# Understanding the drivers of marine liquid-water cloud occurrence and properties with global observations using neural networks

# — RESPONSE TO REFEREE 1 —

contact: hendrik.andersen@kit.edu

We would like to thank referee 1 for her/his review of the manuscript and her/his constructive criticism. Comments by the referee are colored in blue, our replies are colored in black.

We have thoroughly considered and discussed your input and after careful analysis of each review point concur with you that we have indeed not sufficiently 'articulate[d] what the new thing is that [we] bring to the table', as you state as your 'overarching concern'. The work in this manuscript has a history of several years, over which we have discussed ideas and results with peers and internally many times, so that in writing the manuscript we may have taken several points for granted that are in fact new to a reader confronted with the study for the first time. In this spirit, we have now attempted, guided by your suggestions, to more carefully explain the whats, hows and whys of our research, as well as what is new, and what is not.

I appreciate that the authors have attempted to diversify the ACI investigation field with the use of neural networks. It is often difficult with studies such as this that attempt a new analysis method to create a coherent message. However, I do not think this paper can be published in its current form. My overarching concern in this paper is that the authors do not articulate what the new thing is that they bring to the table besides the black box of a neural network.

General response: See above. Figure 1 included in this document is intended to illustrate the concept of our study schematically: Frequently, aerosol-cloud interactions are studied in a rather isolated manner (in red). At the same time, it is commonly acknowledged that the influence of aerosols is modulated by many environmental factors. With this study, we aim at analyzing the aerosol-cloud-climate system in its entirety. This includes all variations in the environmental conditions, including the seasonal cycle (and its variability) of clouds and meteorology. Our first aim therefore is to find a way to statistically capture this system as completely as possible, including seasonality. Then, in a second step, we focus on and try to separate the effects of aerosols on cloud occurrence and properties from everything else. Our work is not intended to refute previous work done in this field. On the contrary: We would argue that most of the results presented within the study confirm many known aspects of the aerosol-cloud-climate system. But the fact that we were able to find these relationships in a statistical approach considering much more than only aerosol and cloud properties adds an additional line of independent evidence that strengthens the confidence in the existing system understanding. However, this is achieved without isolating specific processes of interest but rather by viewing the system in its entirety. Accordingly, these are the main new things we 'bring to the table': Confidence that the observation data sets considered in a multivariate statistical approach capture the natural variability, and that aerosol effects similar to those found in other studies can be identified in this system. No more, no less.

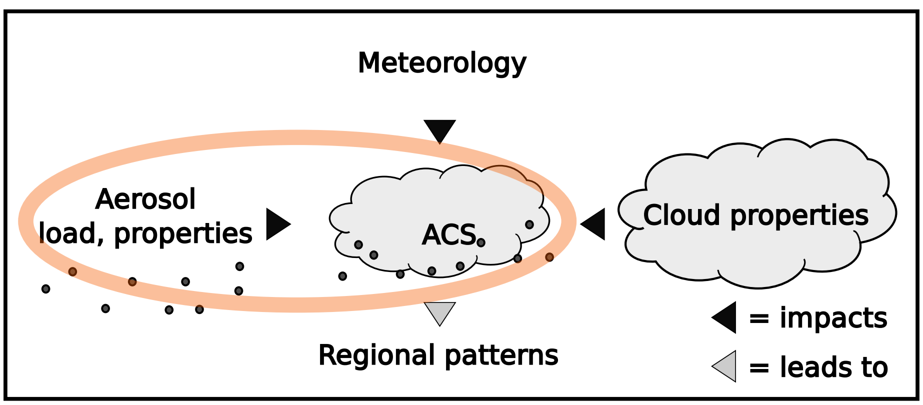


Figure 1: A schematic illustration of the concept of this study (ACS: aerosol-cloud sensitivity).

I have grouped my concerns about this paper into the following categories:

# Statistical evaluation

1) It is unclear to me why doing multiple neural networks on sub regions on monthly data tells us anything useful about what is going on. I need some sort of confidence that a high R2 model cannot be created by a large neural network using a collection of meteorological predictors picked at random. Monthly data has the issue of being driven by the seasonal cycle, which will drive almost everything else, and making it regional will mean that the neural network doesn't need to tell us anything particularly meaningful about how the clouds are driven by their environment. The authors should consider using anomalies relative to the seasonal mean, or simply using annual means. Either of these options would be better than the approach taken in this paper. Admittedly the authors talk about this on page 2 line 25, but they don't provide any convincing proof that they haven't just created a regional seasonal cycle simulator.

This is related to what we argue above: We intend to model liquid-water clouds including their seasonal cycle by using information on aerosol loading and a set of meteorological drivers that were identified as main drivers of liquid-water clouds after careful study of current literature. One could probably create a relatively high $R^2$ model with a very large array of randomly selected predictors due to spurious covariation of seasonal cycles between predictors and predictands. However, in this study, we avoid this by capturing the aerosol-cloud-climate system with a small number of the known main drivers of cloud occurrence and properties. Within this modeled system we then try to understand the effects of each driver and its regional patterns. We argue that regionally specific neural networks are needed to capture the regional variability of liquid-water clouds. Regional patterns exist due to regional differences in cloud type, aerosol composition, meteorology and the respective seasonal cycles.

2) On page 4/line 10 the authors note that they throw out models that have a low R2. I'm not sure why this is ok to do.

We have identified $R^2$ and the root mean square error relative to the mean as good indicators for model skill. We are interested in understanding predictor-predictand relationships by analyzing their respective sensitivities, however, we choose to trust only models that can adequately represent the observed cloud patterns. We prefer to err on the side of caution to avoid reaching conclusions based on inadequate statistical relationships; thus we exclude models that in our opinion are not capable of representing the system well enough. We are open to other ideas regarding alternative ways to ensure adequate model skill.

3) On page 6 I find this something of a straw man. A better test would be to compare multiple linear regression of all the predictors to the ANN, as opposed to a regression on AOD alone. Or to compare the ANN trained using only AOD. I think that the paper would actually be vastly improved by just repeating the analysis with a multiple linear regression to demonstrate to skeptical readers why their paper brings anything new to the table as compared to the numerous previous papers that have looked at ACI and low cloud variability in the past.

We probably did not communicate the intention of this figure with sufficient clarity: This figure is intended to show how well a combination of aerosol and meteorological conditions can explain the variance of cloud properties (multi-variate statistics) as opposed to a simple bivariate approach. We have added results of a multiple linear regression using all the ANN predictors to the figure (2). The comparison of the results of the multiple linear regression and the ANN

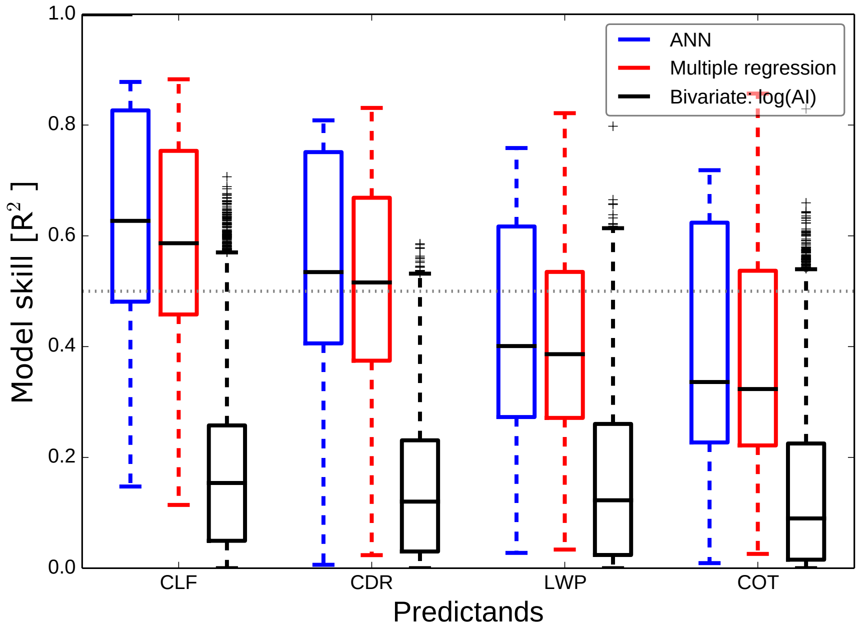


Figure 2: Predictand correlation with ANN (multivariate) test output, multiple linear regression (multivariate) and log(AI) (bivariate). The median is represented by the black horizontal line, framed by the interquartile range (boxes), whiskers expand the boxes by 1.5 interquartile ranges.

suggest that the ANN is an appropriate method to be used in this context.

Neural networks were our statistical method of choice, as they have the advantage of not being reliant on statistical assumptions on predictor and predictand distributions and they are capable of modeling nonlinear relationships. That being said, we agree that other multivariate methods (e.g. multiple linear regression) could also have been used.

4) Figure 5- If the error bars give the range in sensitivity does that mean that nothing except LTS and AOD have a robust relationship with cloud properties that holds outside of a few regions? Didn't we already know this very well from

simple regression models that were easy to interpret (Klein & Hartmann, 1993; Nakajima, Higurashi, Kawamoto, & Penner, 2001)?

We agree with the referee that many of the results of this study confirm what previous studies have already shown. Since we have reached these conclusions using a different methodology, we add another line of evidence. The lack of other relevant relationships would not have been obvious without such an analysis. In our opinion the value of our study is that the results were produced by looking at the entire system at once rather than at isolated relationships. Using this method, we can compare the relevance of each predictor to each predictand including spatial patterns.

5) Choice of predictor/predictands: The choice of predictors by the authors is not appropriate for a paper in the last decade. Why have the authors chosen AOD to be a CCN proxy? AOD is not equivalent to CCN since it has a large contribution from larger, non-CCN relevant aerosols. Why don't the authors use AI, which is far more relevant and typical of more recent studies (Patel, Quaas, & Kumar, 2017)? The authors acknowledge this, but then shrug this off because papers from almost a decade ago do it. In a similar vein, why do the authors use effective radius instead of CDNC? Effective radius for a fixed CCN increases with increasing LWC, making it sensitive to meteorological drivers. The authors do acknowledge this in page 10, section 25 noting that the interaction between inversion strength and effective radius is most likely driven by variations in LWC. This makes the interpretation of the CDR as a proxy for aerosol-cloud effects muddied. Further, the authors use LTS. Why not use EIS, which is used by every study investigating low cloud in the last decade (Myers & Norris, 2015; Qu, Hall, Klein, & Caldwell, 2014; Seethala, Norris, & Myers, 2015; Webb, Lambert, & Gregory, 2013)? Finally, I am concerned with the use

of RH. Clouds and RH are a semi-equivalent quantity, which may just mean that they are comparing ECMWF-interim's cloud cover to MODIS, further aliasing in the seasonal cycle to their prediction model.

AOD vs. AI: For this study, we used the newest version of MODIS products available, collection 6 (C6). In C6, the MODIS Ångström exponent (needed for the computation of the aerosol index as it is the product of AOD and the Ångström exponent) has been discontinued in level 3 (L3) data (p. 3018 Levy et al., 2013). We believe that for this and for other reasons, other recent studies also use the AOD as a proxy for CCN (see: Gryspeerdt and Stier, 2012; Tang et al., 2014; Chakraborty et al., 2016; Stathopoulos et al., 2017; Patel et al., 2017). We agree with referee 1 though, that the aerosol index is an appropriate measure for CCN and have chosen to use it in the ANN. The following figures 3 and 4 are the new results of the ANN when using AI instead of AOD. The spatial patterns in the ANN skill, as well as the mean global sensitivities are nearly identical (compare with figures 3 and 5 in the original ACPD manuscript).

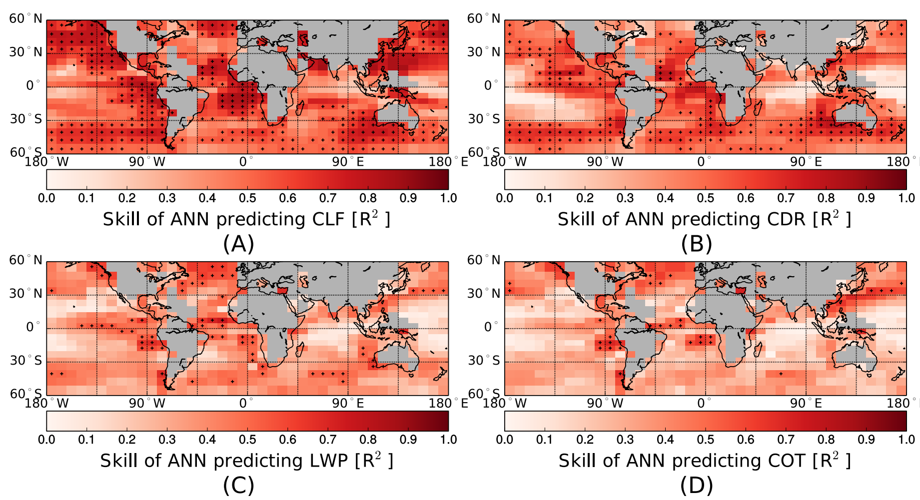


Figure 3: Global patterns of ANN skill as in the manuscript; AI has been used instead of AOD.

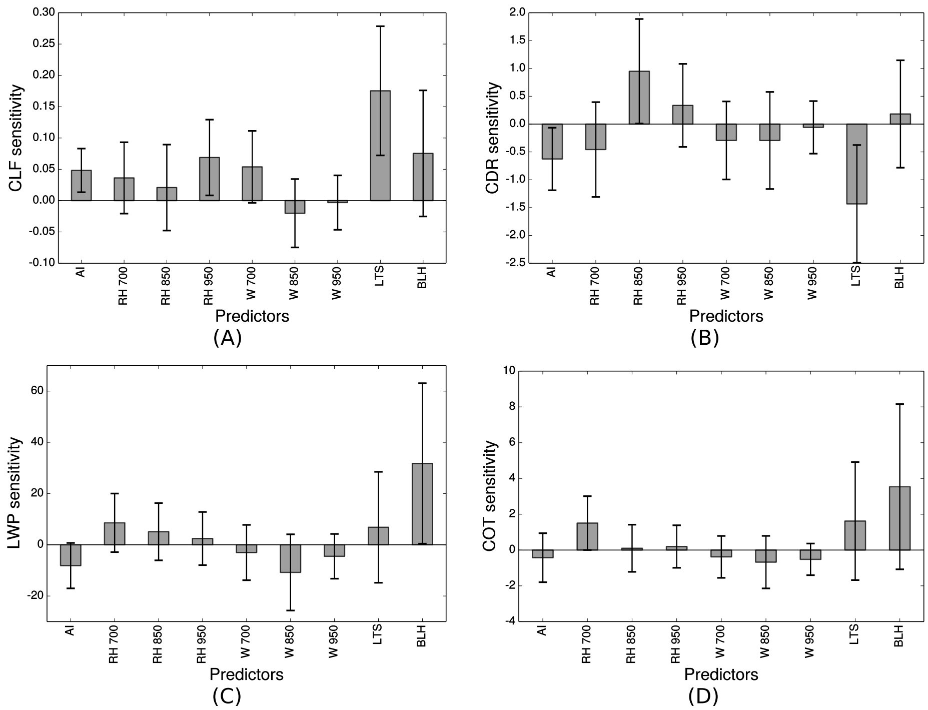


Figure 4: Global mean relative sensitivities as in the manuscript; AI has been used instead of AOD.

Small differences can be observed in the regional patterns of ANN sensitivities (fig. 5 on the following page). The CLF sensitivity to AI is higher in the Southeast Atlantic than its sensitivity to AOD in that specific region. The Southeast Atlantic is of course dominated by biomass burning aerosol, which are mostly in the fine mode and thus feature a relatively larger AI than AOD. The sensitivity of CDR to AI differs from its sensitivity to AOD in regions that are dominated by desert dust. Dust is relatively coarse, so that the AI would be underproportional to the AOD in these regions which might explain the differences between the sensitivities of the two.

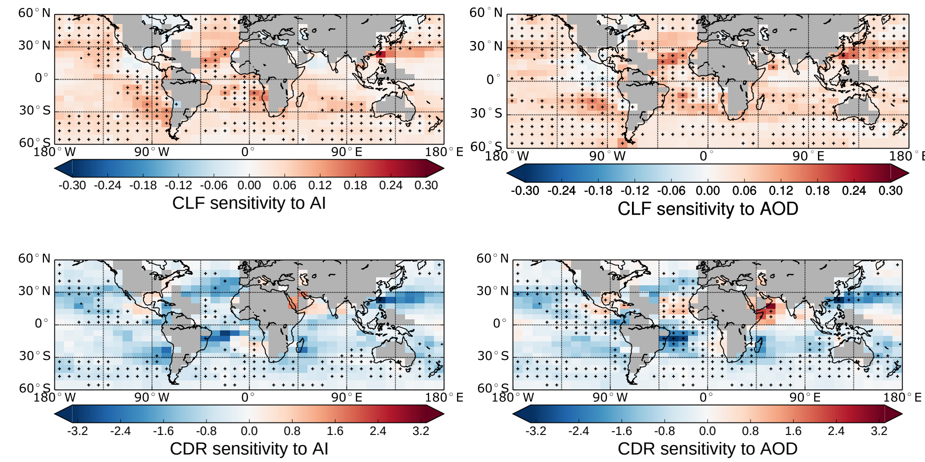


Figure 5: Difference in sensitivities of CLF and CDR to AI (left-hand column) vs. AOD (right-hand column).

CDR vs. CDNC: We agree with the referee that CDNC is a better quantity for the direct analysis of the first aerosol indirect effect, however, its retrieval from satellite is quite problematic, as the retrieval of CDNC requires additional assumptions on the cloud water profile. The commonly-applied adiabatic assumption might be a good proxy for many regions and cloud types (i.e. stratocumulus clouds), however, we are investigating all liquid-water clouds on a global scale. Bennartz and Rausch (2017) showed that the uncertainties in the CDNC retrievals are significantly increased in non-stratocumulus regions. As we are investigating global patterns for various liquid-water cloud types, we came to the conclusion that the uncertainty related to the CDNC retrievals outweighs the theoretical advantages of using CDNC rather than CDR.

LTS vs. EIS: We do not see a specific advantage of using EIS over LTS, as e.g. Lacagnina and Selten (2013) found that for the Californian stratus, LTS is a better predictor than EIS. Some other recent studies that use LTS are e.g. George and Wood (2010); Chen et al. (2014); Gryspeerdt et al. (2014, 2016); Painemal et al. (2014a,b); Adebiyi et al. (2015); Adebiyi and Zuidema (2016); Coopman et al. (2016); Eastman et al. (2016); Ghan et al. (2016). That being said, we would agree that EIS is an appropriate alternative measure for large-scale thermodynamics.

RH: As pointed out above, our intention is to capture the entire aerosol-cloud-climate system and in our opinion, relative humidity has a key role within this system. Thus, the inclusion of RH in the model was a necessity.

# Writing:

The writing is rushed and hard to follow. Clearly expressing why the methodology is valid is crucial for this study and as such the writing needs to be tightened up substantially to clarify their ideas.

See our comment at the beginning of this letter. We will attempt to describe the reason for the methodology, the hypotheses and the relevance of our work more clearly in the revised manuscript.

# Summary:

The authors articulate their guiding hypotheses, which I think is a good thing to do. I am not sure why (1) is a hypothesis. It seems to be more of a statement about neural networks and is worrisome since I am still concerned that the neural network is just looking at the seasonal cycle and is guaranteed to get a high R2. (2) is odd. Why would we have regional patterns? I could see it if this was a regime-dependent analysis (eg stratus vs convection), but the use of w and LTS as predictors in the neural network should mean that the authors can create a single neural network that effectively does this for them. Why is this not the case? What makes a specific lat-lon box a natural choice.

(3) seems to imply that meteorology plays a secondary role to aerosols, which is not true. We don't expect aerosol to tell us where convection and stratus are, for instance.

1) Neural networks have not been used in this context before, so their capabilities in this context were not quite clear. This is also the case for the separation of aerosol and meteorological effects.

2) While this study does not contrast e.g. stratus vs. convection, we analyze all liquid-water clouds globally. It is clear that these feature different cloud types in different regions and that different processes drive these different clouds. This is shown in figure 6. Regional patterns in aerosol-cloud sensitivity exist. They have been shown to be dependent on meteorology and aerosol species composition (e.g. Andersen et al., 2016). If we created a single neural network, all of the regional characteristics and regionally specific sensitivities (c.f. figure 6) would be blurred or missed completely.

3) Our third hypothesis is certainly not intended to imply that meteorology plays a secondary role to aerosols. We will change the wording for clarity in the revised manuscript.

# References

Adebiyi, A. A. and Zuidema, P. (2016). 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(697):1574–1589.

Adebiyi, A. A., Zuidema, P., and Abel, S. J. (2015). The Convolution of Dynamics and Moisture with the Presence of Shortwave Absorbing Aerosols over the Southeast Atlantic. *Journal of Climate*, 28(5):1997–2024.

Andersen, H., Cermak, J., Fuchs, J., and Schwarz, K. (2016). Global observations of cloud-sensitive aerosol loadings in low-level marine clouds. *Journal of Geophysical Research: Atmospheres*, 121(21):12936–12946.

Bennartz, R. and Rausch, J. (2017). Global and regional estimates of warm cloud droplet number concentration based on 13 years of aqua-modis observations. *Atmospheric Chemistry and Physics Discussions*, 2017:1–32.

Chakraborty, S., Fu, R., Massie, S. T., and Stephens, G. (2016). Relative influence of meteorological conditions and aerosols on the lifetime of mesoscale convective systems. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27):7426–7431.

Chen, Y.-C., Christensen, M. W., Stephens, G. L., and Seinfeld, J. H. (2014). Satellite-based estimate of global aerosol-cloud radiative forcing by marine warm clouds. *Nature Geoscience*, 7(9):643–646.

Coopman, Q., Garrett, T. J., Riedi, J., Eckhardt, S., and Stohl, A. (2016). Effects of long-range aerosol transport on the microphysical properties of low-level liquid clouds in the Arctic. *Atmospheric Chemistry and Physics*, 16(7):4661–4674.

Eastman, R., Wood, R., Eastman, R., and Wood, R. (2016). Factors Controlling Low-Cloud Evolution over the Eastern Subtropical Oceans: A Lagrangian Perspective Using the A-Train Satellites. *Journal of the Atmospheric Sciences*, 73(1):331–351.

George, R. C. and Wood, R. (2010). Subseasonal variability of low cloud radiative properties over the southeast Pacific Ocean. *Atmospheric Chemistry and Physics*, 10(8):4047–4063.

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. (2016). Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability. *Proceedings of the National Academy of Sciences of the United States of America*, 113(21):5804–5811.

Gryspeerdt, E., Quaas, J., and Bellouin, N. (2016). Constraining the aerosol influence on cloud fraction. *Journal of Geophysical Research: Atmospheres*, 121:3566–3583.

Gryspeerdt, E. and Stier, P. (2012). Regime-based analysis of aerosol-cloud interactions. *Geophysical Research Letters*, 39(21):L21802.

Gryspeerdt, E., Stier, P., and Partridge, D. G. (2014). Satellite observations of cloud regime development: the role of aerosol processes. *Atmospheric Chemistry and Physics*, 14(3):1141–1158.

Lacagnina, C. and Selten, F. (2013). A novel diagnostic technique to investigate cloud-controlling factors. *Journal of Geophysical Research: Atmospheres*, 118(12):5979–5991.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, 6(11):2989–3034.

Painemal, D., Kato, S., and Minnis, P. (2014a). 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):11288–11302.

Painemal, D., Road, N., and Stop, M. (2014b). Mean Structure and diurnal cycle of Southeast Atlantic boundary layer clouds: Insights from satellite observations and multiscale modeling framework simulations. *journal of climate*, pages 1–51.

Patel, P. N., Quaas, J., and Kumar, R. (2017). A new statistical approach to improve the satellite based estimation of the radiative forcing by aerosol–cloud interactions. *Atmospheric Chemistry and Physics Discussions*, 17:3687–3698.

Stathopoulos, S., Georgoulias, A., and Kourtidis, K. (2017). Space-borne observations of aerosol - cloud relations for cloud systems of different heights. *Atmospheric Research*, 183:191–201.

Tang, J., Wang, P., Mickley, L. J., Xia, X., Liao, H., Yue, X., Sun, L., and Xia, J. (2014). Positive relationship between liquid cloud droplet effective radius and aerosol optical depth over Eastern China from satellite data. *Atmospheric Environment*, 84:244–253.