Analysis of cloud fraction adjustment to aerosols and its dependence on meteorological controls using explainable machine learning

Yichen Jia [1, 2], Hendrik Andersen [1, 2], and Jan Cermak [1, 2]

[1]Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Karlsruhe, Germany
[2]Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Karlsruhe, Germany

Correspondence: Yichen Jia (yichen.jia@kit.edu)

Abstract. Aerosol-cloud interactions (ACI) have a pronounced influence on the Earth's radiation budget but continue to pose one of the most substantial uncertainties in the climate system. Marine boundary-layer clouds (MBLCs) are particularly important since they cover a large portion of the Earth's surface. One of the biggest challenges in quantifying ACI from observations lies in isolating adjustments of cloud fraction (CLF) to aerosol perturbations from the covariability and influence of the local meteorological conditions. In this study, this isolation is attempted using nine years (2011–2019) of near-global daily satellite cloud products in combination with reanalysis data of meteorological parameters. With cloud-droplet number concentration (N_d) as a proxy for aerosol, MBLC CLF is predicted by region-specific gradient boosting machine learning models. By means of SHapley Additive exPlanation (SHAP) regression values, CLF sensitivity to N_d and meteorological factors as well as meteorological influences on the N_d–CLF sensitivity are quantified. The regional ML models are able to capture on average 45 % of the CLF variability. Global patterns of CLF sensitivity show that CLF is positively associated with N_d, in particular in the stratocumulus-to-cumulus transition regions and in the southern hemispheric midlatitudes. CLF sensitivity to estimated inversion strength (EIS) is ubiquitously positive and strongest in tropical and subtropical regions topped by stratocumulus and within the midlatitudes. Globally, increased sea-surface temperature (SST) reduces CLF, particularly in stratocumulus regions. The spatial patterns of CLF sensitivity to horizontal wind components in the free troposphere point to the impact of synoptic-scale weather systems and vertical wind shear on MBLCs. The N_d–CLF relationship is found to depend more on the selected thermodynamical variables than dynamical variables, and in particular on EIS and SST. In the midlatitudes, a stronger inversion is found to amplify the N_d–CLF relationship, while this is not observed in the stratocumulus regions. In the stratocumulus-to-cumulus transition regions, the N_d–CLF sensitivity is found to be amplified by higher SSTs, potentially pointing to N_d more frequently delaying this transition in these conditions. The expected climatic changes of EIS and SST may thus influence future forcings from ACIs. The near-global ML framework introduced in this study produces a better quantification of the response of MBLC CLF to aerosols taking into account the covariations with meteorology.

1 Introduction

The emission of aerosols into the atmosphere affects the Earth's climate, in particular by masking part of the warming effect from greenhouse gases by reflecting solar radiation and changing cloud properties. Aerosol-cloud interactions (ACI) can strongly influence the Earth's energy distribution and thus also contribute a substantial uncertainty to past and future climate projections. The effective radiative forcing due to ACI (ERFaci) is assessed as –1.0 Wm^{-2} with an uncertainty range of –1.7 to –0.3 Wm^{-2} (Forster et al., 2021), albeit decades of effort and headway have been made in understanding the complex system of aerosols, clouds and their environmental controls. The correct representation of ACI in Earth system models (ESMs) remains a tremendous challenge because of the lack of accurate global quantification of the cloud-related fine-scale processes, and the lack of larger-scale constraints from the existing measurement systems at the ESM spatiotemporal resolution (Fan et al., 2016; Seinfeld et al., 2016; Sato et al., 2018).

Marine boundary layer clouds (MBLCs) cover over 23 % of the global ocean surface (Wood, 2012). Due to relatively small temperature differences between MBLC top and the sea surface, they only weakly impact outgoing longwave radiation but greatly reflect incoming shortwave radiation, leading to a strong net cooling effect (Hartmann et al., 1992). MBLCs play a critical role in the Earth's radiative balance (Zheng et al., 2021) and, in this regard, are the most important cloud type (Chen et al., 2014). Furthermore, MBLCs are especially susceptible to aerosol perturbations due to their ~~physical properties (Wood et al., 2015)~~relatively low optical depths (Turner, 2007; Leahy et al., 2012) and their formation in environments typically characterized by lower anthropogenic aerosol loading than continental clouds (Platnick and Twomey, 1994). Therefore, a deeper understanding of the aerosol-MBLC interactions is crucial to reduce the uncertainties in climate predictions. Atmospheric aerosols are critical for the formation of clouds as cloud condensation nuclei (CCN). Increases in aerosols are associated with increases in cloud droplet number concentration (N_d). As the cloud water is distributed among more droplets, cloud droplet effective radius (r_e) shrinks at constant liquid water content, resulting in an enhancement of cloud brightness and a negative instantaneous radiative forcing (Twomey, 1977). The likelihood of collision and coalescence subsequently decreases due to smaller drop sizes, hampering rainfall formation, which can prolong cloud lifetime and thus increase cloud fraction (CLF) (Albrecht, 1989). However, the aerosol–CLF relationship is complex, and the sign of the CLF adjustment can also be the opposite. This has been found in particular for non-precipitating clouds, stemming from ~~the intensified evaporation and~~ enhanced entrainment mixing with ambient air over the ~~cloud top (Small et al., 2009; Christensen and Stephens, 2011)~~ clouds owing to shorter evaporation timescales (Wang et al., 2003; Jiang et al., 2006; Small et al., 2009) or reduced sedimentation (Ackerman et al., 2004; Bretherton et al., 2007) because of smaller droplet sizes.

From the perspective of observations at satellite scales, though there are studies suggesting a negative relationship between aerosols and CLF (Dey et al., 2011; Small et al., 2011), it has been documented by multiple studies that the overall CLF increases in response to increasing aerosols ~~(e.g. Kaufman and Koren, 2006; Gryspeerdt et al., 2016; Christensen et al., 2017; Andersen et al~~ (e.g. Kaufman and Koren, 2006; Yuan et al., 2011; Gryspeerdt et al., 2016; Christensen et al., 2017; Andersen et al., 2017; Fuchs et al., 20. Likewise, studies based on ESMs reported substantial negative ERFaci due to liquid water path (LWP) and CLF adjustments (e.g. Zelinka et al., 2014). In spite of the attribution of such adjustments in ESMs primarily to LWP adjustments (Ghan et al.,

2016), a global satellite-based study by Bender et al. (2019) suggested that LWP adjustments are overestimated in ESMs, and that aerosol impact on CLF dominates the negative aerosol forcing. ~~A similar conclusion was also recently drawn~~ This is supported by observational evidence presented by Toll et al. (2019) who also reported an overestimation of LWP adjustment in climate models, and by Chen et al. (2022) who ~~found~~ recently highlighted the role of CLF increases due to aerosols from a large volcano eruption ~~to be~~ as the main cause of the associated forcing. Some large-eddy simulations have however suggested a negative response of CLF of trade wind cumulus to aerosol perturbations (Xue and Feingold, 2006; Seifert et al., 2015). While most studies, both from observational and model points of view, are in agreement that generally CLF increases with increasing aerosols due to a prolonged lifetime (Douglas and L'Ecuyer, 2022), the magnitude of the response of CLF to aerosols and its corresponding adjustments are still highly uncertain. For satellite-based analyses, one of the most challenging aspects in the quantification of CLF adjustment is isolating the influence of the aerosol loading on cloud properties from confounding covariations with meteorological parameters (Andersen et al., 2016; Gryspeerdt et al., 2019; Bellouin et al., 2020), paired with aerosol retrieval issues related to aerosol swelling and 3D radiative effects in the vicinity of clouds (Loeb and Schuster, 2008; Schwarz et al., 2017). Recent observational studies have utilized different methods to tackle this issue. A first approach is to stratify the data by meteorological factors and therefore accounting for local meteorology in the relationships (e.g. Su et al., 2010; Chen et al., 2014; Andersen and Cermak, 2015). Secondly, using N_d as a mediating variable was proposed by Gryspeerdt et al. (2016) to analyze the causal pathway between aerosol optical depth and CLF. Another approach is to use a sampling strategy that applies a cloud–aerosol pairing algorithm (Christensen et al., 2017). However, these methods do not account for aerosol retrieval issues, meteorological influencing factors and confounders at once, which is essential to constrain the CLF adjustment. Recently, several studies have successfully used machine learning (ML) to account for non-linearities and meteorological factors to quantify ACI (Andersen et al., 2017; Fuchs et al., 2018; Dadashazar et al., 2021; Zipfel et al., 2022). ML regression algorithms allow to predict CLF (predictand) on the basis of aerosol and meteorological factors at the same time and treat the aerosol-cloud-meteorology system as a whole. In addition, ML models can represent non-linear interactive systems, which can be analyzed in sensitivity analyses with explainable ML techniques. Explainable ML refers to the techniques explaining the predictions of a trained ML model by explicitly quantifying the relationships, which helps improve the understandability, transparency and trustworthiness of the ML models (Beucler et al., 2020).

In this study, we set up region-specific ML models at a global scale using satellite and reanalysis data sets to predict CLF to analyze N_d-induced changes in MBLCs. The goal of the explainable ML framework is to quantify the global sensitivity patterns of CLF to N_d and meteorological factors. In addition, we aim to estimate the magnitude of the dependence of N_d–CLF sensitivity on the meteorological factors using ~~SHAP~~ SHapley Additive exPlanation (SHAP) interaction values, providing a new and insightful pathway to more profound knowledge of the physical processes relevant to the CLF adjustment and hence to a global constraint on aerosol-induced CLF changes accounting for meteorological covariations. The hypothesis of this study is that the response of cloud fraction of MBLCs to aerosol perturbations is positive, but buffered, i.e. reduced or amplified, by ambient meteorology and both the sensitivities and the interactions with meteorological factors have distinct regional patterns.

2 Data and methods

2.1 Data sets

This work combines nine years (2011-2019) of satellite retrievals from Moderate Resolution Imaging Spectroradiometer (MODIS) and reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) from 60°N to 60°S. ~~The information on cloud optical and physical properties for~~ In this study, MBLCs are defined as single-layer ~~warm MBLCs and~~ warm cloud fields with cloud top temperatures higher than 268 K. To achieve this, the information on CLF (product: Cloud_Retrieval_Fraction_1L_Liquid), r_e (product: Cloud_Effective_Radius_1L_Liquid_Mean), Cloud optical depth (τ_c; product: Cloud_Optical_Thickness_1L_Liquid_Mean), cloud top temperature (CTT; product: Cloud_Top_Temperature_Mean) and satellite viewing geometry are obtained from MODIS level-3 collection-6.1 atmosphere daily ~~product~~ products on the Terra platform (MOD08_D3), which are gridded into 1°× 1°globally from level-2 atmospheric products. ~~Except for CLF~~ CLF serves as the predictand ~~,~~ in this study. The computation of N_d relies on τ_c and r_e, ~~cloud optical depth (τ_c), cloud top temperature (CTT)~~with filtering criteria based on CTT, solar zenith viewing angle and satellite zenith angle~~are used to filter and compute the N_d~~, as elaborated in the following.

The equation used to calculate the MODIS N_d is from Quaas et al. (2006), which depends on the retrievals of r_e and τ_c, so do the uncertainties on the errors propagated from r_e and τ_c:

$$N_d = \alpha \tau_c^{0.5} r_e^{-2.5} \tag{1}$$

where $\alpha = 1.37 \times 10^{-5}$ m$^{-0.5}$ is a constant related to adiabatic growth rate. The uncertainties in N_d retrievals are exhaustively evaluated by (Grosvenor et al., 2018), which suggests that the uncertainties in averaged N_d over 1°× 1°grid box (spatial resolution of the MODIS products used in this study) decrease by over 50 % compared to pixel-level uncertainties. This retrieval approach relies on the assumed adiabaticity in global marine warm clouds where liquid water content and r_e increase monotonically and N_d distributes as constant vertically. Departure from the adiabatic assumption (e.g. due to entrainment) would result in N_d retrieval biases (Merk et al., 2016; Bennartz and Rausch, 2017). The uncertainty related to the estimation of N_d from MODIS also depends on liquid CLF. N_d is less biased in the regions of larger CLF where clouds are more homogeneous, while in the regions with lower CLF N_d retrievals are sparser and less reliable (Grosvenor et al., 2018; Zhu et al., 2018). In such heterogeneous cloud fields, subpixel effects in the retrieval of r_e can negatively bias the retrieved N_d values ~~(Grosvenor et al., 2018)~~(Zhang and Platnick, 2011; Zhang et al., 2012; Grosvenor et al., 2018). Such retrieval biases could cause a bias in the N_d–CLF relationship as well.

Following the screening criteria for more reliable N_d demarcated by Gryspeerdt et al. (2022), only clouds restricted to single-layer in liquid phase with CTT higher than 268 K are considered. As suggested by Quaas et al. (2006), samples with r_e < 4 μm and τ_c < 4 are excluded to cope with the high r_e retrieval uncertainties at low τ_c. In addition, solar and sensor viewing zenith angles respectively greater than 65°and 55°are removed to avoid the large biases in r_e and τ_c retrievals (as in Grosvenor et al., 2018). The pixels selected according to the above sampling strategies generate more reliable N_d estimates.

~~20 atmospheric~~ Atmospheric and oceanic variables are taken from the fifth generation ECMWF atmospheric reanalysis of the global climate (ERA5) at an hourly frequency (Table 1) (Hersbach et al., 2020). The ~~reanalysis~~ ERA5 data sets are ~~subsequently~~ harmonized to fit the level-3 MODIS data by ~~regridding~~ first being resampled to $1° \times 1°$ ~~resolution and collocating to~~ from their default $0.25° \times 0.25°$ spatial resolution using bilinear interpolation, and they are subsequently collocated to Terra MODIS by extracting hourly data to align with the UTC overpass times of the Terra satellite for each grid cell, yielding a spatiotemporally matched MODIS-ERA5 combined data set for ~~the training of~~ training the ML models. ~~Only samples within 1-99 percentiles~~ For N_d retrievals, only samples within 1–99 percentiles are retained to exclude potential unrealistic outliers from r_e and τ_c retrievals (Zipfel et al., 2022). Furthermore, the explanation of ML models in this study relies on using linear regressions to capture the distribution of individual prediction instances, and the extreme values may excessively magnify or reduce the sensitivity or interactive effects quantified by SHAP (shown in Fig. 1 and discussed in Sect. 2.3.2). The threshold of 1–99 percentiles for each predictor ~~are retained~~ is thus adopted to remove the values at the very tails of the specific distribution and to improve the robustness of the estimated sensitivities. To define the sensitivities of CLF and the interactive effects of meteorological factors, the natural logarithm of N_d is taken. Estimated inversion strength (EIS) is calculated based on the formulation from Wood and Bretherton (2006) and in this study, it is dependent only on atmospheric temperatures at 700 hPa and at the level of 1000 hPa. All input data ~~is~~ for each XGB model (i.e. for each $5° \times 5°$ window) are standardized for comparability of the estimates of the sensitivity and the interactive effect with meteorology (Sect. ~~2.3.2~~ 2.3.2).

2.2 Machine learning model setup

Extreme Gradient Boosting (XGB) is a distributed tree boosting algorithm aiming to provide a scalable, portable and flexible library under the Gradient Boosting framework (Chen and Guestrin, 2016). The state-of-the-art XGB can be implemented efficiently in Python and has been recently used to study clouds and ACI (Andersen et al., 2022; Douglas and L'Ecuyer, 2022). As an extension of previous gradient boosting methods, XGB has incorporated regularization techniques which help prevent overfitting and improve model generalization. Besides, the subsampling on training subsets and column (feature) subsampling techniques can shorten the running time and also avert overfitting and hence elevate model performance (Chen and Guestrin, 2016). Relevant regularization and subsampling hyperparameters are tuned using Bayesian Optimization to determine the best combination, see Table 2 for the search space.

Data from 2011 to 2016 are used for training and data from 2017 to 2019 for testing (independent train/test split about 67 %/33 %). As the data set is chronologically split, we ensure that the models are not trained on neighbouring data points, which would not be fully independent due to the high autocorrelation of the model inputs. The data at a $1° \times 1°$ spatial resolution are aggregated in $5° \times 5°$ geographical windows, where an individual, independent XGB model is trained and tested for each "window". Hereby a region-specific ML framework is established to potentially capture regionally-specific relationships and characteristics and thus the regional patterns of CLF adjustment. The coarser $5° \times 5°$ spatial resolution of the modelling grid increases the sample size by a factor of $\approx$25 which is helpful to establish robust sensitivity estimates. In addition, at the spatial resolution of $1° \times 1°$ summarized in $5° \times 5°$ degree windows, the spatial scale is adequate for ACI sensitivity estimation (Grandey and Stier, 2010). ~~Only the geographical windows with a number of available observations higher than 6000 data~~

Table 1. Summary of the predictors from ERA5 reanalysis.

Predictor Name	**Abbreviation**	**Units**
Instantaneous pressure level parameters (at 700 hPa, 850 hPa)		
Relative humidity	RH_{700}, RH_{850}	%
Specific humidity	SH_{700}, SH_{850}	kgkg^{-1}
Temperature	t_{700}, t_{850}	K
Vertical velocity	ω_{700}, ω_{850}	Pas^{-1}
Eastward wind component	u_{700}, u_{850}	ms^{-1}
Northward wind component	v_{700}, v_{850}	ms^{-1}
Surface and single level parameters (instantaneous or mean rates/fluxes)		
Eastward and northward wind component at 10 m	u_{10}, v_{10}	ms^{-1}
Boundary-layer height	BLH	m
Convective available potential energy	CAPE	Jkg^{-1}
Sea surface temperature	SST	K
Total column water vapour	TCWV	kgm^{-2}
Mean large-scale precipitation fraction	PF	Proportion
Mean surface sensible/latent heat flux	SHF/LHF	Wm^{-2}
Calculated		
Estimated inversion strength	EIS	K

Table 2. Overview of the hyperparameters tuned for regional Extreme Gradient Boosting models using Bayesian Optimization.

Hyperparameter name	**Search Space**
learning_rate	0.01–0.5
max_depth	3–10
min_child_weight	1–10
subsample	0.5–1
colsample_bytree	0.5–1
gamma	0–10
alpha	0–10
lambda	0–10

points are retained to To ensure a sufficient data amount for training and testing the XGB models, only the geographical windows with over 6000 available data points are retained. Consequently, 34 out of 1190 oceanic windows have been excluded. These windows located between 47.5°W–122.5°E and 52.5°S–57.5°S in the Southern Ocean (Fig. 2) contain fewer than 6000 valid samples due to the screening for N_d retrievals. For each model, the hyperparameters are tuned by implementing Bayesian optimization, which uses a Gaussian process prior distribution over hyperparameters to initialize a probabilistic model for the objective function to be optimized. After the initialization, the probabilistic model is updated iteratively and Bayesian optimization suggests the optimal combination of hyperparameters to try for the next iteration according to the previous one and samples gathered from the search space (Table 2) (Snoek et al., 2012). Each iteration is evaluated by 5-fold cross-validation using root mean square error (RMSE) as score. The number of boosting rounds (the number of trees) for each XGB model is then determined by the early stopping technique to further avoid overfitting i.e. the training of the model will stop early once it is monitored that the score of cross-validation does not improve within 20 iteration rounds.

2.3 Explaining the machine learning models

2.3.1 SHapley Additive exPlanation (SHAP) values

SHAP values were proposed by Lundberg and Lee (2017) on the basis of cooperative game theory to explain the outputs of ML models. It provides a novel and model-agnostic way to measure the relative contributions of N_d (as a surrogate for aerosols) and meteorological factors to CLF changes. The SHAP approach has been implemented with XGB in Python and it has been reported that outputs from XGB models with various number of trees can be well explained by the SHAP framework in different subject areas (e.g. Padarian et al., 2020; Lundberg et al., 2018, 2020; Kim et al., 2021) (e.g. Padarian et al., 2020; Lundberg et al., 2018, 2020; Kim et al., 2021; Li et al., 2022). The contribution of a predictor value for a specific "local" prediction is calculated as the difference between the predictions of the model in the presence and absence of this specific predictor for all possible predictor value combinations, and hence it is possible to gain insights into how a certain model outcome is achieved besides global feature importance.

SHAP values can be positive and negative (i. e. positive/negativecontribution to the final model prediction) . The final model prediction is obtained by adding the sum of all SHAP values to the base value (the mean of all model predictions) The base value in the context of SHAP values is typically computed as the average of all predictions by ML models over the entire training data points. Positive (negative) SHAP values indicate that the specific feature value increases (decreases) the prediction compared to this base value. In other words, the base value serves as the reference point against which the contributions of individual features are measured. SHAP values for all features will always sum up to the difference between the base value and the final model prediction so that SHAP values are additive and internally consistent. Furthermore, the quantification of the influence of meteorology on the N_d–CLF relationship can be analysed using SHAP interaction values, which are an extension of SHAP values. They measure the difference between the SHAP values for a feature when another (secondary) feature is included versus when it is not included, offering a potential tool for insights into feature interactions captured by the tree ensembles. SHAP values have been applied to study atmospheric aerosols in the context of air pollution Stirnberg et al. (2021), and have been

used by Zipfel et al. (2022) to explore satellite-observed N_d-LWP relationship in MBLCs in the Southeast Atlantic, finding that meteorological variables have considerable influences on the N_d-LWP relationship using SHAP interactive values. Moreover, the use of SHAP interaction values in these studies allows for a more profound and in-depth comprehension of the underlying processes with respect to local meteorology. ~~This state-of-art~~ SHAP values provide insights into the behaviour of the XGB models, and as all statistical/ML models, they may not necessarily reflect real-world physical causality. Nevertheless, this state-of-the-art technique allows us to account for meteorological covariations when deriving sensitivities and to appraise to what extent the meteorological predictors interact with and influence the N_d–CLF relationship beyond traditional global-level feature attributions.

2.3.2 Quantification of sensitivities and interactive effects

Figure. 1 is an exemplary graph for a regional XGB model at a specific 5°× 5°window (27.5°S–32.5°S, 122.5°W–127.5°W). SHAP values and SHAP interaction values are used to explain this XGB model and to quantify and isolate the CLF sensitivity to N_d and the interactive effects of meteorological factors (here ~~SST~~sea-surface temperature (SST)). Each dot in Fig. 1 represents an individual data instance (i.e. a single observation at a specific grid cell and time step) and shows how individual N_d values impact the CLF prediction. Figure 1 (a) illustrates that increased N_d values lead to an increase in the predicted CLF. The ~~isolated~~ CLF sensitivity to N_d is estimated as the slope of the linear regression between SHAP values and ln N_d (0.098 ~~).~~ CLF σ^{-1}). Note that all the input data have been standardized and thus sensitivities are expressed with the unit of cloud fraction change per standard deviation (CLF σ^{-1}). The vertical dispersion around the N_d–CLF relationship captured by the SHAP dependence plot is due to the dependence of the N_d contribution to the predicted CLF on meteorological factors (e.g. SST) in the model, which is captured by SHAP interaction values, as displayed in Fig. 1 (b). The colouring of the data points by SST illustrates how interactions with SST split up the N_d–CLF relationship, with low SST values amplifying the N_d contribution and vice versa. To quantify this interaction effect, the meteorological data are then divided into a group of above-average feature values and a group of below-average feature values. A linear regression is fit to the N_d values and the SHAP interaction values in each group. An interaction index (IAI) is derived from these regression fits ~~,~~and defined as the ~~difference between the slopes~~ slope for the high-value group (> mean) ~~and~~ minus the slope for the low-value group (< mean)~~, with IAI~~ :

$$IAI = \beta_{x,high} - \beta_{x,low} \tag{2}$$

where β is the slope of the linear regression between SHAP interaction values and ln N_d values, and the subscripts denote the high-value group and the low-value group for a specific meteorological variable x (SST in the example), respectively. At the exemplary geographical window, the influence of SST on N_d–CLF sensitivity is quantified by IAI = ~~-0.029 for SST at the examplegeographical window~~ -0.029 CLF σ^{-1} (Fig. 1 (b)). Similar to sensitivities, the unit of IAIs is also CLF σ^{-1}. Therefore, for a positive sensitivity such as the N_d–CLF sensitivity shown in Fig. 1 (a), a negative IAI value means that the N_d–CLF sensitivity is larger with low feature values, as shown in Fig. 1 (b) (the positive relationship is weakened by high SST values). On the contrary, a positive IAI value is corresponding to a larger positive sensitivity with high feature values. ~~Note that all~~

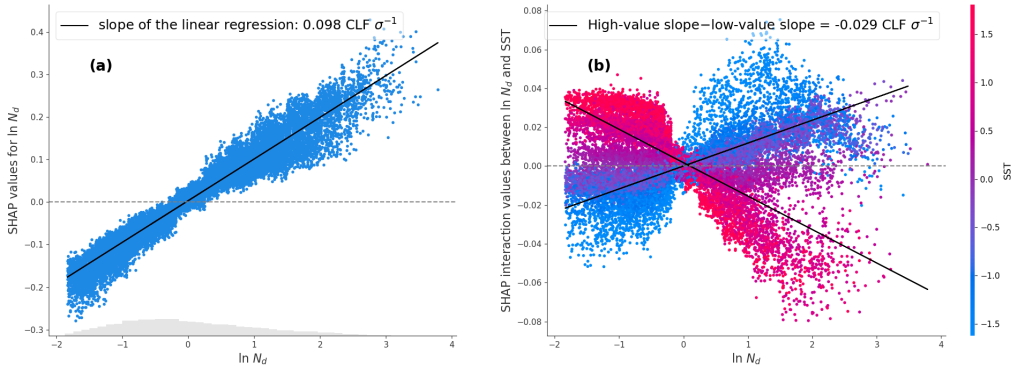

Figure 1. SHAP dependence plots of cloud droplet number concentration (N_d) in the region from 27.5°S to 32.5°S and from 122.5°W to 127.5°W. (a) SHAP values showing ln N_d–CLF relationship with the corresponding sensitivity defined as the slope of linear regression. (b) SHAP interaction values coloured by sea surface temperature (SST) showing the dependence of ln N_d–CLF relationship on the interactive effects of SST. The interaction values are further divided into two groups by the mean feature value of SST. Linear regressions are performed respectively for the high-value group and low-value group and the Interaction Index (IAI) is defined as the slope for the high-value group subtracting the slope for the low-value group. The horizontal dashed line is a demarcation between negative and positive SHAP (interaction) values.

~~the input data have been standardized and thus sensitivities and IAIs are expressed with the unit of cloud fraction change per standard deviation (CLF σ^{-1}).~~

3 Results and discussion

3.1 Model performance

The skills of the region-specific XGB models in predicting CLF are evaluated by the coefficient of determination (R^2) on the unseen hold-out test data. The global weighted mean R^2 is 0.45 (about 45 % on weighted average and up to 73.57 % of the variability in CLF prediction is explained) and a standard deviation of 0.10. While this means that on average, about half of the variability in CLF cannot be explained by the machine learning models, this is expected as previous studies have shown that the performance of statistical models decreases when going from monthly to daily data (Andersen et al., 2017; Fuchs et al., 2018; Dadashazar et al., 2021), and the performance is on par with that reported by Dadashazar et al. (2021), who used machine learning models to predict N_d with daily reanalysis data. The models in tropical regions in the Indian Ocean and the Western Pacific relatively poorly explain the variability in CLF, while XGB models perform well in the stratocumulus regions in the subtropics near the continents, and in the midlatitudes, particularly the southern hemispheric midlatitudes. The high skill of predicting CLF in the southern hemispheric midlatitudes is in contrast to a recent study where this region has been found to be particularly difficult to model statistically with monthly data ~~(?)~~(Andersen et al., 2023). In this region, the day-to-day CLF

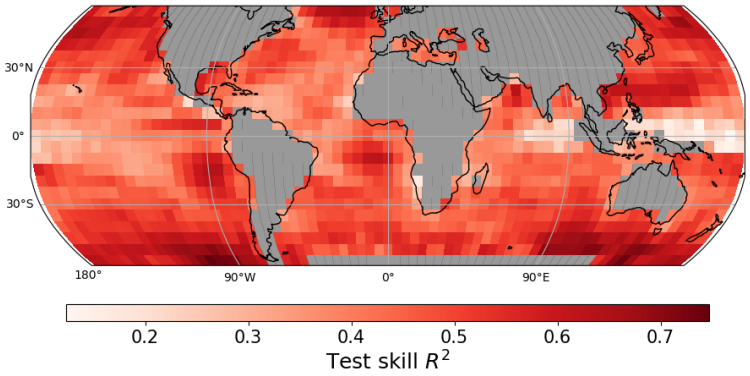

Figure 2. R^2 score of regional Extreme Gradient Boosting models predicting cloud fraction of marine boundary layer clouds in the independent test data set (2017–2019).

variability is high due to the large influence of synoptic-scale weather systems, and hence data at the daily resolution is more adequate to represent the CLF variability in these regions.

3.2 CLF sensitivity: global perspectives and regional characteristics

3.2.1 Global overview of CLF sensitivities

Figure 3 summarises the means and distributions of the near-global sensitivities of CLF to all predictors. The sensitivities are estimated as described in Sect~~2.3.2.~~ 2.3.2. The sequence is sorted by descending mean values of the absolute sensitivities (i.e. by feature importance) of the predictor variables. A strong and consistently positive N_d–CLF sensitivity is found. The fact that CLF is most sensitive to N_d is to be expected, as cloud observations from the same sensor are more directly related than a reanalysis product, so that their overall magnitude should not be compared (Zipfel et al., 2022). The entrainment of relatively dry air from the free troposphere into the MBL will be impeded by a stronger inversion (i.e. higher EIS), resulting in a shallower, better-mixed and more humid MBL conducive to stratocumulus clouds (Bretherton and Wyant, 1997; Wood and Hartmann, 2006; Qu et al., 2015a; Myers et al., 2021). The salient positive sensitivity to EIS is in accordance with the links found in previous studies (e.g. Klein and Hartmann, 1993; Qu et al., 2015b; Andersen et al., 2017) suggesting that EIS is a crucial controlling factor for low-marine cloud cover. Note that in some studies the strength of the inversion over the boundary layer is measured by lower tropospheric stability, which can be regarded as a similar metric outperformed by EIS (Wood and Bretherton, 2006). ~~PF~~ Precipitation fraction is the fraction of the original ERA5 grid box covered by large-scale precipitation. The strong positive CLF sensitivity to ~~PF~~ precipitation fraction is likely caused by the ML model learning that precipitation can be viewed as a proxy for cloudiness, rather than being an indicator of the physical processes via which precipitation exerts controls on the macrophysics of MBLCs. Humidity shows positive CLF sensitivities greater at 850 hPa, where cloud tops are often located (Gryspeerdt and Stier, 2012), than at 700 hPa which is typically in the free troposphere above the MBLCs (Myers and Norris, 2013). Likewise, the atmospheric temperature at 850 hPa (t_{850}) presents stronger CLF sensitivity than the

temperature at 700 hPa (t_{700}). Nonetheless, in the case of winds the 700 hPa pressure level is more relevant than that at 850 hPa. A relatively pronounced negative sensitivity to the eastward wind component at 700 hPa (u_{700}) indicates that clouds are depleted due to more westerlies at this level. CLF exhibits negative sensitivities to vertical pressure velocities both at 850 and 700 hPa, showing that large-scale ascending motion is connected to increases in MBLCs (Myers and Norris, 2013; Bretherton et al., 2013; Blossey et al., 2013). In general, the global averages of CLF sensitivity in terms of dynamical predictors (i.e. 3-D winds at surface and pressure levels) vary in sign and are less strong. A marked negative sensitivity of CLF to SST is found, which is in agreement to many prior studies (e.g. Qu et al., 2015b; Scott et al., 2020), where increases in SST have been found to lead to low cloud breakup and dissipation due to a number of processes as described in e.g. Scott et al. (2020). One of these is that the associated enhancement of ~~LHF~~ mean surface latent heat flux (LHF) deepens MBL and facilitates buoyancy and thus the entrainment of dry free-tropospheric air (Rieck et al., 2012; Andersen et al., 2022). However, CLF is much less sensitive to LHF than to SST, which may indicate that this mechanism is less important at the spatial and time scales considered in this study. CLF exhibits a considerable negative sensitivity to ~~SHF~~mean surface sensible heat flux (SHF), which quantifies an increase in CLF with increasing SHF (upward SHF are negative). While increased SHF can promote the transition from decks of stratus or stratocumulus clouds (high CLF) to more convective clouds (low CLF) due to the deepening of the boundary layer (Fan et al., 2016) potentially leading to a positive SHF–CLF relationship, increased SHF is associated to situations of cold air advection where turbulent surface fluxes are enhanced, leading to marked increases in CLF (Miyamoto et al., 2018; Zelinka et al., 2018; Grise and Kelleher, 2021)

3.2.2 Spatial patterns of the CLF sensitivity to N_d

The sensitivity of MBLC fraction associated with the aerosol proxy N_d are ubiquitously positive, in accordance with the global correlations or sensitivities found in (e.g. Gryspeerdt et al., 2016; Andersen et al., 2017). This is presumably due to the lifetime effect, but could also partially result from N_d retrieval biases discussed in ~~See~~Sect. 2.1. The global weighted mean value of the N_d–CLF sensitivity is 0.074 CLF σ^{-1} with a standard deviation of 0.036 CLF σ^{-1}. CLF is particularly sensitive to N_d in the regions of frequent stratocumulus to cumulus transition off the western continental coasts. These marked positive N_d–CLF sensitivities may be caused by high N_d delaying the transition from stratocumulus to cumulus clouds (Gryspeerdt et al., 2016; Christensen et al., 2020). The N_d–CLF sensitivity is also pronounced in the southern hemispheric midlatitudes, where stratiform clouds dominate. The N_d–CLF sensitivity is weak and close to 0 in the tropics, in particular in the deep convective warm pool region. These spatial patterns of N_d–CLF sensitivity resemble those found by Gryspeerdt et al. (2016), in particular the ones where they mediated the AOD–CLF relationship by N_d, but are more pronounced in the southern hemispheric midlatitudes. This difference in estimated sensitivity seems noteworthy and should thus be investigated in future work. As N_d retrievals tend to negatively bias at lower CLF and positively bias at higher CLF, the N_d–CLF sensitivity may be overestimated, and at the scales considered here, should be interpreted as an upper bound to the physical N_d–CLF sensitivity.

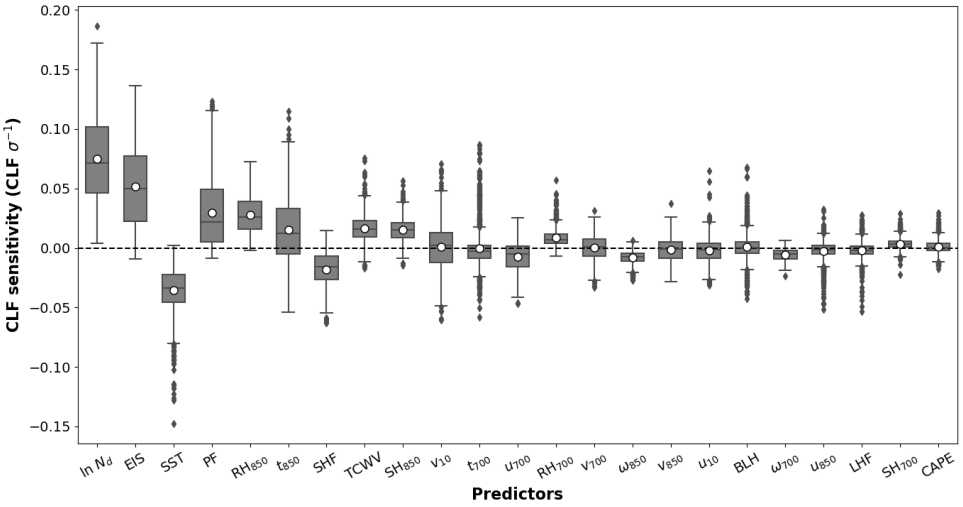

Figure 3. The distribution of the sensitivities of cloud fraction to all predictors as depicted in Table 1. Boxes represent the interquartile range which is extended by whiskers up to 1.5 interquartile ranges, with outliers shown as the points outside the range. The solid line and white dot in each box show the median and mean values of the sensitivities, respectively. Predictors are sorted by the mean values of absolute sensitivity values. The dashed line across the figure differentiates positive and negative sensitivity values.

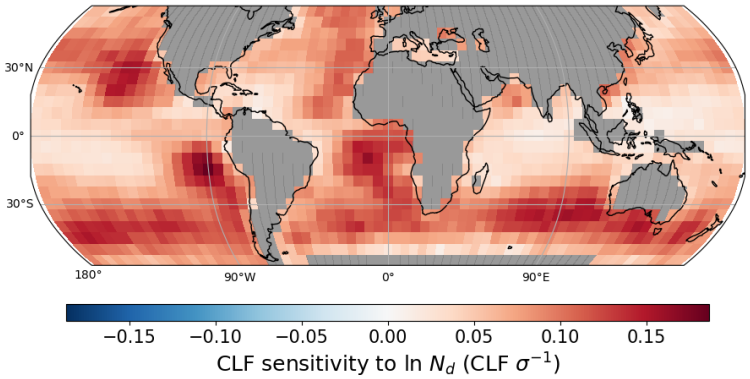

Figure 4. Sensitivity of marine boundary layer cloud fraction to ln N_d.

3.2.3 Spatial patterns of the CLF sensitivity to thermodynamical drivers

There has been a strong consensus that EIS and SST are the two important determinants of cloud fraction of marine boundary clouds and their corresponding radiative effects across different geographical regions and on varying time scales (e.g. Bretherton, 2015; Myers and Norris, 2015; McCoy et al., 2017; Wall et al., 2017). Stronger inversions capping MBL (i.e. higher EIS) will hamper the entrainment of aloft dry air from the troposphere and thus lead to a shallower MBL and more moisture trapped

within MBL, promoting the development and maintenance of low-level clouds (Andersen et al., 2017). The regional EIS–CLF sensitivity patterns (Fig. 5(a)) show that marine low cloud fraction increases ubiquitously in response to stronger EIS, in particular in the tropical and subtropical stratocumulus-capped regions and within the midlatitudes. The sensitivity pattern is in good agreement with that found by Scott et al. (2020) and **?**, related studies at different time scales (Grise and Medeiros, 2016; Kelleher and Grise, 2019; de Szoeke et al., 2016).

MBLC cover reduces globally in response to increased SST, particularly pronounced in the stratocumulus regions over eastern oceanic basins (Fig. 5(b)), consistent well with (Scott et al., 2020). SST can favour MBLC dissipation through increasing surface latent heat fluxes and deepening MBL, facilitating dry entrainment and eventually desiccating the MBL and clouds (Rieck et al., 2012; Qu et al., 2015b). Yet as stated in Sect. 3.2.1, the weak CLF sensitivity to LHF in relation to the strong sensitivity to SST may imply that the other process makes more substantial contributions, namely that the higher moisture gradient between the troposphere and MBL arising from the increased SST makes the entrained air more efficient in evaporating cloud water (van der Dussen et al., 2015; Qu et al., 2015b). This process has been shown to be the driving mechanism for the observed reduction in marine low cloud cover near the coast of Baja California (Andersen et al., 2022).

Figure. 5 (c) shows that low marine cloud fraction increases with negative (upward) SHF most markedly in the stratocumulus regions. CLF can increase in response to increased surface fluxes in situations of cold advection (Zelinka et al., 2018). Over the South Indian Ocean, a marked SHF–CLF sensitivity is also found. Here, enhancements of SHF due to the subtropical anticyclone and midlatitude storm-track activity have been found to increase CLF (Miyamoto et al., 2018). The results suggest that the increase of CLF due to increased SHF (e.g. due to cold advection) outweighs the influence of SHF on CLF by controlling the transition from marine stratocumulus to open-cellular marine clouds (Kazil et al., 2014; Fan et al., 2016) in the core stratocumulus regions. Consequently, the SHF–CLF sensitivity is less pronounced in regions of frequent closed- to open-cell and cumulus transitions. ~~RH_{850}~~ Relative humidity at 850 hPa (RH_{850}) is positively related to marine low liquid cloud fraction across the globe. The positive sensitivity is particularly strong in the trade cumulus regions where the 850 hPa level is representative of the boundary layer. In the coastal stratocumulus regions, clouds are frequently below this level (Adebiyi and Zuidema, 2016), so that clouds are not as sensitive to variability in RH at that level.

3.2.4 Spatial patterns of the CLF sensitivity to dynamical drivers

Large-scale circulations and dynamical conditions play an essential role in controlling cloud fraction and the indirect effects of aerosols (Su et al., 2010; Small et al., 2011). The large-scale dynamics are represented by the horizontal and vertical winds at 700 hPa and 850 hPa, which display clear and distinct regional patterns of CLF sensitivity (Fig. 6). It can also be seen that at the considered scales and pressure levels, horizontal wind vectors have stronger CLF sensitivities than large-scale vertical motion. There is a coherent pattern of negative CLF sensitivity to the zonal wind at 700 hPa in the stratocumulus-dominated regions (also apparent at 850 hPa), and the southern hemispheric midlatitudes, indicating a decrease in MBLCs with westerly anomalies at this pressure level. Recently, a study using monthly data has also found a similar sensitivity pattern of stratocumulus clouds to zonal wind at 700 hPa, finding that the reduced CLF is related to increased vertical wind shear (as the boundary layer flow is easterly), leading to increased turbulence and dry-air entrainment (**?**). Using monthly data, **?** did not find a similar CLF

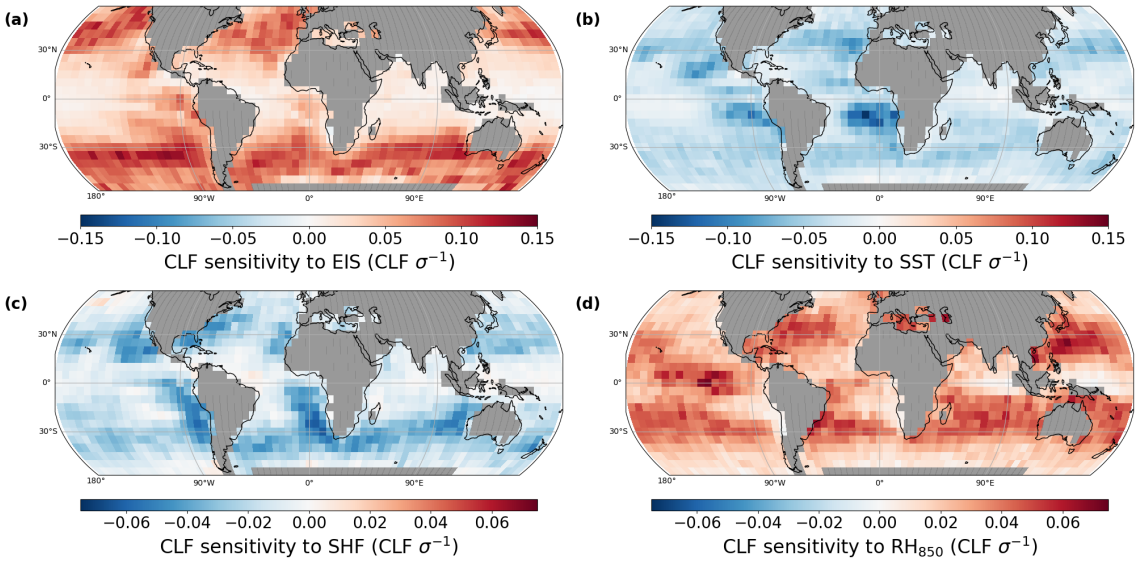

Figure 5. Sensitivity of marine boundary layer cloud fraction to the estimated inversion strength (EIS), sea surface temperature (SST), sensible heat flux (SHF) and relative humidity at 850 hPa (RH_{850}). Note that the range of colourbars of SHF and RH_{850} (-0.075–0.075) are narrower than EIS and SST (-0.15–0.15).

sensitivity to zonal winds in the southern hemispheric midlatitudes, though. As the CLF sensitivity to u_{700} in the southern hemispheric midlatitudes is only apparent using daily data and only at 700 hPa, it seems likely that it is related to synoptic variability that drives day-to-day variability in MBLCs in this region (Kelleher and Grise, 2019). Positive CLF sensitivities to u_{700} (higher CLF with westerly anomalies) and to a lesser degree u_{850} are found off the eastern Asian and North-American continents. CLF increases due to cold-air outbreaks in NW Atlantic and NW Pacific may be the reason for these positive sensitivities. Cold-air outbreaks occur during winter as cold continental air moves over warmer SSTs, increasing moisture and heat fluxes into the MBL so that the formation of MBLCs is favoured (Young et al., 2002). This leads to wintertime maxima in CLF in these regions (Yuan and Oreopoulos, 2013).

The sensitivity of CLF to the meridional winds at 700 hPa exhibits two bands straddling the subtropical regions between about 15 °and 35 °in both hemispheres but opposite in sign (positive in the Northern Hemisphere and negative in the Southern Hemisphere), illustrating that in these regions the poleward winds are associated with an increase in low cloud fraction. The bands are still apparent at 850 hPa, while the negative band in the Southern Hemisphere extends northward to tropical areas. These hemispheric sensitivity bands to ~~v_{700}~~ the v wind component at 700 hPa closely resemble those found in **?**, with their analysis suggesting that the poleward winds on the eastern side of midlatitude cyclones may be related to warm and moist advection, increasing CLF. However, they also find a strong correlation of these free-tropospheric poleward winds with large-scale ascending air motion making the assertion of causality difficult. Poleward winds are also found to decrease CLF over the southern hemispheric midlatitudes.

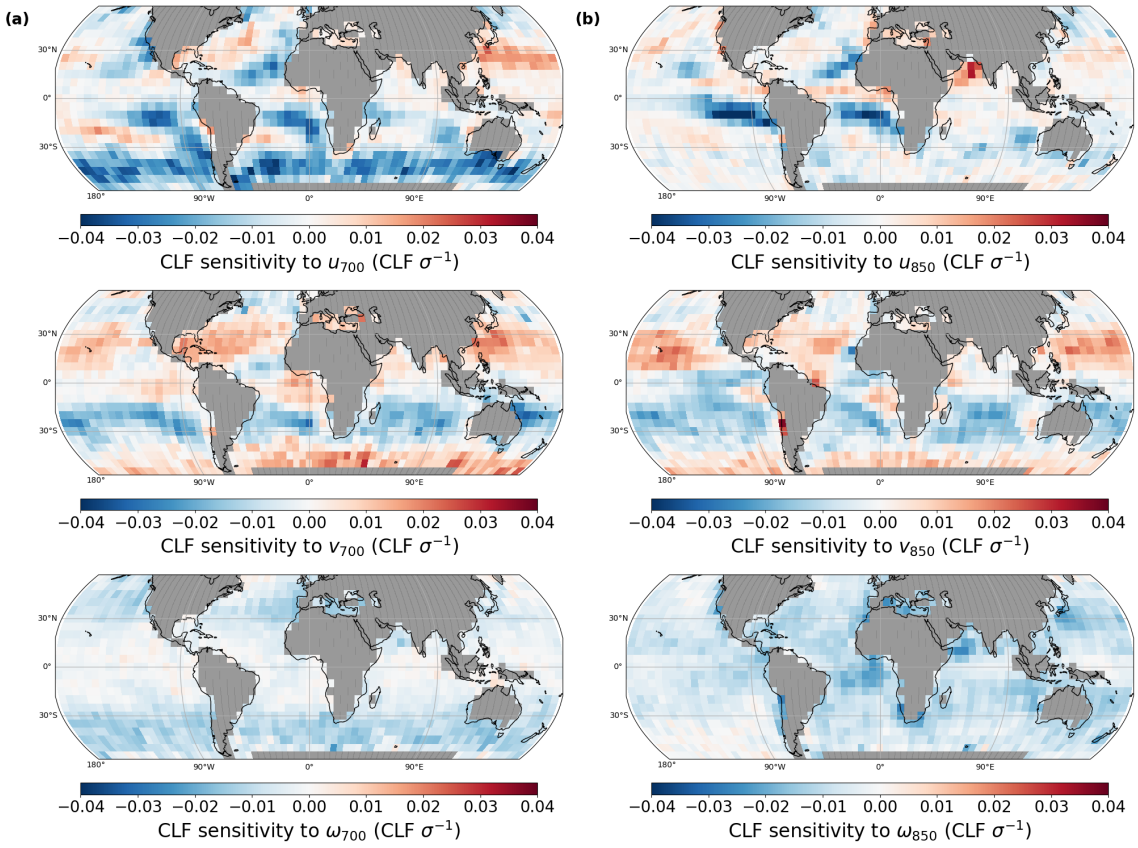

Figure 6. Sensitivity of cloud fraction to u, v wind component vectors and vertical velocities at 700 hPa (column (a)) and 850 hPa (column (b)). Note that the range of the colourbar is in general smaller (-0.04–0.04) than in Fig. 5.

CLF is negatively connected to the vertical pressure velocity both at 700 hPa and 850 hPa (ω_{700} and ω_{850}) over the entire Earth, indicating that ascending large-scale air motion enhances the cover of MBLCs globally. It is shown in the bottom of Fig. 6 column (a) that the CLF sensitivity to ~~ω at 700 hPa~~ ω_{700} is larger in the midlatitude ocean basins, whereas the CLF sensitivity to ~~ω at 850 hPa~~ ω_{850} is larger in the subtropical oceans where subsidence is climatologically prevalent (Myers and Norris, 2015, 2016; Scott et al., 2020). This seems indicative of CLF being most sensitive to large-scale ascending motion at the typical altitude of the clouds. It is interesting to note that between 30°N and 30°S, no marked CLF sensitivity to ~~ω at 700 hPa~~ ω_{700} is found, contrasting the finding of enhanced subsidence at this level reducing MBLCs by Myers and Norris (2013). This effect is likely better described in the ω_{850} data which is more related to the altitude of the cloud top.

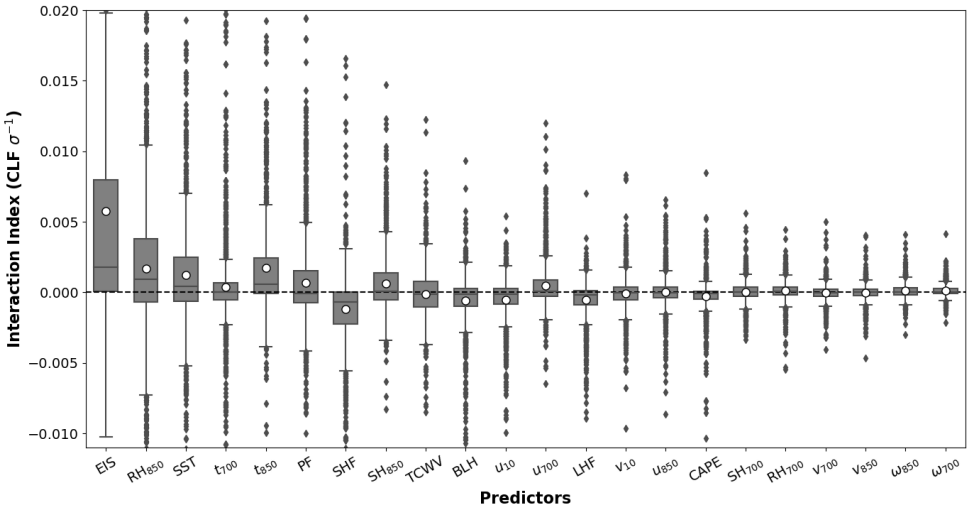

Figure 7. Similar to Fig. 3 but for the interaction effect of N_d with all environmental parameters, quantified by the Interaction Index (CLF σ^{-1}).

3.3 Dependence of N_d–CLF relationship on meteorology

3.3.1 Global overview of the Interaction Indices

In this section, we use the IAI as defined in Sect. 2.3.2 to quantitatively show how the response of MBLC fraction attributed to the aerosol proxy N_d varies with the meteorological factors. As discussed in Sect. 2.3.2, since the sensitivity related to N_d is positive across the globe (Fig. 5 (d)), a positive IAI can be interpreted as an amplification of the N_d–CLF sensitivity with high (above-average) feature values of a meteorological variable, whereas a negative IAI signifies an amplification of the sensitivity at low feature values.

In Fig. 7, analogous to Fig. 3, the features along the x-axis are arranged in descending order based on their averaged absolute IAIs i.e. by the strength of the impact of each meteorological feature on the N_d–CLF sensitivity. Similar to the feature importance summarized by Fig. 3, EIS, SST, RH_{850} and SHF have relatively large strength of interaction effect and thus can be regarded as critical controlling factors not only for marine low cloud cover but also for their response to changes in N_d (and in extension aerosols). Compared to the CLF-sensitivities, the IAIs associated with atmospheric temperatures at 700 and 850 hPa have greater strengths. Furthermore, it can also be seen that the vertical and horizontal winds at the surface and different pressure levels are ranked generally lower. In general, the thermodynamical factors have a stronger influence on the N_d–CLF sensitivity than the dynamical factors.

3.3.2 Spatial patterns of the Interaction Indices

Coherent and distinct spatial distributions of the impact of selected meteorological parameters on the N_d–CLF relationship can be observed. Hereafter we show the regional characteristics of the interaction effects of EIS and SST which are the two most important meteorological factors for CLF in MBLCs and have the greatest absolute strengths of IAI. EIS exerts the most noticeable positive IAIs over the midlatitude oceanic areas (Figure. 8 (a)), reflecting that stronger temperature inversions capping the MBL over these regions will amplify the positive N_d–CLF relationship. The results suggest that in these regions, through hampering the entrainment of drier air from the free troposphere, the stronger inversion and more stable conditions are capable of trapping more moisture within a shallower MBL and could thus weaken the evaporation-entrainment feedback, ultimately favouring a more positive N_d–CLF relationship (Chen et al., 2014; Christensen et al., 2020). It is interesting to note that these interactions are not apparent in the stratocumulus regions where EIS is a strong control of CLF, and in the stratocumulus-to-cumulus transition regions, where Christensen et al. (2020) found the aerosol effect on this transition to be confined to stable atmospheric conditions. This may suggest that the suggested entrainment effect is dependent on the EIS, and stronger at slightly lower EIS values typically found in the midlatitudes (Scott et al., 2020). The observed impact of EIS on the N_d–CLF relationship found in the midlatitudes also has implications within the context of climate change. While in the subtropics, global climate models predict an increase in EIS with a warming climate, in the midlatitudes EIS is predicted to decrease (Myers et al., 2021), potentially decreasing the sensitivity of CLF to N_d there.

Fig. 8 (b) shows that higher SSTs are found to amplify the positive N_d–CLF relationship (positive IAI) in the regions of frequent stratocumulus-to-cumulus transition (Cesana and Del Genio, 2021). Here, higher SSTs tend to lead to the transition from stratocumulus clouds to shallow convective clouds (Cesana et al., 2019), however, this transition has been found to be delayed when aerosol is increased (Goren et al., 2019; Christensen et al., 2020). The positive IAIs in these transition regions may thus point to increased control of N_d on CLF at higher SST values, as these are the situations where transitions typically occur and when increased N_d can act to delay this transition. In these regions, higher SSTs in the future may thus increase the sensitivity of MBLC CLF to aerosols. It should be noted that the quantification of the dependence of the N_d–CLF relationship on meteorological factors (EIS, SST discussed in this section) is also likely subject to the biases in the N_d–CLF sensitivity caused by the N_d retrieval biases as a function of CLF. This would potentially contribute to the non-causal facets of the relationships and interactive effects quantified by SHAP values.

4 Conclusions

In this study, nine years (2011-2019) of daily satellite and reanalysis data have been analyzed to better understand the effect of N_d on CLF in MBLC, and its dependence on meteorological factors. We have established a near-global machine learning framework to predict the cloud fraction of marine boundary clouds using regionally-specific XGB regression models. The explainable machine learning technique of SHAP regression values has been used to explain the regional XGB models, to quantify the CLF sensitivity to all cloud controlling factors with a specific focus on N_d, moreover, to quantify the meteorological influence on the N_d–CLF relationship at a global scale. The main findings of this study are stated as follows:

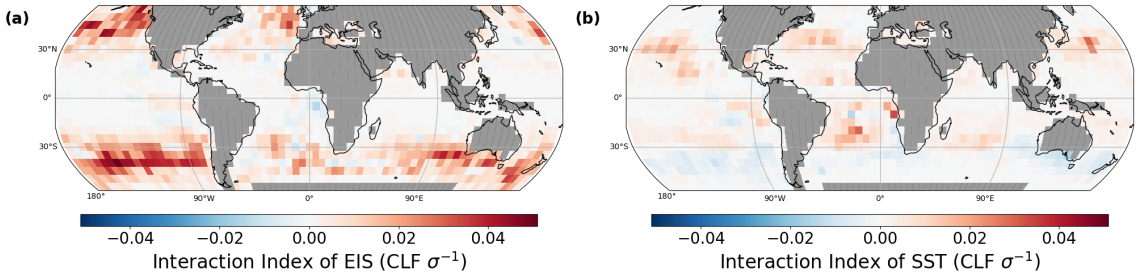

Figure 8. Patterns of the Interaction Index showing the dependence of the N_d–CLF relationship on estimated inversion strength (EIS) (a) and sea surface temperature (SST) (b).

1. Marine boundary layer cloud fraction is the most sensitive to N_d (surrogate for aerosols) in the regions of stratocumulus to cumulus transition, which may arise from the high N_d delaying this transition. The N_d–CLF sensitivity in the southern hemispheric midlatitudes is higher than in previous studies, which should be investigated in future work. The ubiquitous positive sensitivity reveals that aerosol has a considerable impact on MBL cloudiness, even though a positive retrieval bias of N_d at high CLF on the estimated N_d–CLF sensitivity is likely to lead to an overestimation.

2. Consistent with the literature, EIS and SST are two important determinants for marine low clouds by regulating surface fluxes and dry-air entrainment processes. In addition, strong negative CLF sensitivity and spatial patterns for SHF are also found, suggesting that the effect of cold air advection may surpass the SHF enhancement of closed- to open-cell and cumulus transitions. Dynamic drivers (meridional and zonal winds) indicate that midlatitude synoptic-scale disturbances and vertical wind shear make considerable contributions to marine low cloud amounts.

3. In general, thermodynamical parameters exert a more important influence on the N_d–CLF relationship than dynamical parameters. EIS, RH_{850}, SST, temperatures at 700 and 850 hPa have the strongest effect on the N_d–CLF sensitivity. In the midlatitudes, higher EIS is found to amplify the positive N_d–CLF sensitivity which may be related to a reduced entrainment feedback in these conditions. Whereas higher SST is found to amplify the N_d–CLF sensitivity in stratocumulus-to-cumulus transition regions potentially because the transition induced by higher SSTs may be delayed by increased N_d. These findings have implications for possible future changes in the sensitivity of CLF to aerosols.

4. For the dynamical and thermodynamical factors shown here, both CLF sensitivities and the interactive effects (dependence of N_d–CLF relationship on meteorology) exhibit distinct regional patterns. These coherent spatial patterns indicate that the proposed explainable machine learning framework is not only capable of skillfully predicting CLF for marine low clouds but also captures regional characteristics of the relation between CLF and N_d, and meteorological influences in a physically meaningful way.

In the future, the observation-based sensitivities and interactive effects quantified by the ML framework here will be compared to those in ESMs, which have the potential to evaluate ESM parameterizations related to ACI and even help gain insights into how the models could be tuned in this respect.

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

Data availability. All data sets used in this study are publicly available. The MODIS data set (https://dx.doi.org/10.5067/MODIS/MOD08_M3.061) was acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) (https://ladsweb.modaps.eosdis.nasa.gov/, last access: 09 June 2021); The hourly reanalysis data at single levels (https://doi.org/10.24381/cds.adbb2d47) and pressure levels (https://doi.org/10.24381/cds.bd0915c6, last access: 22 Feb 2023) are obtained from the Copernicus Climate Change Service (C3S) Climate Date Store: https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset.

Author contributions. HA and JC designed the initial research idea. YJ, HA and JC developed the study concept and methodology. YJ and HA obtained and analysed the data sets. YJ implemented the explainable machine learning framework, performed visualization and wrote the original draft. All authors contributed to interpreting the results, reviewing and improving the manuscript.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. The (co-) authors have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 821205 (FORCeS) and the Deutsche Forschungsgemeinschaft (DFG) in the project Constraining Aerosol-Low cloud InteractionS with multi-target MAchine learning (CALISMA), project number 440521482. We thank two anonymous reviewers whose helpful comments contributed to improving the manuscript.

References

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014–1017, https://doi.org/10.1038/nature03174, 2004.

Adebiyi, A. A. and Zuidema, P.: The role of the southern African easterly jet in modifying the southeast Atlantic aerosol and cloud environments, Quarterly Journal of the Royal Meteorological Society, 142, 1574–1589, https://doi.org/https://doi.org/10.1002/qj.2765, 2016.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

Andersen, H. and Cermak, J.: How thermodynamic environments control stratocumulus microphysics and interactions with aerosols, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/2/024004, 2015.

Andersen, H., Cermak, J., Fuchs, J., and Schwarz, K.: Global observations of cloud-sensitive aerosol loadings in low-levelmarine clouds, Journal of Geophysical Research, 121, 936–12, https://doi.org/10.1002/2016JD025614, 2016.

Andersen, H., Cermak, J., Fuchs, J., Knutti, R., and Lohmann, U.: Understanding the drivers of marine liquid-water cloud occurrence and properties with global observations using neural networks, Atmospheric Chemistry and Physics, 17, 9535–9546, https://doi.org/10.5194/acp-17-9535-2017, 2017.

Andersen, H., Cermak, J., Zipfel, L., and Myers, T. A.: Attribution of Observed Recent Decrease in Low Clouds Over the Northeastern Pacific to Cloud-Controlling Factors, Geophysical Research Letters, 49, 1–10, https://doi.org/10.1029/2021gl096498, 2022.

Andersen, H., Cermak, J., Douglas, A., Myers, T. A., Nowack, P., Stier, P., Wall, C. J., and Wilson Kemsley, S.: Sensitivities of cloud radiative effects to large-scale meteorology and aerosols from global observations, Atmospheric Chemistry and Physics, 23, 10 775–10 794, https://doi.org/10.5194/acp-23-10775-2023, 2023.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. L., Dufresne, J. L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, 1–45, https://doi.org/10.1029/2019RG000660, 2020.

Bender, F. A., Frey, L., McCoy, D. T., Grosvenor, D. P., and Mohrmann, J. K.: Assessment of aerosol–cloud–radiation correlations in satellite observations, climate models and reanalysis, Climate Dynamics, 52, 4371–4392, https://doi.org/10.1007/s00382-018-4384-z, 2019.

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmospheric Chemistry and Physics, 17, 9815–9836, https://doi.org/10.5194/acp-17-9815-2017, 2017.

Beucler, T., Ebert-Uphoff, I., Rasp, S., Pritchard, M., and Gentine, P.: Machine Learning for Clouds and Climate, Clouds and Climate, pp. 1–27, 2020.

Blossey, P. N., Bretherton, C. S., Zhang, M., Cheng, A., Endo, S., Heus, T., Liu, Y., Lock, A. P., de Roode, S. R., and Xu, K.-M.: Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison, Journal of Advances in Modeling Earth Systems, 5, 234–258, https://doi.org/https://doi.org/10.1002/jame.20025, 2013.

Bretherton, C. S.: Insights into low-latitude cloud feedbacks from high-resolution models, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20140 415, https://doi.org/10.1098/rsta.2014.0415, 2015.

Bretherton, C. S. and Wyant, M. C.: Moisture Transport, Lower-Tropospheric Stability, and Decoupling of Cloud-Topped Boundary Layers, Journal of the Atmospheric Sciences, 54, 148–167, https://doi.org/https://doi.org/10.1175/1520-0469(1997)054<0148:MTLTSA>2.0.CO;2, 1997.

Bretherton, C. S., Blossey, P. N., and Uchida, J.: Cloud droplet sedimentation, entrainment efficiency, and subtropical stratocumulus albedo, Geophysical Research Letters, 34, https://doi.org/https://doi.org/10.1029/2006GL027648, 2007.

Bretherton, C. S., Blossey, P. N., and Jones, C. R.: Mechanisms of marine low cloud sensitivity to idealized climate perturbations: A single-LES exploration extending the CGILS cases, Journal of Advances in Modeling Earth Systems, 5, 316–337, https://doi.org/https://doi.org/10.1002/jame.20019, 2013.

Cesana, G., Del Genio, A. D., Ackerman, A. S., Kelley, M., Elsaesser, G., Fridlind, A. M., Cheng, Y., and Yao, M.-S.: Evaluating models' response of tropical low clouds to SST forcings using CALIPSO observations, Atmospheric Chemistry and Physics, 19, 2813–2832, https://doi.org/10.5194/acp-19-2813-2019, 2019.

Cesana, G. V. and Del Genio, A. D.: Observational constraint on cloud feedbacks suggests moderate climate sensitivity, Nature Climate Change, 11, 213–218, https://doi.org/10.1038/s41558-020-00970-y, 2021.

Chen, T. and Guestrin, C.: XGBoost: A scalable tree boosting system, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-Augu, 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Chen, Y., Haywood, J., Wang, Y., Malavelle, F., Jordan, G., Partridge, D., Fieldsend, J., De Leeuw, J., Schmidt, A., Cho, N., Oreopoulos, L., Platnick, S., Grosvenor, D., Field, P., and Lohmann, U.: Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover, Nature Geoscience, 15, 609–614, https://doi.org/10.1038/s41561-022-00991-6, 2022.

Chen, Y. C., Christensen, M. W., Stephens, G. L., and Seinfeld, J. H.: Satellite-based estimate of global aerosol-cloud radiative forcing by marine warm clouds, Nature Geoscience, 7, 643–646, https://doi.org/10.1038/ngeo2214, 2014.

Christensen, M. W. and Stephens, G. L.: Microphysical and macrophysical responses of marine stratocumulus polluted by underlying ships: Evidence of cloud deepening, Journal of Geophysical Research Atmospheres, 116, 1–10, https://doi.org/10.1029/2010JD014638, 2011.

Christensen, M. W., Neubauer, D., Poulsen, C. A., Thomas, G. E., McGarragh, G. R., Povey, A. C., Proud, S. R., and Grainger, R. G.: Unveiling aerosol–cloud interactions – Part 1: Cloud contamination in satellite products enhances the aerosol indirect forcing estimate, Atmospheric Chemistry and Physics, 17, 13 151–13 164, https://doi.org/10.5194/acp-17-13151-2017, 2017.

Christensen, M. W., Jones, W. K., and Stier, P.: Aerosols enhance cloud lifetime and brightness along the stratus-to-cumulus transition, Proceedings of the National Academy of Sciences of the United States of America, 117, 17 591–17 598, https://doi.org/10.1073/pnas.1921231117, 2020.

Dadashazar, H., Painemal, D., Alipanah, M., Brunke, M., Chellappan, S., Corral, A. F., Crosbie, E., Kirschler, S., Liu, H., Moore, R. H., Robinson, C., Scarino, A. J., Shook, M., Sinclair, K., Thornhill, K. L., Voigt, C., Wang, H., Winstead, E., Zeng, X., Ziemba, L., Zuidema, P., and Sorooshian, A.: Cloud drop number concentrations over the western North Atlantic Ocean: seasonal cycle, aerosol interrelationships, and other influential factors, Atmospheric Chemistry and Physics, 21, 10 499–10 526, https://doi.org/10.5194/acp-21-10499-2021, 2021.

de Szoeke, S. P., Verlinden, K. L., Yuter, S. E., and Mechem, D. B.: The Time Scales of Variability of Marine Low Clouds, Journal of Climate, 29, 6463–6481, https://doi.org/https://doi.org/10.1175/JCLI-D-15-0460.1, 2016.

Dey, S., Di Girolamo, L., Zhao, G., Jones, A. L., and McFarquhar, G. M.: Satellite-observed relationships between aerosol and trade-wind cumulus cloud properties over the Indian Ocean, Geophysical Research Letters, 38, https://doi.org/10.1029/2010GL045588, 2011.

Douglas, A. R. and L'Ecuyer, T.: Possible evidence of increased global cloudiness due to aerosol-cloud interactions, Atmospheric Chemistry and Physics Discussions, 2022, 1–22, https://acp.copernicus.org/preprints/acp-2022-688/, 2022.

Fan, J., Wang, Y., Rosenfeld, D., and Liu, X.: Review of aerosol-cloud interactions: Mechanisms, significance, and challenges, https://doi.org/10.1175/JAS-D-16-0037.1, 2016.

Forster, P. M., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: Chapter 7: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, 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, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., p. in press, Cambridge University Press, 2021.

Fuchs, J., Cermak, J., and Andersen, H.: Building a cloud in the southeast Atlantic: Understanding low-cloud controls based on satellite observations with machine learning, Atmospheric Chemistry and Physics, 18, 16 537–16 552, https://doi.org/10.5194/acp-18-16537-2018, 2018.

Ghan, S., Wang, M., Zhang, S., Ferrachat, S., Gettelman, A., Griesfeller, J., Kipling, Z., Lohmann, U., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., Wang, H., and Zhang, K.: Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability, Proceedings of the National Academy of Sciences, 113, 5804–5811, https://doi.org/10.1073/PNAS.1514036113, 2016.

Goren, T., Kazil, J., Hoffmann, F., Yamaguchi, T., and Feingold, G.: Anthropogenic Air Pollution Delays Marine Stratocumulus Breakup to Open Cells, Geophysical Research Letters, 46, 14 135–14 144, https://doi.org/https://doi.org/10.1029/2019GL085412, 2019.

Grandey, B. S. and Stier, P.: A critical look at spatial scale choices in satellite-based aerosol indirect effect studies, Atmospheric Chemistry and Physics, 10, 11 459–11 470, https://doi.org/10.5194/acp-10-11459-2010, 2010.

Grise, K. M. and Kelleher, M. K.: Midlatitude Cloud Radiative Effect Sensitivity to Cloud Controlling Factors in Observations and Models: Relationship with Southern Hemisphere Jet Shifts and Climate Sensitivity, Journal of Climate, 34, 5869–5886, https://doi.org/https://doi.org/10.1175/JCLI-D-20-0986.1, 2021.

Grise, K. M. and Medeiros, B.: Understanding the Varied Influence of Midlatitude Jet Position on Clouds and Cloud Radiative Effects in Observations and Global Climate Models, Journal of Climate, 29, 9005–9025, https://doi.org/https://doi.org/10.1175/JCLI-D-16-0295.1, 2016.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409–453, https://doi.org/10.1029/2017RG000593, 2018.

Gryspeerdt, E. and Stier, P.: Regime-based analysis of aerosol-cloud interactions, Geophysical Research Letters, 39, 1–5, https://doi.org/10.1029/2012GL053221, 2012.

Gryspeerdt, E., Quaas, J., and Bellouin, N.: Constraining the aerosol influence on cloud fraction, Journal of Geophysical Research, 121, 3566–3583, https://doi.org/10.1002/2015JD023744, 2016.

Gryspeerdt, E., Goren, T., Sourdeval, O., Quaas, J., Mülmenstädt, J., Dipu, S., Unglaub, C., Gettelman, A., and Christensen, M.: Constraining the aerosol influence on cloud liquid water path, Atmospheric Chemistry and Physics, 19, 5331–5347, https://doi.org/10.5194/acp-19-5331-2019, 2019.

Gryspeerdt, E., McCoy, D. T., Crosbie, E., Moore, R. H., Nott, G. J., Painemal, D., Small-Griswold, J., Sorooshian, A., and Ziemba, L.: The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data, Atmospheric Measurement Techniques, 15, 3875–3892, https://doi.org/10.5194/amt-15-3875-2022, 2022.

Hartmann, D. L., Ockert-Bell, M. E., and Michelsen, M. L.: The Effect of Cloud Type on Earth's Energy Balance: Global Analysis, Journal of Climate, 5, 1281–1304, https://doi.org/10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2, 1992.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Jiang, H., Xue, H., Teller, A., Feingold, G., and Levin, Z.: Aerosol effects on the lifetime of shallow cumulus, Geophysical Research Letters, 33, https://doi.org/https://doi.org/10.1029/2006GL026024, 2006.

Kaufman, Y. J. and Koren, I.: Smoke and Pollution Aerosol Effect on Cloud Cover, Science, 313, 655–658, https://doi.org/10.1126/science.1126232, 2006.

Kazil, J., Feingold, G., Wang, H., and Yamaguchi, T.: On the interaction between marine boundary layer cellular cloudiness and surface heat fluxes, Atmospheric Chemistry and Physics, 14, 61–79, https://doi.org/10.5194/acp-14-61-2014, 2014.

Kelleher, M. K. and Grise, K. M.: Examining Southern Ocean Cloud Controlling Factors on Daily Time Scales and Their Connections to Midlatitude Weather Systems, Journal of Climate, 32, 5145–5160, https://doi.org/https://doi.org/10.1175/JCLI-D-18-0840.1, 2019.

Kim, M., Brunner, D., and Kuhlmann, G.: Importance of satellite observations for high-resolution mapping of near-surface NO2 by machine learning, Remote Sensing of Environment, 264, 112 573, https://doi.org/10.1016/j.rse.2021.112573, 2021.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587–1606, https://doi.org/https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

Leahy, L. V., Wood, R., Charlson, R. J., Hostetler, C. A., Rogers, R. R., Vaughan, M. A., and Winker, D. M.: On the nature and extent of optically thin marine low clouds, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/https://doi.org/10.1029/2012JD017929, 2012.

Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., and Orth, R.: Widespread increasing vegetation sensitivity to soil moisture, Nature Communications, 13, 3959, https://doi.org/10.1038/s41467-022-31667-9, 2022.

Loeb, N. G. and Schuster, G. L.: An observational study of the relationship between cloud, aerosol and meteorology in broken low-level cloud conditions, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/https://doi.org/10.1029/2007JD009763, 2008.

Lundberg, S. M. and Lee, S. I.: A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems, 2017-Decem, 4766–4775, 2017.

Lundberg, S. M., Erion, G. G., and Lee, S.-I.: Consistent Individualized Feature Attribution for Tree Ensembles, ArXiv, abs/1802.0, http://arxiv.org/abs/1802.03888, 2018.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nature Machine Intelligence, 2, 56–67, https://doi.org/10.1038/s42256-019-0138-9, 2020.

McCoy, D. T., Eastman, R., Hartmann, D. L., and Wood, R.: The Change in Low Cloud Cover in a Warmed Climate Inferred from AIRS, MODIS, and ERA-Interim, Journal of Climate, 30, 3609–3620, https://doi.org/https://doi.org/10.1175/JCLI-D-15-0734.1, 2017.

Merk, D., Deneke, H., Pospichal, B., and Seifert, P.: Investigation of the adiabatic assumption for estimating cloud micro- and macrophysical properties from satellite and ground observations, Atmospheric Chemistry and Physics, 16, 933–952, https://doi.org/10.5194/acp-16-933-2016, 2016.

Miyamoto, A., Nakamura, H., and Miyasaka, T.: Influence of the Subtropical High and Storm Track on Low-Cloud Fraction and Its Seasonality over the South Indian Ocean, Journal of Climate, 31, 4017–4039, https://doi.org/https://doi.org/10.1175/JCLI-D-17-0229.1, 2018.

Myers, T. A. and Norris, J. R.: Observational Evidence That Enhanced Subsidence Reduces Subtropical Marine Boundary Layer Cloudiness, Journal of Climate, 26, 7507–7524, https://doi.org/https://doi.org/10.1175/JCLI-D-12-00736.1, 2013.

Myers, T. A. and Norris, J. R.: On the Relationships between Subtropical Clouds and Meteorology in Observations and CMIP3 and CMIP5 Models, Journal of Climate, 28, 2945–2967, https://doi.org/https://doi.org/10.1175/JCLI-D-14-00475.1, 2015.

Myers, T. A. and Norris, J. R.: Reducing the uncertainty in subtropical cloud feedback, Geophysical Research Letters, 43, 2144–2148, https://doi.org/10.1002/2015GL067416, 2016.

Myers, T. A., Scott, R. C., Zelinka, M. D., Klein, S. A., Norris, J. R., and Caldwell, P. M.: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity, Nature Climate Change, 11, 501–507, https://doi.org/10.1038/s41558-021-01039-0, 2021.

Padarian, J., McBratney, A. B., and Minasny, B.: Game theory interpretation of digital soil mapping convolutional neural networks, Soil, 6, 389–397, https://doi.org/10.5194/soil-6-389-2020, 2020.

Platnick, S. and Twomey, S.: Determining the Susceptibility of Cloud Albedo to Changes in Droplet Concentration with the Advanced Very High Resolution Radiometer, Journal of Applied Meteorology and Climatology, 33, 334–347, https://doi.org/https://doi.org/10.1175/1520-0450(1994)033<0334:DTSOCA>2.0.CO;2, 1994.

Qu, X., Hall, A., Klein, S. A., and Caldwell, P. M.: The strength of the tropical inversion and its response to climate change in 18 CMIP5 models, Climate Dynamics, 45, 375–396, https://doi.org/10.1007/s00382-014-2441-9, 2015a.

Qu, X., Hall, A., Klein, S. A., and Deangelis, A. M.: Positive tropical marine low-cloud cover feedback inferred from cloud-controlling factors, Geophysical Research Letters, 42, 7767–7775, https://doi.org/10.1002/2015GL065627, 2015b.

Quaas, J., Boucher, O., and Lohmann, U.: Constraining the total aerosol indirect effect in the LMDZ and ECHAM4 GCMs using MODIS satellite data, Atmospheric Chemistry and Physics, 6, 947–955, https://doi.org/10.5194/acp-6-947-2006, 2006.

Rieck, M., Nuijens, L., and Stevens, B.: Marine Boundary Layer Cloud Feedbacks in a Constant Relative Humidity Atmosphere, Journal of the Atmospheric Sciences, 69, 2538–2550, https://doi.org/https://doi.org/10.1175/JAS-D-11-0203.1, 2012.

Rosenfeld, D., Zhu, Y., Wang, M., Zheng, Y., Goren, T., and Yu, S.: Aerosol-driven droplet concentrations dominate coverage and water of oceanic low-level clouds, Science, 363, https://doi.org/10.1126/science.aav0566, 2019.

Sato, Y., Goto, D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H., and Nakajima, T.: Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model, Nature Communications, 9, 1–7, https://doi.org/10.1038/s41467-018-03379-6, 2018.

Schwarz, K., Cermak, J., Fuchs, J., and Andersen, H.: Mapping the Twilight Zone—What We Are Missing between Clouds and Aerosols, Remote Sensing, 9, https://doi.org/10.3390/rs9060577, 2017.

Scott, R. C., Myers, T. A., Norris, J. R., Zelinka, M. D., Klein, S. A., Sun, M., and Doelling, D. R.: Observed Sensitivity of Low-Cloud Radiative Effects to Meteorological Perturbations over the Global Oceans, Journal of Climate, 33, 7717–7734, https://doi.org/https://doi.org/10.1175/JCLI-D-19-1028.1, 2020.

Seifert, A., Heus, T., Pincus, R., and Stevens, B.: Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection, Journal of Advances in Modeling Earth Systems, 7, 1918–1937, https://doi.org/10.1002/2015MS000489, 2015.

Seinfeld, J. H., Bretherton, C., Carslaw, K. S., Coe, H., DeMott, P. J., Dunlea, E. J., Feingold, G., Ghan, S., Guenther, A. B., Kahn, R., Kraucunas, I., Kreidenweis, S. M., Molina, M. J., Nenes, A., Penner, J. E., Prather, K. A., Ramanathan, V., Ramaswamy, V., Rasch, P. J., Ravishankara, A. R., Rosenfeld, D., Stephens, G., and Wood, R.: Improving our fundamental understanding of the role of aerosol-cloud interactions in the climate system, Proceedings of the National Academy of Sciences of the United States of America, 113, 5781–5790, https://doi.org/10.1073/pnas.1514043113, 2016.

Small, J. D., Chuang, P. Y., Feingold, G., and Jiang, H.: Can aerosol decrease cloud lifetime?, Geophysical Research Letters, 36, 1–5, https://doi.org/10.1029/2009GL038888, 2009.

Small, J. D., Jiang, J. H., Su, H., and Zhai, C.: Relationship between aerosol and cloud fraction over Australia, Geophysical Research Letters, 38, https://doi.org/https://doi.org/10.1029/2011GL049404, 2011.

Snoek, J., Larochelle, H., and Adams, R. P.: Practical Bayesian Optimization of Machine Learning Algorithms, in: Advances in Neural Information Processing Systems, edited by Pereira, F., Burges, C. J., Bottou, L., and Weinberger, K. Q., vol. 25, Curran Associates, Inc., https://proceedings.neurips.cc/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf, 2012.

Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J. E., and Favez, O.: Meteorology-driven variability of air pollution (PM1) revealed with explainable machine learning, Atmospheric Chemistry and Physics, 21, 3919–3948, https://doi.org/10.5194/acp-21-3919-2021, 2021.

Su, W., Loeb, N. G., Xu, K.-M., Schuster, G. L., and Eitzen, Z. A.: An estimate of aerosol indirect effect from satellite measurements with concurrent meteorological analysis, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/https://doi.org/10.1029/2010JD013948, 2010.

Toll, V., Christensen, M., Quaas, J., and Bellouin, N.: Weak average liquid-cloud-water response to anthropogenic aerosols, Nature, 572, 51–55, https://doi.org/10.1038/s41586-019-1423-9, 2019.

Turner, D. D.: Improved ground-based liquid water path retrievals using a combined infrared and microwave approach, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2007JD008530, 2007.

Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, Journal of the Atmospheric Sciences, 34, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034<1149:tiopot>2.0.co;2, 1977.

van der Dussen, J. J., de Roode, S. R., Dal Gesso, S., and Siebesma, A. P.: An LES model study of the influence of the free tropospheric thermodynamic conditions on the stratocumulus response to a climate perturbation, Journal of Advances in Modeling Earth Systems, 7, 670–691, https://doi.org/https://doi.org/10.1002/2014MS000380, 2015.

Wall, C. J., Hartmann, D. L., and Ma, P.-L.: Instantaneous Linkages between Clouds and Large-Scale Meteorology over the Southern Ocean in Observations and a Climate Model, Journal of Climate, 30, 9455–9474, https://doi.org/https://doi.org/10.1175/JCLI-D-17-0156.1, 2017.

Wang, S., Wang, Q., and Feingold, G.: Turbulence, Condensation, and Liquid Water Transport in Numerically Simulated Non-precipitating Stratocumulus Clouds, Journal of the Atmospheric Sciences, 60, 262–278, https://doi.org/https://doi.org/10.1175/1520-0469(2003)060<0262:TCALWT>2.0.CO;2, 2003.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373–2423, https://doi.org/10.1175/MWR-D-11-00121.1, 2012.

Wood, R. and Bretherton, C. S.: On the relationship between stratiform low cloud cover and lower-tropospheric stability, Journal of Climate, 19, 6425–6432, https://doi.org/10.1175/JCLI3988.1, 2006.

Wood, R. and Hartmann, D. L.: Spatial Variability of Liquid Water Path in Marine Low Cloud: The Importance of Mesoscale Cellular Convection, Journal of Climate, 19, 1748–1764, https://doi.org/https://doi.org/10.1175/JCLI3702.1, 2006.

Wood, R., Wyant, M., Bretherton, C. S., Rémillard, J., Kollias, P., Fletcher, J., Stemmler, J., de Szoeke, S., Yuter, S., Miller, M., Mechem, D., Tselioudis, G., Chiu, J. C., Mann, J. A. L., O'Connor, E. J., Hogan, R. J., Dong, X., Miller, M., Ghate, V., Jefferson, A., Min, Q., Minnis, P., Palikonda, R., Albrecht, B., Luke, E., Hannay, C., and Lin, Y.: Clouds, Aerosols, and Precipitation in the Marine Boundary Layer: An Arm Mobile Facility Deployment, Bulletin of the American Meteorological Society, 96, 419–440, https://doi.org/10.1175/BAMS-D-13-00180.1, 2015.

Xue, H. and Feingold, G.: Large-Eddy Simulations of Trade Wind Cumuli: Investigation of Aerosol Indirect Effects, Journal of the Atmospheric Sciences, 63, 1605–1622, https://doi.org/10.1175/JAS3706.1, 2006.

Young, G. S., Kristovich, D. A. R., Hjelmfelt, M. R., and Foster, R. C.: ROLLS, STREETS, WAVES, AND MORE: A Review of Quasi-Two-Dimensional Structures in the Atmospheric Boundary Layer, Bulletin of the American Meteorological Society, 83, 997–1002, https://doi.org/https://doi.org/10.1175/1520-0477(2002)083<0997:RSWAMA>2.3.CO;2, 2002.

Yuan, T. and Oreopoulos, L.: On the global character of overlap between low and high clouds, Geophysical Research Letters, 40, 5320–5326, https://doi.org/https://doi.org/10.1002/grl.50871, 2013.

Yuan, T., Remer, L. A., and Yu, H.: Microphysical, macrophysical and radiative signatures of volcanic aerosols in trade wind cumulus observed by the A-Train, Atmospheric Chemistry and Physics, 11, 7119–7132, https://doi.org/10.5194/acp-11-7119-2011, 2011.

Zelinka, M. D., Andrews, T., Forster, P. M., and Taylor, K. E.: Quantifying components of aerosol-cloud-radiation interactions in climate models, Journal of Geophysical Research: Atmospheres, 119, 7599–7615, https://doi.org/10.1002/2014JD021710, 2014.

Zelinka, M. D., Grise, K. M., Klein, S. A., Zhou, C., DeAngelis, A. M., and Christensen, M. W.: Drivers of the Low-Cloud Response to Poleward Jet Shifts in the North Pacific in Observations and Models, Journal of Climate, 31, 7925–7947, https://doi.org/https://doi.org/10.1175/JCLI-D-18-0114.1, 2018.

Zhang, Z. and Platnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2011JD016216, 2011.

Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R., and Xue, H.: Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/https://doi.org/10.1029/2012JD017655, 2012.

Zheng, G., Wang, Y., Wood, R., Jensen, M. P., Kuang, C., McCoy, I. L., Matthews, A., Mei, F., Tomlinson, J. M., Shilling, J. E., Zawadowicz, M. A., Crosbie, E., Moore, R., Ziemba, L., Andreae, M. O., and Wang, J.: New particle formation in the remote marine boundary layer, Nature Communications, 12, https://doi.org/10.1038/s41467-020-20773-1, 2021.

Zhu, Y., Rosenfeld, D., and Li, Z.: Under What Conditions Can We Trust Retrieved Cloud Drop Concentrations in Broken Marine Stratocumulus?, Journal of Geophysical Research: Atmospheres, 123, 8754–8767, https://doi.org/10.1029/2017JD028083, 2018.

Zipfel, L., Andersen, H., and Cermak, J.: Machine-Learning Based Analysis of Liquid Water Path Adjustments to Aerosol Perturbations in Marine Boundary Layer Clouds Using Satellite Observations, Atmosphere, 13, https://doi.org/10.3390/atmos13040586, 2022.