Karlsruhe Institute of Technology

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 821205

FORCeS

Constraints on cloud fraction adjustment to aerosols using explainable machine learning

A novel approach to address the interactions between aerosol-cloud interactions and meteorology

Authors: Yichen Jia (contact: Yichen Jia — yichen.jia@kit.edu), Hendrik Andersen, Jan Cermak

1. Motivation and Aim

- Separating the influence of aerosols from meteorological parameters in a complex buffered aerosol-cloud-climate system remains a challenge.
- Using cloud droplet number concentration (N_d) as a proxy for aerosol, the employed explainable Machine Learning (ML) framework aims to **quantify the marine boundary layer cloud fraction (CLF) sensitivity to aerosols and meteorological factors, and to estimate the dependence of N_d–CLF relationship on meteorology by investigating interaction effects in the trained ML models**.

2. Data and Methods

Satellite and reanalysis datasets

- Nine years (2011-2019) daily data from Terra platform (N_d, CLF) and ERA5 reanalysis (21 meteorological variables), spatiotemporally harmonized.
- Each original 1°× 1°grid is aggregated to 5°× 5°"windows" and one ML model is trained and tested for each specific window.

Machine learning and SHapley Additive exPlanation (SHAP) values

- Extreme Gradient Boosting (XGB) models are used to predict CLF.
- SHAP regression values are used to quantify the contribution of each predictor to each individual model prediction.
- Figure 1 a: CLF sensitivity is defined as the slope of the linear regression between the SHAP values (isolated contributions to the model predicted CLF) and feature values of a specific predictor.
- Figure 1 b: Interaction Index (IAI) is defined as the difference between the slopes of linear regressions of the SHAP interaction values and the features for high SST values (above-average) and low SST values (below-average).

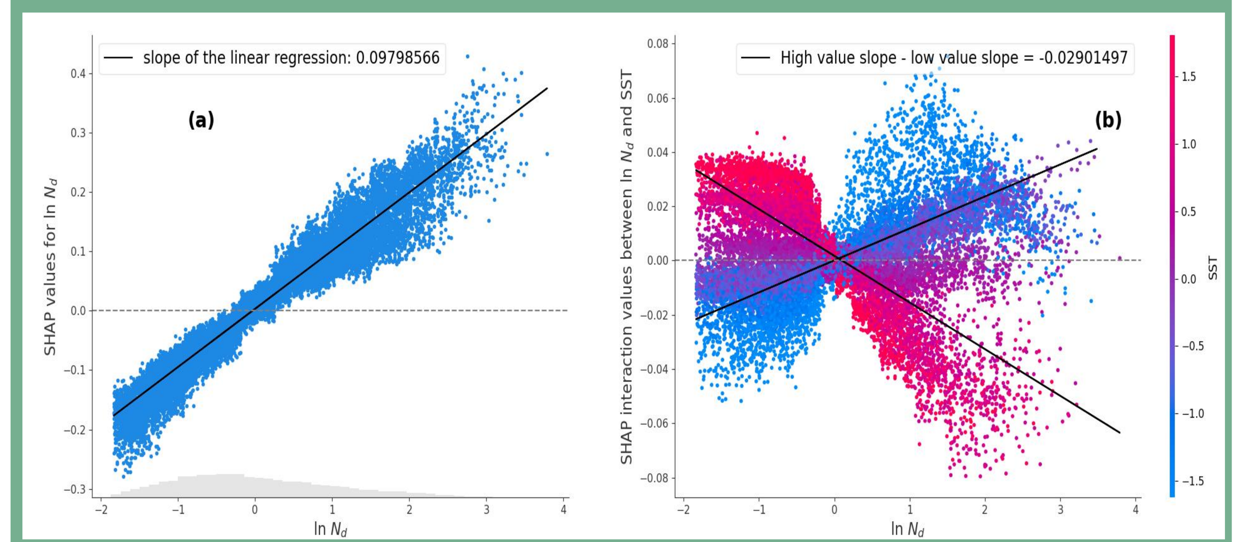

Figure 1: An example of CLF sensitivity to SST (a) and interaction effects between SST and N_d colored by N_d (b)

- **Negative IAI**: sensitivity stronger with low (< mean) feature values
- **Positive IAI**: sensitivity stronger with high (> mean) feature values

3. Results and Discussions

XGB model performance

XGB model skills range from around 0.1 to 0.75, higher in the regions dominated by stratiform cloud systems

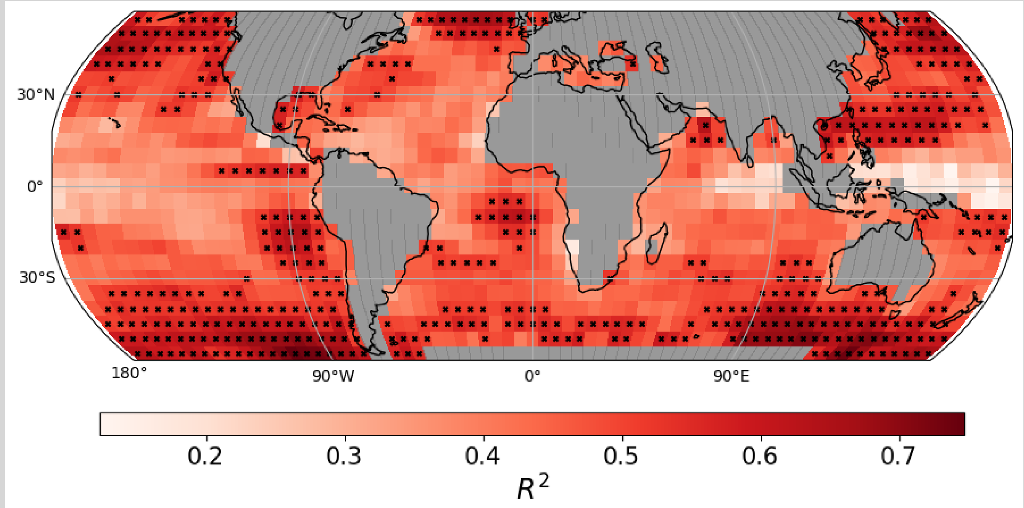

Figure 2: Performance of regional XGB models predicting CLF.

CLF sensitivity: global perspective and regional characteristics

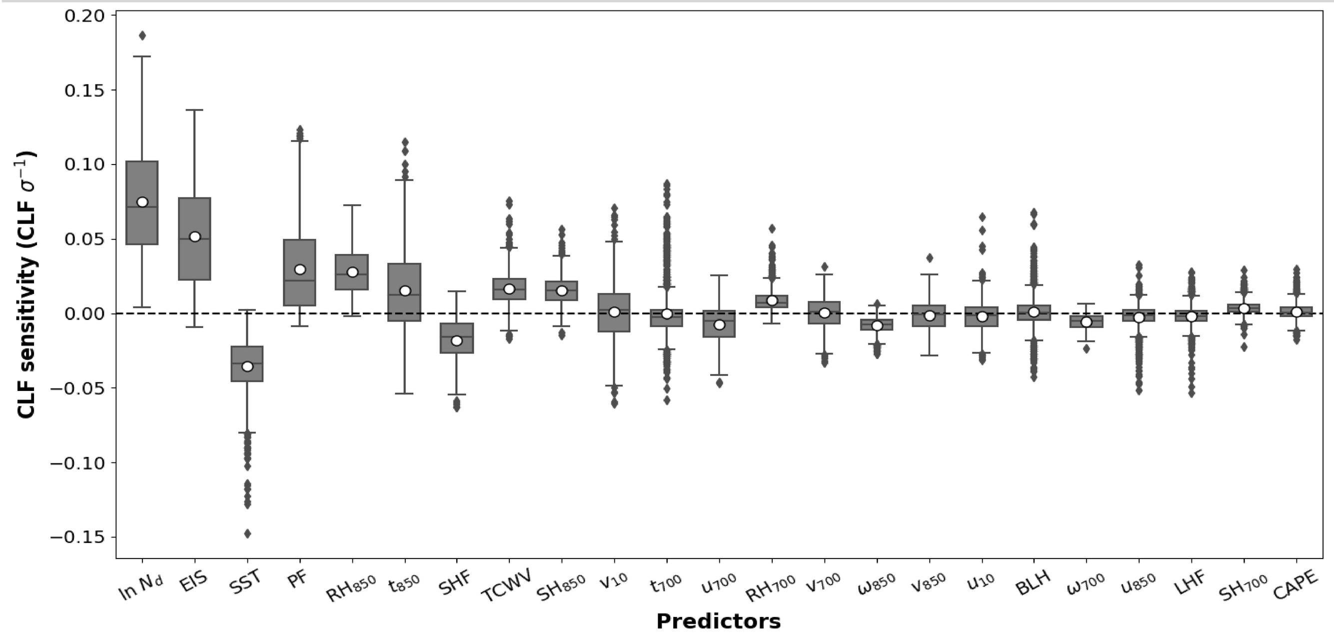

Figure 3: The distribution of the sensitivities of CLF to all predictors. The sequence is sorted descendingly by the mean values of the absolute sensitivities.

- CLF is the most sensitive to N_d, estimated inversion strength (EIS), precipitation fraction (PF), relative humidity at 850 hPa (RH_{850}) (positive) and sea surface temperature (SST), surface sensible heat flux (SHF) (negative).
- Strong CLF sensitivity to N_d not directly comparable to meteorological sensitivities (N_d observations more closely linked to CLF observations than reanalysis).

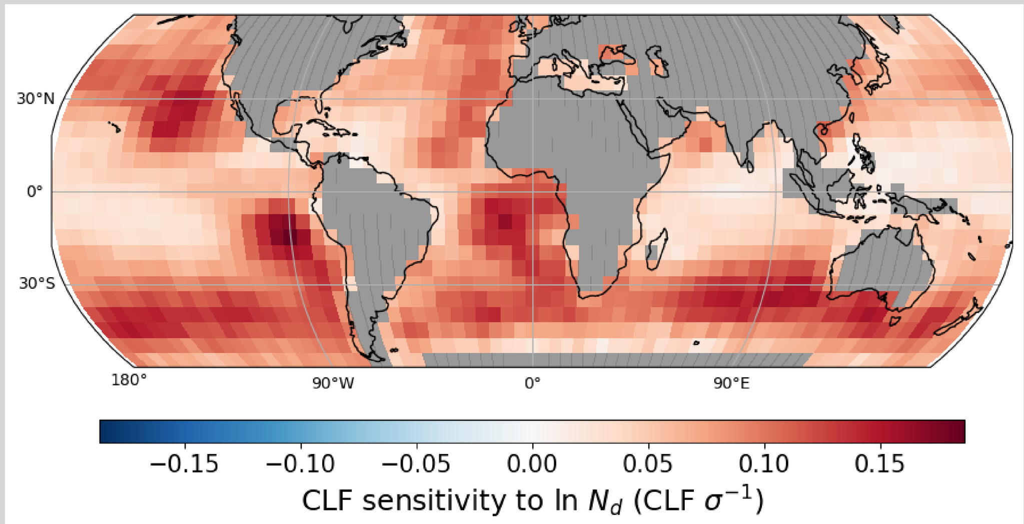

Figure 4: Sensitivity of marine boundary layer cloud fraction to ln N_d

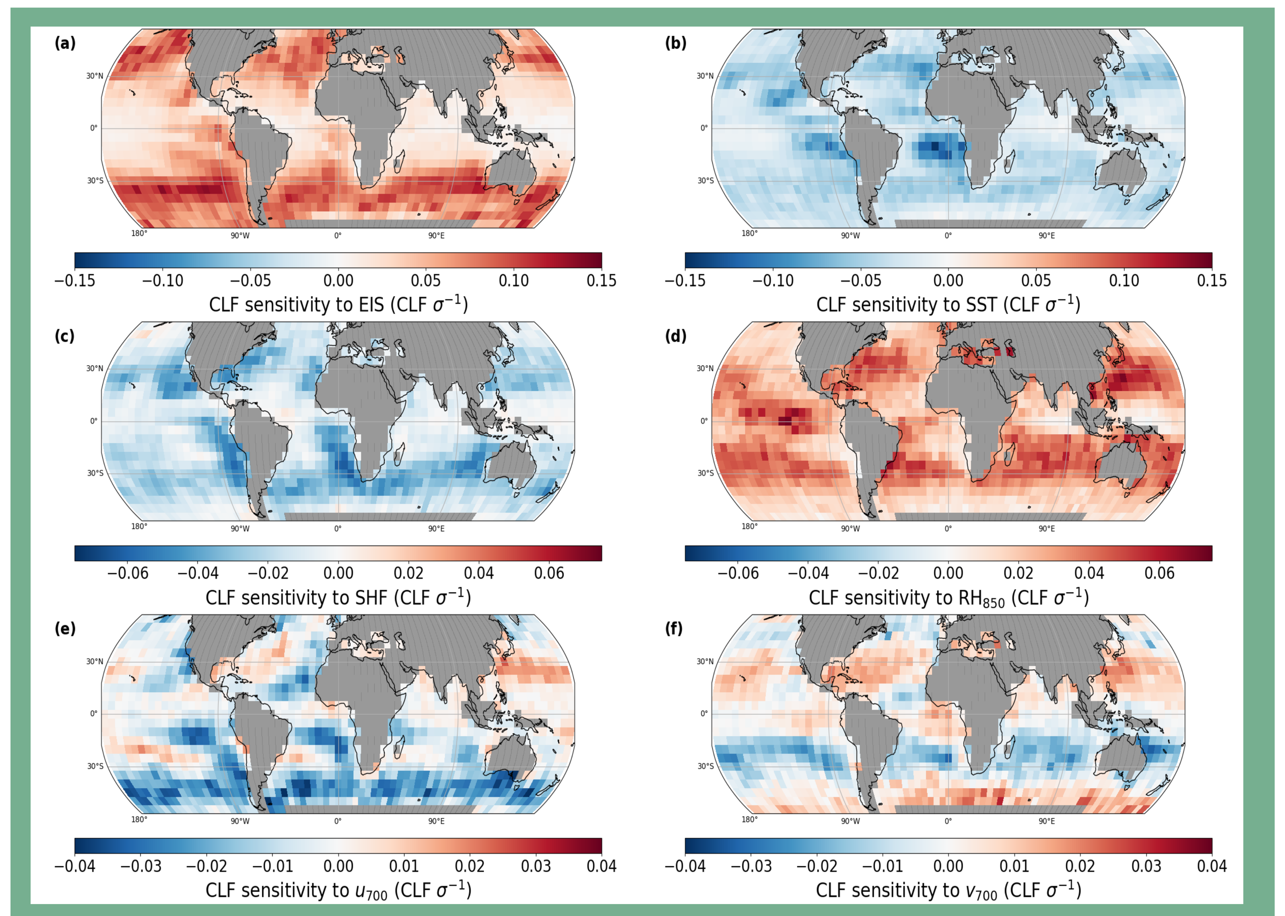

Figure 5: Geographical patterns of the sensitivity of marine boundary layer cloud fraction to EIS, SST, SHF and N_d, u_{700}, v_{700}. Note that the range of colourbars vary.

- CLF is positively associated with EIS and RH_{850} ubiquitously, strongest in the stratus and stratocumulus regions for EIS; and in the trade cumulus and tropical ascent cumulus regions for RH_{850}.
- CLF is negatively asscoiated with SHF and SST globally, strongest in the stratocumulus regions.
- Distinct spatial patterns for horizontal wind components at 700 hPa indicate the importance of midlatitude synoptic forcing.

Dependence of N_d-CLF relationship on ambient meteorology

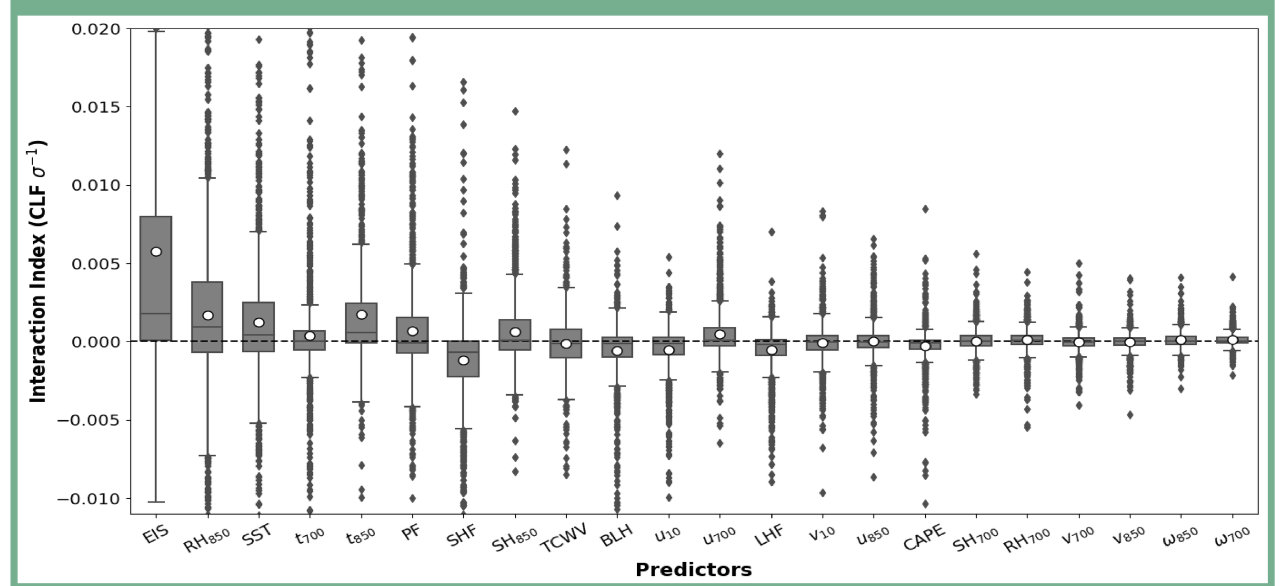

Figure 6: SImilar to Fig. 3 but for the Interaction Indices.

- EIS, RH_{850}, SST, temperatures at 700 and 850 hPa have the most significant influences on the N_d-CLF relationship.
- In general, thermodynamical factors have more interactions with the N_d-CLF sensitivity than dynamical factors.

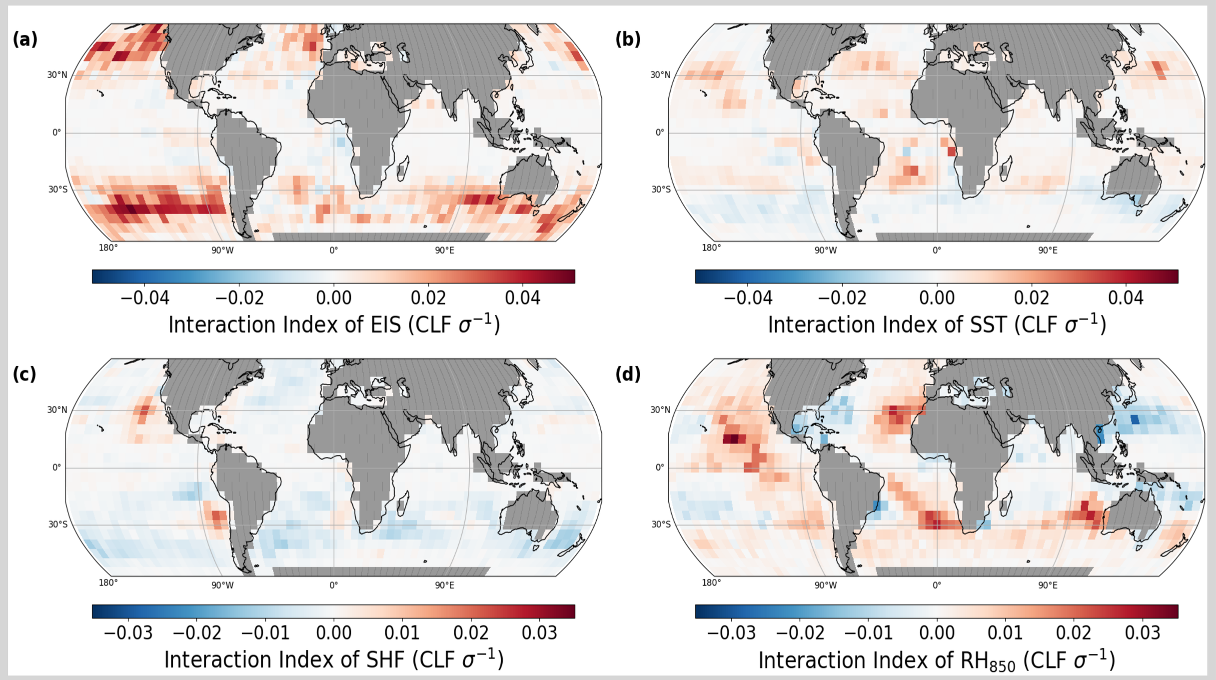

Figure 7: Global patterns of IAIs for the interactive effect between N_d and EIS (a), SST (b), SHF (c), RH_{850} (d). Note that the range of colourbars vary.

- EIS exerts positive IAIs over the midlatitude oceans, reflecting that stronger inversions capping these areas will amplify the positive N_d-CLF relationship.
- In trade cumulus regions, higher SSTs are found to amplify the positive N_d-CLF relationship.
- A positive IAI is apparent in the stratocumulus-topped eastern Pacific oceanic basins, which indicates that increasing upward SHF is associated with a weaker CLF enhancement from N_d.
- The N_d-CLF relationship in stratocumulus and trade cumulus regions tends to be stronger in humid boundary layers.

4. Conclusions and Outlook

- Marine boudary layer cloud fraction is the most sensitive to N_d (surrogate of aerosol) in regions dominated by stratiform clouds.
- EIS and SST are important determinants, dynamic drivers indicate that mid-latitude synoptic-scale disburtances make considerable contributions.
- In general, thermodynamical parameters exert more important influence on the N_d-CLF relationship than dynamical parameters.
- Both CLF sensitivities and the interactions with meteorology exhibit distinct regional patterns.

Future work:

The observation-based sensitivity and interactions investigated by the ML framework in this study will be compared to ESM-based ones with identical ML setup, offering a novel way to evaluate ESM parameterizations related to ACI help gain insights into how they could be tuned.

Abstract

EGU Outstanding Student & PhD candidate Presentation contest

Sharing is encouraged