the strict shipping fuel regulations in 2020. Although only accounting for the contributions from changes in SW cloud radiative effect from three marine low-cloud regions, our estimate is in line with recent assessments taking different observational and modeling approaches (Fig. 2)[17,35-43]. We extend the focus beyond individual shipping lanes or shipping corridors and find regionally distinct patterns of warming, with a lane-shaped pattern evident in the SEA region that aligns well with the main shipping lane (Fig. 1), providing confidence in the robustness of the data-driven approach. Strong spatial variations in the forcing point to the importance of accounting for co-variations between aerosol perturbations and cloud susceptibility when scaling up radiative forcing from ACI[33,54].

Our radiative forcing estimate represents area-weighted contributions from changes in only the low clouds over the three stratocumulus regions where shipping routes prevail, suggesting that an even stronger warming effect due to IMO2020 is plausible, e.g., with contributions from the lessening of aerosol scattering of sunlight and the adjustments in the trade cumulus regime[36]. This positions our study to suggest that substantial warming from IMO2020 might have contributed to 2023 being the warmest year on record, adding to the existing lines of evidence supporting the postulation of accelerated GHG warming in the coming decades if aggressive aerosol emission regulations are pursued[3,5,55].

An important finding of this work is the strong masking effect of natural variability on the detectability of the radiative impact from IMO2020. By comparing the observational record of cloud properties since 2020 with the NNE-constructed counterfactual cloud fields (signal significance) and the observational records before and after 2020 (signal detection against natural variability), we have demonstrated an overall low detectability in cloud radiative effect of the IMO2020 event, manifested as concurrent changes in true detection rate and false detection rate, despite the substantial decreases in N_d, which has a higher chance of being detected (Figs. 4 and 5). This has important implications for the detectability of large-scale radiative perturbations of both negative and positive sign. For example, if marine cloud brightening were to be attempted our findings imply a strong, regionally-dependent masking of MCB detectability by the natural variability in cloud radiative effect, and that an ambitious seeding strategy might be required. Regionally, the higher detectability for the southeastern Atlantic stratocumulus deck speaks to the benefit of seeding a concentrated area, instead of a widespread area to boost signal detection. The low detectability of the radiative perturbation from IMO2020 at scales larger than ship track or shipping lane speaks to the substantial challenge in validating the viability of MCB at regional-to-global scales[44,45].

Methods
Training data

We use 8 cloud controlling factors (CCFs) and low cloud properties, including relative cloud radiative effect (rCRE), low-cloud fraction (LCF), low-cloud albedo (A_c) and liquid cloud droplet number concentration (N_d) to train the Neural Network (NN). The 8-CCF consists of 8 large-scale meteorological variables (MVs) acquired from the ERA5 reanalysis[56-58] every hour at 0.25° resolution, which are then averaged to monthly gridded (1°) values. These MVs are: pressure velocity at 500 hPa (ω_{500}), relative humidity (RH) at 850 hPa and 925 hPa (RH$_{850}$, RH$_{925}$), potential temperature (θ) at 1000 hPa, 850 hPa and 700 hPa (θ_{1000}, θ_{850}, θ_{700}), column water vapor (CWV), and latent heat flux (LHF). Simply including all available MVs in the training process does not lead to a higher predictive skill but increases the risk of overfitting. Therefore, we select MVs based on the best-performing combination of the 8-MV sets in predicting monthly gridded low-cloud properties by testing all possible combinations of 14 MVs using an ensemble-NN approach[25]. By doing this, we implicitly select MVs that may already be highly correlated with other MVs that are not included as predictors. In other words, the model has indirect access to information not explicitly present as predictors. Essentially, we select the minimum number of MVs that well represent the large-scale meteorological co-variations in order to avoid overfitting. This explains why some seemingly important cloud controlling factors, e.g., surface winds, are not included in the predictors.

The choice of rCRE instead of cloud radiative effect (CRE), is to enable attribution of the radiative effect to low clouds only, instead of all clouds, without the confounding effects of solar angle or surface albedo. rCRE, originally introduced as effective cloud albedo[59], is essentially the effect of cloud on the shortwave radiation received at surface, (i.e., 1–transmittance), calculated as[59,60]

$$\text{rCRE} = 1 - \frac{F_{SW,all}}{F_{SW,clr}} \tag{1}$$

where F_{SW} denotes net SW fluxes at the surface, *all* denotes all sky, and *clr* denotes clear sky. As cloud absorption is generally much smaller than cloud albedo[61], Eq. 1 can be further reduced to

$$\text{rCRE} = f_{cld} \cdot A_{cld} \tag{2}$$

where f_{cld} and A_{cld} denote cloud fraction and cloud albedo. rCRE for low clouds is calculated by replacing f_{cld} and A_{cld} with LCF and A_c, which are obtained from monthly gridded Clouds and the Earth's Radiant Energy System (CERES) Flux By Cloud Type (FBCT) ed. 4.1 product[62,63] where TOA radiative fluxes are binned into 7 cloud effective pressure and 6 cloud optical depth bins. Liquid-cloud areal fraction from the lowest two pressure bins (1100–680 hPa), including all optical depth bins, are summed to obtain LCF, and the shortwave (SW) cloud albedo in these bins are averaged by weighting by their liquid-cloud areal fraction to obtain A_c.

Daily gridded N_d estimates from the Moderate Resolution Imaging Spectroradiometer (MODIS)[64,65] are calculated using the level-2 3.7-μm cloud optical depth (τ) and effective radius (r_e) retrievals[66]. This is an extended version of ref. [64] to include years after 2020. The MODIS N_d estimates can be biased when the scene includes multi-layer or broken cloud fields, sufficiently small τ or r_e, inhomogeneous sub-pixels, or high solar zenith angle[67]. Therefore, we further apply the G18 sampling in ref. [64] to avoid these potentially biased cases. Daily values of N_d are averaged across the satellite platforms (i.e., between Terra and Aqua) over 1-month periods, weighted by the number of available N_d retrievals after the G18 sampling.

We extract these data for oceanic grids covering the northeastern Pacific stratocumulus (NEP) region ($-155°$ to $-115°$ W and 10° to 40° N), the southeastern Pacific (SEP) stratocumulus region ($-100°$ to $-70°$ W and $-40°$ to 0° S), and the southeastern Atlantic (SEA) stratocumulus region ($-15°$ W to 15° E and $-25°$ to 0° S) for this study.

All input data, i.e., MVs, are standardized before training by removing their means and scaling to unit variance. We remove all land-affected grid points using the land-mask variable provided in the ERA5 dataset. Data covering 2003–2018 (inclusive) is used to train the NNs. In total, we have 532,931 data points for all 3 stratocumulus decks (trained all together), which are randomly split into 68%:17%:15% for training, validation, and testing, respectively. Training each stratocumulus region individually does not provide improvement in the NN-performance.

Neural network ensemble architecture

A densely connected feedforward NN is used in this work[25]. It maps input variables, as a vector ($\overrightarrow{m}_0$) containing the 8 MVs, to corresponding target variables (rCRE as the target variable, for example), using a series of operations:

$$\text{rCRE} \sim g_{N+1}(g_N(...(g_2(g_1(\overrightarrow{m}_0))))). \tag{3}$$

where $\overrightarrow{m}_0$ is the input vector and g_N indicates the activation function g for the Nth hidden layer. Taking the first hidden layer as an example, $\overrightarrow{m}_0$ is mapped to n_1 nodes:

$$\overrightarrow{m}_1 = g_1(\overrightarrow{m}_0) = \varphi_1(W_1 \cdot \overrightarrow{m}_0 + \overrightarrow{b}_1), \tag{4}$$

where $\overrightarrow{m}_1$ is an $n_1 \times 1$ vector and φ_1 indicates a nonlinear function, whose weight matrix and bias vector are W_1 and $\overrightarrow{b}_1$, respectively.

Supported by performance tests, we set the NN architecture to have 4 to 8 hidden layers from input to output with the number of node in each layer ranging from 15 to 44. We use the SWISH activation function for all hidden layers[68], which is expressed as

$$\varphi(x) = \frac{x}{1 + e^{-\beta x}}, \tag{5}$$

where the parameter β is set to 1. Given the bounded nature of our target variables (between 0 and 1), the output layer uses the tanh activation function after variables have been transformed to between -1 and 1. We note all outputs from the NN are scaled back to the target's raw unit for the rest of the analyses, including the forcing estimate and the statistical significance testing. The NN is constructed and trained with TensorFlow v2.2[69] using the He uniform variance scaling initializer[70] and the ADAM optimizer[71] for 60 epochs at a learning rate of 10^{-3}, which further drops to 10^{-4} if the mean squared error (MSE; used in the loss function) for validation plateaus.

We generate a 100-member NN ensemble to boost the robustness of the NN prediction and assess the uncertainty of the NN-framework[25]. The diversity of the NN-ensemble members comes from i) the random selection of the training data (68% of all), ii) the number of hidden layers is randomly selected between 4 to 8, iii) the number of nodes in each hidden layer is randomly selected between 15 and 44, and iv) the total number of parameters ranges from 4641 to 4841. We believe this NN ensemble ensures diverse-enough training architectures, to the extent that their predictive skills are not compromised, and provides a robust uncertainty quantification of this framework. The performance of the 100-member NNE, including training, validation, and testing, is shown in Supplementary Fig. 1. The relatively narrow spread in the ensemble performance indicates that the input meteorological variables are well-suited and that the training data and its spatiotemporal scale are optimized for the task of predicting cloud properties at that scale, without any overfitting (Supplementary Fig. 2).

Prior studies have demonstrated the existence of a climatological relationship between cloud properties (e.g., N_d and cloud fraction) and large-scale meteorological conditions[25,27,72]. We here demonstrated the high skill of this 8-MVs-to-cloud-field NN model (Supplementary Fig. 1 and 2). However, when high spatial correlations in aerosol perturbations and cloud responses exist, e.g., over the SEA owing to the alignment between the

shipping route and the prevailing boundary layer wind, the NN model tends to smooth out the rather sharp gradient in N_d and rCRE in and out of the shipping lane. This creates a linear bias across the shipping lane region, particularly over the SEA. Therefore, we apply a linear bias correction to the NN model using the training and testing data samples from 2003 to 2018. The bias-corrected NN ensemble is used to predict the target variables to construct the counterfactual cloud fields for the IMO2020 impacted year (2020–2022).

A known issue with using NN to perform predictions given inputs variables outside the training datasets is extrapolation. Based on the two-dimensional probability density distributions of the 8-MV combinations (Supplementary Fig. 3) between the training and testing time period (2003-2018, shades of gray) and the IMO2020 time period (2020–2022, orange contours), we conclude that the range of the input variables for predictions is within that of the training datasets, i.e., we are not extrapolating.

Definitions of perturbation-significance and perturbation-detectability

The key message we want to convey in this study is that significant perturbations to a cloud field can be hard to detect at times due to coincident changes from other drivers (e.g., meteorological conditions). This involves a key concept, that is the *detectability* of a signal, which is defined differently from the *significance* of a signal in this work. The former assesses whether an observational record is significantly (in a statistical sense) different from a reference observational record, taking into account all potential drivers of the variability in the target, whereas the latter assesses the strength of the signal from a single driver, e.g., in our case aerosol signal from the IMO2020 regulation, all else equal. Note the *detectability* assessment is purely a comparison between two observational records, whereas the *significance* assessment is a comparison between an observational record and a counterfactual record, holding all else equal except the aerosol field.

First, we construct the counterfactual cloud field (NNE_IMO) by predicting the target variables based on the exact meteorological fields of IMO2020-affected years (2020-2022), using the NNE. NNE_IMO represents the aerosol-alone impact of the IMO2020 event, a business-as-usual scenario where only the aerosol field is altered. Therefore, the difference between NNE_IMO and the CERES observational record during 2020-2022 (OBS_IMO) indicates the impact of the IMO2020 event that abruptly and globally affected aerosol emissions, assuming other short-lived abrupt or long-term gradual changes in aerosol emissions during this time are negligible compared to the signal of IMO2020 (see *Discussion*). For the counterfactual values, NN ensemble members are first averaged to produce NNE-mean values for each grid at each month during 2020-2022.

The non-parametric, paired Wilcoxon signed-rank test (WSR test)[73] is used to test if OBS_IMO (3-year time series) is significantly different from NNE_IMO (3-year time series), evaluating the *significance* of the aerosol signal due to IMO2020. The WSR test is selected for this task because (i) it does not assume an underlying probability distribution of the target variables, and (ii) it includes temporal information in the test (i.e., a paired test), which fits better in the context of aerosol-cloud interactions given that cloud susceptibility to aerosol perturbations varies seasonally in response to the seasonality in large-scale meteorological conditions[33,54].

Next, we evaluate if a there is a detectable difference, in the observational record, post- and pre- IMO2020 (OBS_IMO versus OBS_pre-IMO), including the impact from all cloud variability drivers, using the WSR test. Since the WSR test is a paired test, we need to construct a 3-year time series to use as the pre-IMO2020 reference record. An obvious choice is the 3-year period that immediately proceeds the IMO2020 event (2017–2019). This 2017–2019 time period is used to produce the results shown in Fig. 4. We also explored the dependence of the IMO2020 *detectability* on the number of years that goes into the construction of the 3-year pre-IMO2020 record by averaging every 3-year periods that proceed 2020 with an 1-year interval. Essentially, the first pre-IMO2020 record is simply the 2017–2019 record, the second is the average between 2016–2018 and 2017–2019, the third is the average across 2015–2017, 2016–2018, and 2017–2019, and so on. In total, 15 3-year pre-

IMO2020 records are constructed and used to produce the results shown in Fig. 5, with the 15th being the average across 2003–2005, ..., and 2017–2019.

Statistical significance based on the WSR test is established at the 95% confidence interval (CI), or P-value less than 0.05. Two tests, OBS_{IMO} versus NNE_{IMO} (examining the *significance* of the aerosol signal) and OBS_{IMO} versus $OBS_{pre-IMO}$ (examining whether the observational records are different after and before IMO2020), yield four possible outcomes (Fig. 3): (i) *true detection*—when the aerosol signal is significant while pre- and post-IMO2020 observational records are different, (ii) *no detection*—when neither the aerosol signal is significant nor the two observational records are different, (iii) *missed detection*—when the aerosol signal is significant while the post-IMO2020 record is indifferent from the pre-IMO2020 record, and (iv) *false detection*—when the post-IMO2020 record is different from the pre-IMO2020 record while the aerosol signal is insignificant. The latter two scenarios can be attributed to the influence of natural variability driven by changes in the meteorological conditions, where natural variability is either falsely taken as a perturbation (when the aerosol signal is weak) or strong enough to obscure a significant aerosol signal. The similarity between these four outcomes and the outcomes of a binary classifier (i.e., true positive, false positive, false negative, and true negative) enables us to construct the equivalent confusion matrix and the receiver operating characteristic (ROC) curve in Fig. 5. In our case, true detection rate is the fraction of significant IMO2020 (aerosol-induced) perturbation being detected and false detection rate is the fraction of insignificant IMO2020 (aerosol-induced) perturbation being detected.

Radiative forcing and susceptibility calculations

Perturbations in rCRE, LCF and A_c are translated into perturbations in shortwave radiation at the top-of-atmosphere (ΔSW_{TOA}) using

$$\Delta SW_{TOA}|_{rCRE} = -\langle \Delta rCRE|_{lon,lat,month} \times SW_{in}|_{lon,lat,month} \rangle, \quad (6)$$

$$\Delta SW_{TOA}|_{LCF} = -\langle \Delta LCF|_{lon,lat,month} \times A_c|_{lon,lat,month} \times SW_{in}|_{lon,lat,month} \rangle, \quad (7)$$

$$\Delta SW_{TOA}|_{A_c} = -\langle \Delta A_c|_{lon,lat,month} \times LCF|_{lon,lat,month} \times SW_{in}|_{lon,lat,month} \rangle, \quad (8)$$

where $SW_{in|lon,lat,mounth}$ denotes incoming solar radiation at a given longitude, latitude and month and angle brackets indicate temporal averaging. Positive ΔSW_{TOA} indicates warming. These radiative perturbations are scaled by the fractional areal coverage of the domain relative to the globe to derive global forcings that can be compared to estimates from other studies[17,36]. Using the 100 ensemble members from the NNE, we estimate the uncertainty ranges for the mean radiative forcing are ±0.002 W m^{-2}, ±0.005 W m^{-2}, and ±0.007 W m^{-2}, at 68%, 95%, and 99% confidence intervals, respectively, equivalent to 1-σ, 2-σ, and 3-σ of a normal distribution.

A top-down rough estimation of cloud susceptibility to N_d perturbations is provided by calculating the linear regression slope between percentage changes in cloud properties ($\Delta rCRE$, ΔLCF, and ΔA_c) and percentage changes in N_d. Cloud susceptibilities indicate the percentage change in cloud properties given a unit percentage change in N_d. These values are reported in Supplementary Fig. 4.

Data availability

The CERES FBCT data were obtained from the NASA's Langley Research Center CERES ordering tool at https://ceres.larc.nasa.gov/data/ (downloaded 21 February 2024), and are available at https://doi.org/10.5067/Terra-Aqua/CERES/FLUXBYCLDTYP-MONTH_L3.004A. The fifth-generation ECMWF (ERA5) atmospheric reanalyses of the global climate data were obtained from the Copernicus Climate Change Service at https://cds.climate.copernicus.eu/ (downloaded 21 February 2024), and are available at https://doi.org/10.24381/cds.6860a573 and https://doi.org/10.24381/cds.f17050d7. The MODIS-derived N_d data are available from the Centre for Environmental Data Analysis at https://doi.org/10.5285/864a46cc65054008857ee5bb772a2a2b.

Code availability

Code is available from the corresponding author on reasonable request.

References

1. Stephens, G. L. et al. An update on earth's energy balance in light of the latest global observations. *Nat. Geosci.* **5**, 691–696 (2012).
2. Forster, P. et al. in *Climate Change 2021: The Physical Science Basis, Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. Masson-Delmotte, V. et al.) 923–1054 (Cambridge University Press, 2021).
3. Wood, R., Vogt, M. A. & McCoy, I. L. Aggressive aerosol mitigation policies reduce chances of keeping global warming to below 2C. *Earth's Future* **12**, https://doi.org/10.1029/2023ef004233 (2024).
4. Nair, H. R. C. R. et al. Aerosol demasking enhances climate warming over south asia. *npj Clim. Atmos. Sci.* **6**, 39 (2023).
5. Wang, P. et al. Aerosols overtake greenhouse gases causing a warmer climate and more weather extremes toward carbon neutrality. *Nat. Commun.* **14**, 7257 (2023).
6. Twomey, S. Pollution and the planetary albedo. *Atmos. Environ.* **8**, 1251–1256 (1974).
7. Albrecht, B. A. Aerosols, cloud microphysics, and fractional cloudiness. *Science* **245**, 1227–1230 (1989).
8. Caldwell, P., Bretherton, C. S. & Wood, R. Mixed-layer budget analysis of the diurnal cycle of entrainment in southeast pacific stratocumulus. *J. Atmos. Sci.* **62**, 3775–3791 (2005).
9. Wang, S., Wang, Q. & Feingold, G. Turbulence, condensation, and liquid water transport in numerically simulated nonprecipitating stratocumulus clouds. *J. Atmos. Sci.* **60**, 262–278 (2003).
10. Bretherton, C. S., Blossey, P. N. & Uchida, J. Cloud droplet sedimentation, entrainment efficiency, and subtropical stratocumulus albedo. *Geophys. Res. Lett.* **34**, L03813 (2007).
11. Zhang, J., Chen, Y.-S., Yamaguchi, T. & Feingold, G. Cloud water adjustments to aerosol perturbations are buffered by solar heating in non-precipitating marine stratocumuli. *Atmos. Chem. Phys.* **24**, 10425–10440 (2024).
12. Conover, J. H. Anomalous cloud lines. *J. Atmos. Sci.* **23**, 778–785 (1966).
13. Twomey, S., Howell, H. B. & Wojciechowski, T. A. Comments on "anomalous cloud lines''. *J. Atmos. Sci.* **25**, 333–334 (1968).
14. Ramachandran, P., Zoph, B. & Le, Q. V. *Resolution mepc.305(73) (Adopted on 26 October 2018) Amendments to the Annex of the Protocol of 1997 to Amend the International Convention for the Prevention of Pollution from Ships, 1973, as Modified by the Protocol of 1978 Relating Thereto* (2019).
15. Watson-Parris, D. et al. Shipping regulations lead to large reduction in cloud perturbations. *Proc. Natl Acad. Sci.* **119**, e2206885119 (2022).
16. Yuan, T. et al. Global reduction in ship-tracks from sulfur regulations for shipping fuel. *Sci. Adv.* **8**, eabn7988 (2022).
17. Diamond, M. S. Detection of large-scale cloud microphysical changes within a major shipping corridor after implementation of the international maritime organization 2020 fuel sulfur regulations. *Atmos. Chem. Phys.* **23**, 8259–8269 (2023).
18. Manshausen, P., Watson-Parris, D., Christensen, M. W., Jalkanen, J.-P. & Stier, P. Rapid saturation of cloud water adjustments to shipping emissions. *Atmos. Chem. Phys.* **23**, 12545–12555 (2023).
19. Gryspeerdt, E. et al. Constraining the aerosol influence on cloud liquid water path. *Atmos. Chem. Phys.* **19**, 5331–5347 (2019).
20. Diamond, M. S., Director, H. M., Eastman, R., Possner, A. & Wood, R. Substantial cloud brightening from shipping in subtropical low clouds. *AGU Adv.* **1**, e2019AV000111 (2020).

21. Seidel, D. J., Feingold, G., Jacobson, A. R. & Loeb, N. Detection limits of albedo changes induced by climate engineering. *Nat. Clim. Change* **4**, 93–98 (2014).

22. National Academies of Sciences, Engineering, and Medicine (NASEM) report. *Reflecting Sunlight: Recommendations for Solar Geoengineering Research and Research Governance* (The National Academies Press, 2021). https://doi.org/10.17226/25762.

23. Latham, J. & Smith, M. H. Effect on global warming of wind-dependent aerosol generation at the ocean surface. *Nature* **347**, 372–373 (1990).

24. Latham, J. et al. Marine cloud brightening. *Philos. Trans. Math. Phys. Eng. Sci.* **370**, 4217–4262 (2012).

25. Chen, Y.-S., Yamaguchi, T., Bogenschutz, P. A. & Feingold, G. Model evaluation and intercomparison of marine warm low cloud fractions with neural network ensembles. *J. Adv. Model. Earth Syst.* **13**, e2021MS002625 (2021).

26. Andersen, H., Cermak, J., Fuchs, J., Knutti, R. & Lohmann, U. Understanding the drivers of marine liquid-water cloud occurrence and properties with global observations using neural networks. *Atmos. Chem. Phys.* **17**, 9535–9546 (2017).

27. Chen, Y. et al. Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover. *Nat. Geosci.* **15**, 609–614 (2022).

28. Andersen, H., Cermak, J., Zipfel, L. & Myers, T. A. Attribution of observed recent decrease in low clouds over the northeastern pacific to cloud-controlling factors. *Geophys. Res. Lett.* **49**, e2021GL096498 (2022).

29. Andersen, H. et al. Sensitivities of cloud radiative effects to large-scale meteorology and aerosols from global observations. *Atmos. Chem. Phys.* **23**, 10775–10794 (2023).

30. Jia, Y., Andersen, H. & Cermak, J. Analysis of cloud fraction adjustment to aerosols and its dependence on meteorological controls using explainable machine learning. *EGUsphere* **2023**, 1–25 (2023).

31. Yamaguchi, T., Feingold, G. & Kazil, J. Stratocumulus to cumulus transition by drizzle. *J. Adv. Model. Earth Syst.* **9**, 2333–2349 (2017).

32. Wall, C. J., Storelvmo, T. & Possner, A. Global observations of aerosol indirect effects from marine liquid clouds. *Atmos. Chem. Phys.* **23**, 13125–13141 (2023).

33. Zhang, J. & Feingold, G. Distinct regional meteorological influences on low-cloud albedo susceptibility over global marine stratocumulus regions. *Atmos. Chem. Phys.* **23**, 1073–1090 (2023).

34. Bellouin, N. et al. Bounding global aerosol radiative forcing of climate change. *Rev. Geophys.* **58**, e2019RG000660 (2020).

35. Gettelman, A. et al. Has reducing ship emissions brought forward global warming? *Geophys. Res. Lett.* **51**, https://doi.org/10.1029/2024gl109077 (2024).

36. Yuan, T. et al. Abrupt reduction in shipping emission as an inadvertent geoengineering termination shock produces substantial radiative warming. *Commun. Earth Environ.* **5**, 281 (2024).

37. Yoshioka, M., Grosvenor, D. P., Booth, B. B. B., Morice, C. P. & Carslaw, K. S. Warming effects of reduced sulfur emissions from shipping. *EGUsphere* **2024**, 1–19 (2024).

38. Quaglia, I. & Visioni, D. Modeling 2020 regulatory changes in international shipping emissions helps explain 2023 anomalous warming. *EGUsphere* **2024**, 1–19 (2024).

39. Skeie, R. B., Byrom, R., Hodnebrog, Ø., Jouan, C. & Myhre, G. Multi-model effective radiative forcing of the 2020 sulphur cap for shipping. *EGUsphere* **2024**, 1–14 (2024).

40. Jordan, G. & Henry, M. IMO2020 regulations accelerate global warming by up to 3Â Years in UKESM1. *Earth's. Future* **12**, e2024EF005011 (2024).

41. Jin, Q., Grandey, B. S., Rothenberg, D., Avramov, A. & Wang, C. Impacts on cloud radiative effects induced by coexisting aerosols converted from international shipping and maritime DMS emissions. *Atmos. Chem. Phys.* **18**, 16793–16808 (2018).

42. Sofiev, M. et al. Cleaner fuels for ships provide public health benefits with climate tradeoffs. *Nat. Commun.* **9**, 406 (2018).

43. Bilsback, K. R. et al. Beyond Sox reductions from shipping: assessing the impact of nox and carbonaceous-particle controls on human health and climate. *Environ. Res. Lett.* **15**, 124046 (2020).

44. Diamond, M. S. et al. To assess marine cloud brightening's technical feasibility, we need to know what to study–and when to stop. *Proc. Natl Acad. Sci.* **119**, e2118379119 (2022).

45. Feingold, G. et al. Physical science research needed to evaluate the viability and risks of marine cloud brightening. *Sci. Adv.* **10**, eadi8594 (2024).

46. Zhou, X., Zhang, J. & Feingold, G. On the importance of sea surface temperature for aerosol–induced brightening of marine clouds and implications for cloud feedback in a future warmer climate. *Geophys. Res. Lett.* **48**, https://doi.org/10.1029/2021gl095896 (2021).

47. Zhang, J. & Zuidema, P. Sunlight-absorbing aerosol amplifies the seasonal cycle in low-cloud fraction over the southeast Atlantic. *Atmos. Chem. Phys.* **21**, 11179–11199 (2021).

48. Bauer, S. E. et al. The turning point of the aerosol era. *J. Adv. Model. Earth Syst.* **14**, e2022MS003070 (2022).

49. Fasullo, J. T., Rosenbloom, N. & Buchholz, R. A multiyear tropical pacific cooling response to recent Australian wildfires in cesm2. *Sci. Adv.* **9**, eadg1213 (2023).

50. Forster, P. M. et al. Current and future global climate impacts resulting from COVID-19. *Nat. Clim. Change* **10**, 913–919 (2020).

51. Diamond, M. S. & Wood, R. Limited regional aerosol and cloud microphysical changes despite unprecedented decline in nitrogen oxide pollution during the february 2020 covid-19 shutdown in china. *Geophys. Res. Lett.* **47**, e2020GL088913 (2020).

52. Gettelman, A., Lamboll, R., Bardeen, C. G., Forster, P. M. & Watson-Parris, D. Climate impacts of COVID-19 induced emission changes. *Geophys. Res. Lett.* **48**, e2020GL091805 (2021).

53. Jones, C. D. et al. The climate response to emissions reductions due to Covid-19: Initial results from Covidmip. *Geophys. Res. Lett.* **48**, e2020GL091883 (2021).

54. Zhang, J., Zhou, X., Goren, T. & Feingold, G. Albedo susceptibility of northeastern pacific stratocumulus: the role of covarying meteorological conditions. *Atmos. Chem. Phys.* **22**, 861–880 (2022).

55. Hansen, J. E. et al. Global warming in the pipeline. *Oxf. Open Clim. Change* **3**, kgad008 (2023).

56. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. Roy. Meteor. Soc.* **146**, 1999–2049 (2020).

57. Copernicus Climate Change Service. *ERA5 monthly averaged data on pressure levels from 1940 to present* [Data set], https://doi.org/10.24381/CDS.6860A573 (2019).

58. Copernicus Climate Change Service. *ERA5 monthly averaged data on single levels from 1940 to present* [Data set], https://doi.org/10.24381/CDS.F17050D7 (2019).

59. Betts, A. K. Coupling of water vapor convergence, clouds, precipitation, and land-surface processes. *J. Geophys. Res. -Atmos.* **112**, 1–14 (2007).

60. Liu, Y., Wu, W., Jensen, M. P. & Toto, T. Relationship between cloud radiative forcing, cloud fraction and cloud albedo, and new surface-based approach for determining cloud albedo. *Atmos. Chem. Phys.* **11**, 7155–7170 (2011).

61. Gautier, C. & Landsfeld, M. Surface solar radiation flux and cloud radiative forcing for the atmospheric radiation measurement arm southern great plains (sgp): A satellite surface observations and radiative transfer model study. *J. Atmos. Sci.* **54**, 1289–1307 (1997).

62. Sun, M. et al. Clouds and the Earth's Radiant Energy System (CERES) FluxByCldTyp Edition 4 Data Product. *J. Atmos. Ocean. Tech.* **39**, 303–318 (2022).

63. NASA/LARC/SD/ASDC. CERES Monthly Daytime Mean Regionally Averaged Terra and Aqua TOA Fluxes and Associated Cloud Properties Stratified by Optical Depth and Effective Pressure Edition4A. *[Data set]*, https://doi.org/10.5067/TERRA-AQUA/CERES/FLUXBYCLDTYP-MONTH_L3.004A (2020).

64. Gryspeerdt, E. et al. The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data. *Atmos. Meas. Tech.* **15**, 3875–3892 (2022).

65. Gryspeerdt, E. et al. *Cloud Droplet Number Concentration, Calculated from the Modis (Moderate Resolution Imaging Spectroradiometer) Cloud Optical Properties Retrieval and Gridded Using Different Sampling Strategies [Data set]*, https://doi.org/10.5285/864A46CC65054008857EE5BB772A2A2B (2022).

66. Platnick, S. et al. Modis atmosphere l2 joint atmosphere product. *NASA MODIS Adaptive Processing System* Goddard Space Flight Center, USA, https://doi.org/10.5067/MODIS/MODATML2.006 (2015).

67. Grosvenor, D. P. et al. Remote sensing of droplet number concentration in warm clouds: a review of the current state of knowledge and perspectives. *Rev. Geophys.* **56**, 409–453 (2018).

68. Ramachandran, P., Zoph, B. & Le, Q. V. Searching for activation functions, https://doi.org/10.48550/arXiv.1710.05941 (2017).

69. Abadi, M. et al. TensorFlow: Large-scale machine learning on heterogeneous systems (2015). Software available from tensorflow.org.

70. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *2015 IEEE International Conference on Computer Vision (ICCV)*, 1026–1034, https://doi.org/10.1109/ICCV.2015.123 (2015).

71. Kingma, D. P. & Ba, J. Adam. A method for stochastic optimization, Preprint at https://doi.org/10.48550/arXiv.1412.6980 (2017).

72. Mülmenstädt, J. et al. General circulation models simulate negative liquid water path–droplet number correlations, but anthropogenic aerosols still increase simulated liquid water path. *Atmos. Chem. Phys.* **24**, 7331–7345 (2024).

73. Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics Bull.* **1**, 80–83 (1945).

Acknowledgements

Author contributions

J.Z. conceived the study, performed data analysis, and led the manuscript writing. Y.-S.C. co-conceived the study, provided technical support on building the neural network framework, and contributed to the interpretation of the results and manuscript writing. E.G. provided the gridded N_d dataset and contributed to the interpretation of the results and manuscript writing. T.Y. and G.F. secured the funding for this work, supervised the study, and contributed to the interpretation of the results and manuscript writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01911-9.

Correspondence and requests for materials should be addressed to Jianhao Zhang.

Peer review information *Communications Earth & Environment* thanks Lekha Patel, Robert Wood and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Sylvia Sullivan, Heike Langenberg. [A peer review file is available].

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.