Response to interactive comment of anonymous referee 2 —

Julia Fuchs[1, 2], Jan Cermak[1, 2], and Hendrik Andersen[1, 2]

[1]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

[2]Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

contact: julia.fuchs@kit.edu

This manuscript disentangles aerosol effects on the southeast Atlantic stratocumuls deck from meteorological effects through the use of a machine learning approach labeled Gradient Boosting Regression Trees (GBRTs). It is welcome to see a recognition of both impacts, and the use of an innovate approach to discriminate them. The use of lat_src and lon_src is nice. The results are sensible. I do however feel the study suffers from over-interpretation. One concern is the focus on only the cloud fraction and the cloud effective radius (REF) as the cloud properties. While the REF is influenced by aerosol, it is also a function of the liquid water path. A more straightforward physical relationship is that between AOD (CCN) and the cloud droplet number concentration (Nd), which can be estimated as a function of REF and the cloud optical depth. Cloud deepening is likewise better interpreted through the use of LWP than of REF. Another concern is the lumping of July-August-September. It is by now well appreciated that the biomass-burning aerosol is more likely to be present within the boundary layer in July, moving up in altitude through September, when it is more likely to be above the cloudy boundary layer. Different cloud responses would be anticipated as a function of the month. A useful additional analysis is to examine the GBRT results as a function of month, and interpret them as a function of the varying cloud-aerosol vertical structure.

The study was designed to focus on cloud fraction and cloud effective radius in order to test and interpret the GBRT models on one relevant micro- and one relevant macrophysical cloud property during the biomass-burning season. Using the cloud droplet number concentration is appreciated, however, as it is derived from COT and REF, and based on assumptions on cloud vertical profile, additional uncertainties would be introduced (Grosvenor et al., 2018). We chose to avoid this, because we try to capture the cloud system as completely as possible with the statistical model. As such, we include information on factors that also determine LWP. As variability among these LWP-predictors is simulated in the computation of the sensitivities, we thereby indirectly constrain LWP effects on REF. To account for the referee's suggestion LWP as an essential cloud property is analyzed and results support the interpretation of cloud thickening under stable conditions in all subregions (Fig. 1a), as well as during westerly disturbances, especially in the western subregions (Fig. 1b). In the manuscript the LWP-effects refer to outcomes of a comparable study by Fuchs et al. (2017) where LWP is discussed and 'self-constraining model' is now detailed more clearly.

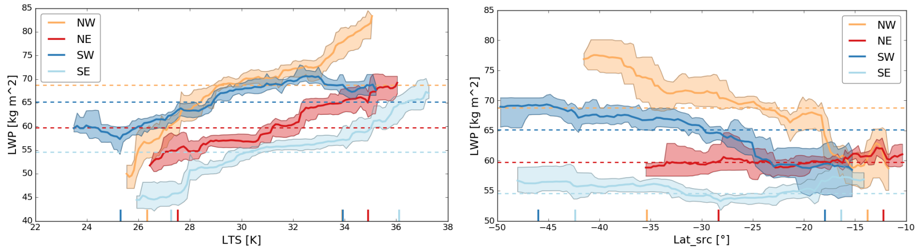
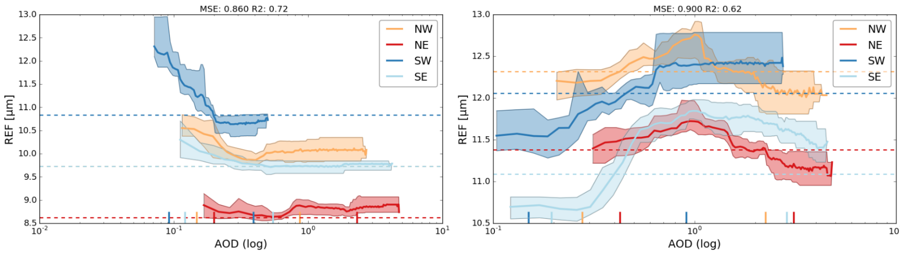


Figure 1: Mean partial dependence of LWP on LTS (left) and source latitude of air mass (right) in the four subregions (colors).

The aggregation of the months July-August-September (JAS) was conducted for better comparability to previous studies (Painemal et al., 2014; Andersen and Cermak, 2015; Adebiyi and Zuidema, 2018) investigating the same season. While we agree with referee 2, changes of the aerosol and boundary layer occur on all scales, so that the assumptions outlined by referee 2 need to be made independent of scale. The intraseasonal variability contained in the training of the GBRT model contribute to the relationships during the investigated season and must be taken into account for the interpretation of results. For this reason, we have now computed monthly GBRT models and included the results concerning the aerosol-cloud relationships in the manuscript. The following figure and text are added to the manuscript. ”Figure 7 shows AOD-REF



(a) (b)

Figure 2: Mean partial dependence of REF on AOD in the four subregions (colors) in July (a) and September (b).

partial dependencies for the months of July and September separately. While during July, REF seems to decrease with increasing AOD, especially in the SW subregion, during September the opposite relationship is found. The contrasting relationships may be related to differences in the vertical distribution of aerosols and clouds in the Southeast Atlantic. During July, aerosol and cloud layers are frequently entangled, facilitating ACI, whereas in September they can be well separated (Adebiyi et al. 2018). During this time, absorbing aerosol may increase the stability and trap humidity in the boundary layer, potentially leading to the observed relationship. The JAS partial dependence between AOD and REF can thus be viewed as a summary of these patterns. However, it is not the study's focus to separate the different aerosol effects mentioned earlier, but to analyze the overall influence of aerosols on clouds during the biomass-burning season.”(p.14, l.10)

Other comments follow:

1. I am not completely comfortable with the use of the 8-day MODIS L3 product used as opposed to shorter time scale, as the 8-day time scale will average over the synoptic time scale and is far longer than the cloud adjustment

time scale of 1-2 days. The authors mention that an 8-day time scale "allows for the large-scale and thermodynamic forcings of cloud properties to be combined", but I remain unclear what this means exactly. In several places in the manuscript the authors refer to processes that occur at much smaller time scales, such as the cloud microphysical response to aerosol. Instead it seems to me the 8-day time scale is primarily capturing a portion of the monthly evolution in the aerosol-cloud vertical structure and seasonal meteorological cycle. Also, the 8-day time scale should be explicitly mentioned in the abstract. The study focuses on processes on aggregated time scales (8-day), assuming that cloud adjustments due to aerosols, though acting on smaller time scales, are detectable in the aggregated data set at the same time as changes of thermodynamic and dynamic conditions. While daily or hourly data might underestimate e.g. the effect of LTS on the cloud cover, the influence of aerosols on the cloud cover might be underestimated by the 8-day aggregation. In particular, since aerosol and cloud properties are not retrieved at the same time in a given location. Eight-day averages are taken to represent the mean states of both at that time scale. These aspects are important and now more explicitly addressed in the manuscript. "The temporal resolution of 8 days allows to combine large-scale, thermodynamic and aerosol forcings of cloud properties simultaneously on a synoptical scale. However, it must be taken into account that clouds adjust on different time scales (hours to several days) to their environment (Klein, 1997; McCoy et al., 2017; Adebiyi and Zuidema, 2018) and thus processes relevant on shorter time scales might be underrepresented in the data set."(p.3, l.8) The 8-day time scale is now introduced in the abstract (p.1, l.5).

2. an issue with using the relative humidity at 950 hPa is that changes in RH are more likely to reflect co-variations with other factors such as the

cold-temperature advection (I suspect this explains the stronger relationship between RH_950hpa and REF in the SE sub-region) and cloud-top inversion strength. Have the authors examined the cross-correlations between their predictors?

Thank you for pointing out this issue. Correlations between the predictors were examined in advance and influenced the choice of predictors to reduce the covariation. Cold-temperature advection and cloud-top inversion strength are not explicitly chosen as predictors, but are assumed to be represented in the data set by other parameters such as wind speed, sea surface temperature and LTS. The manuscript is modified on p.6, l.26: "The predictor set was selected in a way to reduce covariation."

3. how is it that the machine learning approach is able to grasp non-linear relationships? The description of the technique presented on p. 4 still seems to present it as a basically linear technique.

The GBRT algorithm is based on decision trees which are capable of representing non-linear dependencies between predictor and predictand. The parameter space is iteratively split with the goal to minimize a loss function. The sum of the linear decisions of each tree in the ensemble can then represent non-linear relationships. The manuscript is modified on p.3, l.24.

4. It is worth mentioning that the larger region encompassing the 4 subregions has been previously examined in Klein and Hartmann 1993.

This reference is added on p. 3, l.13: "In this study CF and REF are simulated based on a selected predictor set (AOD and meteorological parameters) in the SEA (10°–20° S, 0°–10° E, as analyzed in Klein and Hartmann (1993)) using Gradient Boosting Regression Trees (GBRTs)."

References

Adebiyi, A. A. and Zuidema, P. (2018). Low Cloud Cover Sensitivity to Biomass-Burning Aerosols and Meteorology over the Southeast Atlantic. *Journal of Climate*, 31(11):4329–4346.

Andersen, H. and Cermak, J. (2015). How thermodynamic environments control stratocumulus microphysics and interactions with aerosols. *Environmental Research Letters*, 10(2):024004.

Fuchs, J., Cermak, J., Andersen, H., Hollmann, R., and Schwarz, K. (2017). On the Influence of Air Mass Origin on Low-Cloud Properties in the Southeast Atlantic. *Journal of Geophysical Research: Atmospheres*, 122(20):11,076–11,091.

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. M., Diamond, M. S., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C. L., 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. (2018). Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Reviews of Geophysics*.

Klein, S. and Hartmann, D. (1993). The seasonal cycle of low stratiform clouds. *Journal of Climate*, 6:1587–1606.

Klein, S. A. (1997). Synoptic Variability of Low-Cloud Properties and Meteorological Parameters in the Subtropical Trade Wind Boundary Layer. *Journal of Climate*, 10(8):2018–2039.

McCoy, D. T., Eastman, R., Hartmann, D. L., and Wood, R. (2017). The change in low cloud cover in a warmed climate inferred from AIRS, MODIS, and ERA-interim. *Journal of Climate*, 30(10):3609–3620.

Painemal, D., Kato, S., and Minnis, P. (2014). Boundary layer regulation in the southeast Atlantic cloud microphysics during the biomass burning season as seen by the A-train satellite constellation. *Journal of Geophysical Research: Atmospheres*, 119(19):11,288–11,302.