

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

4 — Editor review —

5 — RESPONSE TO REFEREES 1 & 2 —

6

7 contact: hendrik.andersen@kit.edu

8 We would like to thank referees 1 and 2 for their review of the revised
9 manuscript and their constructive criticism. In the following, comments by the
10 referees are colored blue, our replies are given in black.

11 **Referee 1**

12 The authors have done a lot of work going through my comments, which I
13 appreciate.

14 One remaining issue is reconciling point (3) on page 5 of the author response
15 document and (4) on page 6. I agree that it is interesting to look at these
16 relationships with an NNA, but I don't understand how the complexity of the
17 NNA is justified in this case. It seems that the system is relatively linear based
18 on the new version of Figure 2. As the authors point out in (4) they are clarifying
19 previous work, which I also think is a good thing to do, but if the NNA doesn't
20 greatly alter the model skill beyond a multiple linear regression then it is more
21 difficult to say that this is an independent test of the original results. To be
22 publishable the paper needs to clarify what the NNA can do that the multiple
23 linear regression cannot. I will admit that I am not very familiar with the
24 literature on NNA, so I appreciate the authors taking time to explain what
25 they are doing. Is there some criterion from the literature for when an NNA
26 is an appropriate choice? Eg. something like Bayesian information criterion
27 that would allow an objective statement of 'NNA is better than multiple linear
28 regression'. If some criterion for NNA use from previous investigations can be
29 met it seems like a reasonable analysis.

30

31 The novelty of our study is that we explicitly consider meteorological con-
32 ditions together with aerosol to explain cloud properties in a multi-variate sta-
33 tistical approach. To do this, we chose an ANN for reasons explained below.

34 However, the choice of the ANN vs. other multi-variate methods is not the main
35 innovation or benefit of this research.

36 We chose the ANN instead of the multiple linear regressions to analyze
37 aerosol-cloud-climate interactions because:

38 1. "*Unlike other statistical techniques the multilayer perceptron makes no*
39 *prior assumptions concerning the data distribution*" (Gardner and Dor-
40 ling, 1998, p.2627).

41 2. "*It can model highly non-linear functions and can be trained to accurately*
42 *generalise when presented with new, unseen data*" (Gardner and Dorling,
43 1998, p.2627).

44 Due to these properties, ANNs have generally been shown to better approx-
45 imate relationships than multiple linear regressions (e.g. Sousa et al., 2007;
46 Badr et al., 2014; Hartmann et al., 2016), especially when modeling nonlinear
47 relationships (Badr et al., 2014), as is the case here. In Hartmann et al.
48 (2016), the ANN is used, as its predictive capability is slightly higher than the
49 multiple linear regression (as in our study). While there does not seem to be
50 an objective criterion to decide when to use an ANN, typically the model with
51 the highest predictive capability is chosen, which is the ANN in our case.

52 More generally, Gardner and Dorling (1998, p.2627) state: "*These features*
53 *[1 and 2 above] of the multilayer perceptron make it an attractive alternative*
54 *to developing numerical models, and also when choosing between statistical*
55 *approaches. As will be seen the multilayer perceptron has many applications in*
56 *the atmospheric sciences*". Olden and Jackson (2002) state in their introduction
57 that "*[t]he utility of ANNs for solving complex pattern recognition problems has*
58 *been demonstrated in many terrestrial [sources] and aquatic studies [sources],*
59 *and has led many researchers to advocate ANNs as an attractive, non-linear*
60 *alternative to traditional statistical methods.*"

61 While we feel that the use of the ANNs in the context of aerosol-cloud-climate
62 interactions is well justified, especially in light of the results, other multivariate
63 statistical methods could have been applied as well, as shown in figure 1.
64 We try to clarify these aspects in two sections of the revised version of the
65 manuscript

66 P.2, L.16–17:” *The ANN is chosen as it is capable of modeling highly nonlinear*
67 *functions and does not need any assumptions concerning the data distribution*
68 *(Gardner and Dorling, 1998). ”.*

69 P.6, L.12–13:” *While the ANN is chosen in this study due to its slightly superior*
70 *predictive capabilities, figure 2 suggests that other multi-variate methods would*
71 *have been appropriate as well. ”.*

72

73 In regards to point 2 in the summary on page 12. Do the regional depen-
74 dencies decrease if SST is added as a predictor? Between LTS, w, RH, BLH
75 and SST that should encompass most of the factors that constitute a regime.
76 However, I acknowledge that different factors should drive different regimes so
77 the point that the authors are making seems reasonable.

78 To answer this question, we computed the ANNs including the SST as an ad-
79 ditional predictor. The addition of SST to the ANNs did not meaningfully
80 improve the predictive capability of the ANNs (c.f. fig. 1).

81 Also, regional dependencies are largely unchanged when SST is included.
82 This is illustrated in fig. 2, where the panels show the global distributions of
83 selected input-output sensitivities that feature strong (nearly identical) spatial
84 patterns both without (left-hand panels) and with SST (right-hand panels) in-
85 cluded in the model.

86 In our study design, we excluded SST from the model to prevent colinearity
87 among the predictors (similar to using LTS and RH at various levels instead of

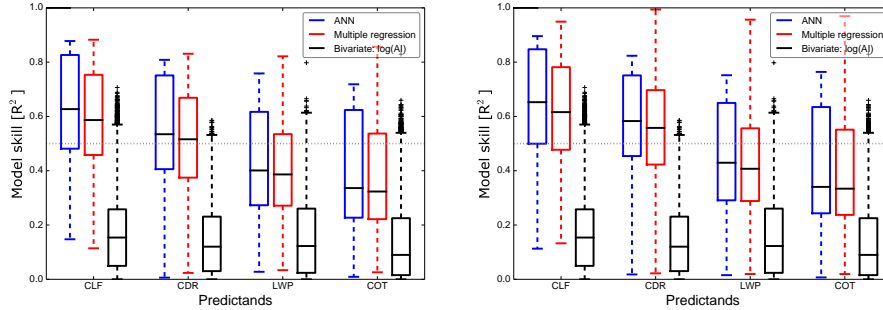


Figure 1: Model skills of ANNs, multiple linear regressions and bivariate (log(AI)- specific cloud property). For ANNs and multiple linear regressions, we included SST (right-hand panel), the version of the manuscript without SST is shown in the left-hand panel.

88 EIS), as a lot of the information content provided by SST is already indirectly
 89 included in the model via the surface temperatures used for the computation of
 90 LTS.
 91 We have added SST as a potential predictor that could have additionally been
 92 included in the model on P.3, L.31.

93 Referee 2

94 I thank the authors for their work in responding to the comments. The majority
 95 of my comments have been addressed, only two things remain. I would suggest
 96 the authors include these at their own discretion. However, some mention of
 97 them might improve the paper still further.

98 The authors have been clear that they would prefer not to include results from
 99 a single global ANN. While creating a single relationship for predicting oceanic
 100 CLF globally may be difficult, this is not as far-fetched as they make it sound,
 101 as this is the central aim of a cloud parametrisation. By making their ANN
 102 regionally dependent, this is similar to including the latitude and longitude in a
 103 parametrisation. I understand that it is difficult to include all of the necessary

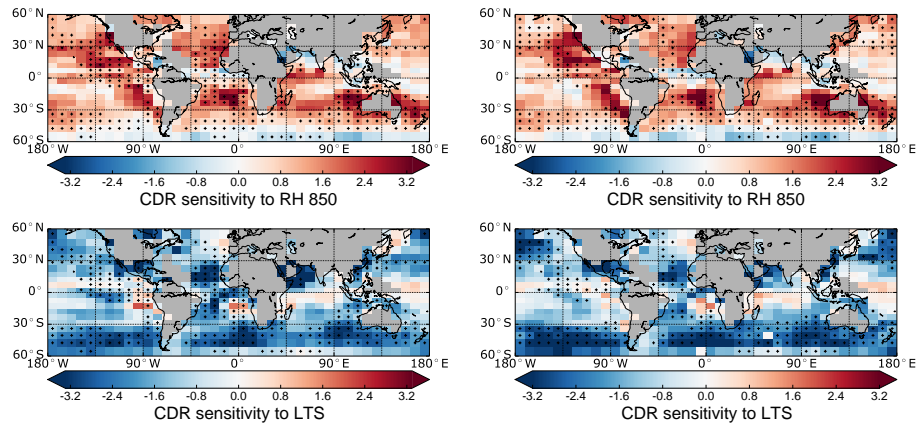


Figure 2: Patterns of the sensitivity of CDR to RH 850 (top) and LTS (bottom). Left-hand panels are computed without SST as in the manuscript, right-hand panels include SST as predictor in the ANN.

104 parameters in a single ANN to predict the CLF (or other parameters), but as
 105 with a cloud parametrisation, the regions or conditions where the global ANN
 106 is deficient would indicate locations for future research.

107 The cloud parameterization is an interesting analogy for using a single global
 108 ANN. When building such a single global ANN, our attempt would be to only
 109 use inputs that are physically related to the output. As such, incoming solar
 110 radiation at the surface could be a good approximation of the latitude. In our
 111 opinion, this is an interesting aspect for future research.

112 Yes, we agree that regions where our current ANN setup does not work quite
 113 as well can be viewed as locations where future research might help to find
 114 out which aspects of the system we are missing. This is an aspect that we
 115 are currently analyzing in a follow-up study. We mention this aspect now in
 116 the revised version of the manuscript (P.6, L.18–P.7, L.3):” *These regions with*
 117 *comparatively low ANN skill may point to regions where the aerosol-cloud-*
 118 *climate system cannot be sufficiently explained with the choice of predictors*
 119 *used in this study and may thus represent regions of interest for future studies.”.*

120

121 Secondly, I think that my previous point about overlying ice cloud was mis-
122 understood. The CLF is the fraction of retrievals where clouds with liquid tops
123 are detected, so any situation where there is overlying ice cloud has the po-
124 tential to reduce the CLF without changing the properties of the underlying
125 liquid cloud. If there are not suitable variables in the ANN to predict the ice
126 cloud fraction, it essentially becomes random noise and would thus limit the
127 predictive ability of the ANN. Using only single-layer cloud retrievals does not
128 address this issue, as they provide only the fraction of retrievals where a single
129 layer liquid water cloud is detected, a subset of CLF. The only way I can see to
130 easily address this issue is to restrict the study to gridboxes where only liquid
131 cloud is detected, ensuring that there is no overlying ice cloud to artificially
132 reduce the CLF. While excluding pixels with ice clouds is not essential (as the
133 ANN clearly already has significant predictive ability), I would recommend that
134 the authors consider it as a method of improving the statistics generated by the
135 ANN.

136 You are right, we misunderstood this point. Yes, the liquid, single layer cloud
137 fraction is the subset of the overall CLF where both liquid cloud tops and single
138 layer situations are detected. As we are using monthly products, it is unclear
139 how we would be able to exclude specific situations as suggested. While we
140 agree that this aspect might weaken the predictive capability of the model, a
141 large effect on the results of this study seems unlikely for two reasons:

- 142 1. Additionally filtering the data for single layer cloud situations (compared
143 to the original manuscript that only filtered for liquid water clouds) did
144 not alter the results meaningfully.
- 145 2. The situations described may average out to some extent in the long-term
146 large-scale data sets used in this study.

147 Also, we are not quantitatively interpreting aerosol effects (e.g. in terms of ra-
148 diative forcing computations), so that small effects do not meaningfully impact
149 our results. We acknowledge and discuss your valid argument in the revised
150 version of the manuscript (P.3, L.7–11):” *One should note that overlying ice-*
151 *water clouds reduce the single-layer liquid-water cloud fraction, without actually*
152 *changing the liquid-water cloud fraction below. This scenario would translate to*
153 *random noise and potentially blur statistical relationships. However, these ef-*
154 *fects are thought to be minor, as these situations are likely to average out to some*
155 *extent in the long-term large-scale data sets used in this study. Also, different*
156 *cloud products were tested for this study that all yielded similar results.”.*

157 **References**

- 158 Badr, H. S., Zaitchik, B. F., Guikema, S. D., Badr, H. S., Zaitchik,
159 B. F., and Guikema, S. D. (2014). Application of Statistical Mod-
160 els to the Prediction of Seasonal Rainfall Anomalies over the Sahel.
161 <http://dx.doi.org/10.1175/JAMC-D-13-0181.1>.
- 162 Gardner, M. and Dorling, S. (1998). Artificial neural networks (the multilayer
163 perceptron) – a review of applications in the atmospheric sciences. *Atmo-*
164 *spheric Environment*, 32(14):2627–2636.
- 165 Hartmann, H., Snow, J. A., Su, B., and Jiang, T. (2016). Seasonal predictions
166 of precipitation in the Aksu-Tarim River basin for improved water resources
167 management. *Global and Planetary Change*, 147:86–96.
- 168 Olden, J. D. and Jackson, D. A. (2002). Illuminating the ”black box”: A
169 randomization approach for understanding variable contributions in artificial
170 neural networks. *Ecological Modelling*, 154(1-2):135–150.

171 Sousa, S., Martins, F., Alvim-Ferraz, M., and Pereira, M. (2007). Multiple linear
172 regression and artificial neural networks based on principal components to
173 predict ozone concentrations. *Environmental Modelling & Software*, 22(1):97–
174 103.

List of important changes

P2 Included sentence for clarity on method selection (Referee 1)

P3 Additional discussion of the properties of the data sets (Referee 2)

P3 Mention of SST as a potential predictor (Referee 1)

P6 Included sentence for clarity on method selection (Referee 1)

P7 Discussion of regions with low ANN skill (Referee 2)

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

Hendrik Andersen^{1,2}, Jan Cermak^{1,2}, Julia Fuchs^{1,2}, Reto Knutti³, and Ulrike Lohmann³

¹Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research

²Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing

³ETH Zürich, Institute of Atmospheric and Climate Science, Zurich, Switzerland

Correspondence to: Hendrik Andersen (hendrik.andersen@kit.edu)

Abstract. The role of aerosols, clouds and their interactions with radiation remain among the largest unknowns in the climate system. Even though the processes involved are complex, aerosol-cloud interactions are often analyzed by means of bivariate relationships. In this study, 15 years (2001–2015) of monthly satellite-retrieved nearly-global aerosol products are combined with reanalysis data of various meteorological parameters to predict satellite-derived marine liquid-water cloud occurrence and properties by means of regionally-specific artificial neural networks. The statistical models used are shown to be capable of predicting clouds, especially in regions of high cloud variability. At this monthly scale, lower tropospheric stability is shown to be the main determinant of cloud fraction and droplet size, especially in stratocumulus regions, while boundary layer height controls the liquid-water amount and thus the optical thickness of clouds. While aerosols show the expected impact on clouds, at this scale they are less relevant than some meteorological factors. Global patterns of the derived sensitivities point to regional characteristics of aerosol and cloud processes.

1 Motivation and aim

Clouds and their microphysical properties play a central role in the Earth’s radiative budget by increasing the albedo but also by interacting with outgoing thermal radiation, leading to a net cooling effect (Boucher et al., 2013). Low-level marine liquid-water clouds are the cloud type with the biggest net cooling effect; their shortwave signal by far exceeds their longwave signal (Hartmann et al., 1992; Wood, 2012; Russell et al., 2013; Chen et al., 2014). A global increase in the occurrence frequency or cooling properties of marine low-level liquid-water clouds could thus offset some of the greenhouse gas warming (Latham et al., 2008). Thus, a complete understanding of the physical processes that determine marine liquid-water clouds and their properties is critical.

Atmospheric aerosols are essential for the formation of clouds, influencing cloud properties as cloud condensation nuclei. An increase in aerosol particles leads to a higher cloud droplet number concentration, and, assuming a constant cloud water content, to smaller droplet radii. This changes the cloud’s radiative properties, as the larger overall droplet surface area increases cloud reflectivity (Twomey, 1977). These changes in droplet number concentration and size are also thought to have ramifications on cloud lifetime (Albrecht, 1989) and cloud vertical extent (Pincus and Baker, 1994). However, these processes are nonlinear (Bréon et al., 2002; Koren et al., 2014; Andersen et al., 2016; Glassmeier and Lohmann, 2016) and dependent on various

environmental conditions that all feature different patterns in time and space (e.g. Loeb and Schuster, 2008; Stevens and Feingold, 2009; Su et al., 2010; Andersen and Cermak, 2015; Andersen et al., 2016).

Even though there have been significant efforts and advances in understanding aerosol-cloud interactions (ACI) over the last 5 decades, the overall scientific understanding is still considered as low (Boucher et al., 2013). This springs from the complexity of ACI and cloud processes themselves, the temporal and spatial scales at which these processes occur, as well as challenges in observing them.

In the satellite observational community, a typical investigative approach to analyze ACI is to directly relate aerosol and cloud observations quantitatively using bivariate statistics, often explicitly considering one or two meteorological variables 10 (e.g. Matsui et al., 2004, 2006; Chen et al., 2014; Andersen and Cermak, 2015). Even though important process inferences have been made on this basis, the limitation of said method set is clearly that the complexity of the processes is not mirrored by the complexity of the statistical method: only selected aspects of the aerosol-cloud system can be analyzed at one time. A multivariate analysis of the relationships between cloud properties and various predictors, including aerosol and meteorological conditions, might be more appropriate for an adequate representation of these atmospheric interactions. In this spirit, this study 15 combines near-global observational and reanalysis data sets as predictors in a multilayer perceptron artificial neural network (ANN) to model near-global marine water cloud occurrence and properties. The [ANN is chosen as it is capable of modeling highly nonlinear functions and does not need any assumptions concerning the data distribution \(Gardner and Dorling, 1998\)](#). The main goal of this study is to identify the main drivers of marine liquid-water cloud occurrence as well as physical and optical properties on a global scale, estimate sensitivities for each predictor, and determine regional patterns therein.

20 The guiding hypotheses are:

1. Neural networks are capable of skillfully modeling cloud patterns on monthly time scales, and allow for a separation and estimates of the relative importance of aerosol and various meteorological factors.
2. Global aerosol and cloud patterns are not only related at a global scale, but regional patterns exist as well.
3. At the spatial and temporal scales considered here, meteorological conditions are more important for cloud occurrence 25 and properties than aerosols.

2 Data and methods

2.1 Data sets

The analysis uses 15 years (2001–2015) of nearly global (60°N–60°S) satellite retrievals and reanalysis fields. Monthly averages of level 3 collection 6 products based on measurements by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on the Terra platform (Levy et al., 2013) are used for information on cloud fraction (CLF; data set: Cloud_Retrieval_Fraction_1L_Liquid_FMean), cloud-top droplet effective radius (CDR; data set: Cloud_Effective_Radius_1L_Liquid_Mean_Mean), cloud liquid water path (LWP; data set: Cloud_Water_Path_1L_Liquid_Mean_Mean) and cloud

optical thickness (COT; data set: Cloud_Optical_Thickness_1L_Liquid_Mean_Mean). While cloud microphysics may also be represented by cloud droplet number concentration, its retrieval requires additional assumptions on vertical cloud water distribution, leading to increased uncertainty (Brenguier et al., 2000), especially in non-stratocumulus cloud regions (Bennartz and Rausch, 2017). As this study investigates liquid-water cloud properties globally, CDR is thus used as a more robust proxy, even though it is also dependent on cloud liquid-water content to some extent (Brenguier et al., 2000). To confine the analysis to liquid-water clouds and to reduce measurement uncertainties due to overlying ice clouds, only single-layer liquid-water cloud products are used. One should note that overlying ice-water clouds reduce the single-layer liquid-water cloud fraction, without actually changing the liquid-water cloud fraction below. This scenario would translate to random noise and potentially blur statistical relationships. However, these effects are thought to be minor, as these situations are likely to average out to some extent in the long-term large-scale data sets used in this study. Also, different cloud products were tested for this study that all yielded similar results. Information on aerosol loading as a proxy for cloud condensation nuclei is provided by aerosol index (AI; computed as a product of the aerosol optical depth (AOD; at $0.55 \mu\text{m}$) and the Ångström exponent (0.55 and $0.867 \mu\text{m}$)). While many studies use the AOD as a proxy for cloud condensation nuclei (e.g. Andreae, 2009; Quaas et al., 2009, 2010; Peters et al., 2012; Koren et al., 2012), the AI has often been found to be a superior measure for this quantity (Stier, 2016), as it weights the fine mode stronger than AOD alone (Nakajima et al., 2001). Some constraints of AI are that it can be affected by aerosol swelling due to hydration in humid environments (Loeb and Schuster, 2008), and that the retrieval describes vertically integrated information and not specifically aerosol at cloud base height where cloud condensation nuclei are typically activated (Shinozuka et al., 2015).

Satellite retrievals are combined with reanalysis data sets from the European Centre for Medium-Range Weather Forecasts (ECMWF) for information on meteorological predictors. The ERA-Interim reanalysis provides data for the time since 1979 and is still continued (Dee et al., 2011). Monthly means of mean daily reanalysis data are used for information on various meteorological predictors at selected atmospheric pressure levels. Meteorological determinants may be grouped into information on relative humidity (RH - at pressure levels 950 hPa (Andersen and Cermak, 2015), 850 hPa (Chen et al., 2014) and 700 hPa (Engström and Ekman, 2010)), vertical velocity (W - at pressure levels 950 hPa, 850 hPa (Kaufman et al., 2005; Engström and Ekman, 2010) and 700 hPa (Engström and Ekman, 2010)), boundary layer height (BLH (Painemal et al., 2014)) and lower tropospheric stability (LTS - computed as the difference in potential temperature between 700 hPa and the surface (Klein and Hartmann, 1993; Chen et al., 2014; Andersen and Cermak, 2015; Andersen et al., 2016)). The reanalysis data used features an original spatial resolution of $0.5^\circ \times 0.5^\circ$ and is subsequently resampled to fit the MODIS $1^\circ \times 1^\circ$ grid.

Typically, clouds form when air cools off, increasing RH. Once supersaturation is reached, water vapor can condense on the CCN. Predictors are selected that are thought to capture this very basic concept well: Vertical velocity and relative humidity are selected as indicators of cloud dynamics and stratification at various pressure levels. CCN are represented by AI, and BLH and LTS describe the large-scale setting. All predictors have been shown to be relevant determinants of liquid-water clouds or their interactions with aerosols in the studies named above. When available, vertically resolved information is preferred to column integrated (e.g. RH at three different pressure levels is preferred to total columnar water vapor), in order to trace processes at various atmospheric levels. While a higher number of reasonable predictors (e.g. sea surface temperature, geopotential height

or horizontal winds as in [Norris and Leovy \(1994\)](#) or Koren et al. (2010)) is likely to marginally increase the skill of the ANN, it would increase model complexity and make interpretation more difficult.

By design the data sets applied in this study average over time and space to specifically study the large-scale changes within the aerosol-cloud-climate system and to allow for future comparisons with global climate models. While on these scales, the causal sequence of cloud processes may not be intact and the processes themselves cannot be observed, their overall ramifications are thought to be represented adequately, in that temporal averaging is intended as a proxy for process relationships.

2.2 Artificial neural networks and study design

10 Basics of artificial neural networks

Machine learning systems consist of a set of numerical operators designed to compute a designated output on given input data. The basic principles, such as the number of numerical links between parameters, are fixed. Artificial neural networks can be described as a branch of machine learning systems. Multilayer perceptrons are a specific type of neural network that are commonly used in the atmospheric sciences and environmental sciences in general, as they are able to model highly nonlinear functions. This type of ANN consists of several layers of interconnected neurons. In general, the architecture of multilayer perceptron ANNs is variable but a typical ANN may consist of an input layer, at least one hidden layer and an output layer. The information from an input pattern is strictly passed from the input layer via the hidden layer(s) to the output layer that yields the desired output pattern (feed-forward ANN). Multilayer perceptron ANNs are fully connected, i.e. each neuron is connected to every neuron in the neighboring layer(s). All connections between neurons in the ANN are specifically weighted so that the information passed to a neuron is the sum of the weighted outputs from the previous layer (net input). The neuron modifies the information by multiplication with a nonlinear transfer function and passes this information through specific weights to all neurons of the following layer (Gardner and Dorling, 1998).

In general, these types of ANNs learn through training. During the training period a subset of the input and output data sets are fed into the ANN. Using this training data, a learning algorithm adjusts the individual weights of each neuron in the network to minimize the error of the output (e.g. the difference between the modeled and observed outputs). The speed of the learning process is adjusted by a learning rate that determines the step size taken during the iterative learning process. While a high learning rate leads to faster convergence, it may miss a global optimum. An additional momentum term adds a fraction of the previous weight change to the current weight change to assist the optimization algorithm out of local minima (Gardner and Dorling, 1998). After the learning algorithm has reached convergence, the predicted output of the network can be compared to the original output for an estimate of model skill. To ensure that the ANN does not only represent the particular data used in the training (overfitting) and is able to generalize the functional relationships underlying the training data, the model is validated using a second independent subset of the input data. If the ANN is able to generalize the relationships between the data sets, the difference between training and validation errors and the overall error are small. The ANN is tested on a third set of independent data to ensure that the model is not overfit to the validation data.

Design of the study and application of the neural network

The ability of the ANN to predict cloud occurrence and properties is dependent not only on an informed choice of predictors, ANN also require sufficient data that fully represent all cases that the ANN is required to generalize, as ANNs perform well for interpolation but poorly for extrapolation (Gardner and Dorling, 1998). In order to circumvent sampling issues and to enable a direct comparability of results in different regions, the near-global data sets are summarized in 40x20 equal area grid cells by aggregating grid cells at the original spatial resolution of $1^\circ \times 1^\circ$. This leads to an increase from the original 180 data points (15 years, 12 months) for each input/output to between 8,000 and 14,000, depending on the number of $1^\circ \times 1^\circ$ pixels that fall into a specific region. Region-specific neural networks are needed to capture regionally varying relationships between cloud properties and their determinants. These relationships feature regional patterns as they depend on liquid-water cloud type, aerosol composition, meteorology and the respective seasonal cycles, all of which exhibit regional patterns (e.g. Stevens and Feingold, 2009; Andersen et al., 2016). These regional characteristics would be blurred or missed completely when using a single global ANN.

The ANN is only applied in grid cells where a minimum of 2000 valid observations exist. In each equal area, an independent ANN is trained over 500 epochs (i.e. number of times the network iterates over the training data) with 60% of the data, validated and tested on 20% of the data each. A simple network topology with one hidden layer consisting of five hidden neurons is applied a) for a more comprehensible model and b) to reduce potential overfitting (Gardner and Dorling, 1998). Multilayer perceptrons with just one hidden layer are frequently used in ecological studies (e.g. Hartmann et al., 2008; Cermak and Knutti, 2009) as they have been shown by several independent studies to be able to approximate any continuous function (Cybenko, 1989; Funahashi, 1989; Hornik et al., 1989; Kecman and Vojislav, 2001; Olden and Jackson, 2002; Di Noia et al., 2013). A hyperbolic tangent is used as the activation function, the weights are initialized randomly from a uniform distribution between -0.1 and 0.1. Gradient descent (Werbos, 1990; Le et al., 2017) is used as the optimization algorithm, with a learning rate of 0.003 and a momentum of 0.01. In-depth testing was undertaken to adjust the details of the model's settings by comparing model skill for a wide number of model setups as in Hartmann et al. (2016). Once the ANN is trained and able to generalize the relationships between the data sets adequately, sensitivity analyses can be conducted. In general, sensitivities are systematically tested by varying each input variable while holding all other input variables constant, e.g., at their average (details see below). In this way the individual contributions of each variable can be analyzed (Olden and Jackson, 2002). A schematic view on the general architecture of the ANN and the training, validation and sensitivity steps is given in Fig. 1.

The ANN skill in modeling the desired outputs is evaluated with the correlation (R^2) between ANN testing output and the corresponding observation data. Sensitivities are only computed for grid cells, where the ANN $R^2 > 0.5$ and the root mean square error relative to the mean (rel. RMSE) $<$ its global average in order to only investigate sensitivities of models that are capable of adequately representing the observed cloud patterns. The derived average sensitivities are only valid for the considered regions and should thus not be interpreted as 'global'. In order to derive a representative and meaningful sensitivity, the mean of ANN-predicted outputs are compared for two groups of input data: all retrievals of a specific predictor smaller than its 25th percentile and all retrievals greater than its 75th percentile; in all cases, the other predictors are held constant at their grid-cell specific mean values. In comparison to a stepwise increase of one specific predictor, a more relevant measure of a

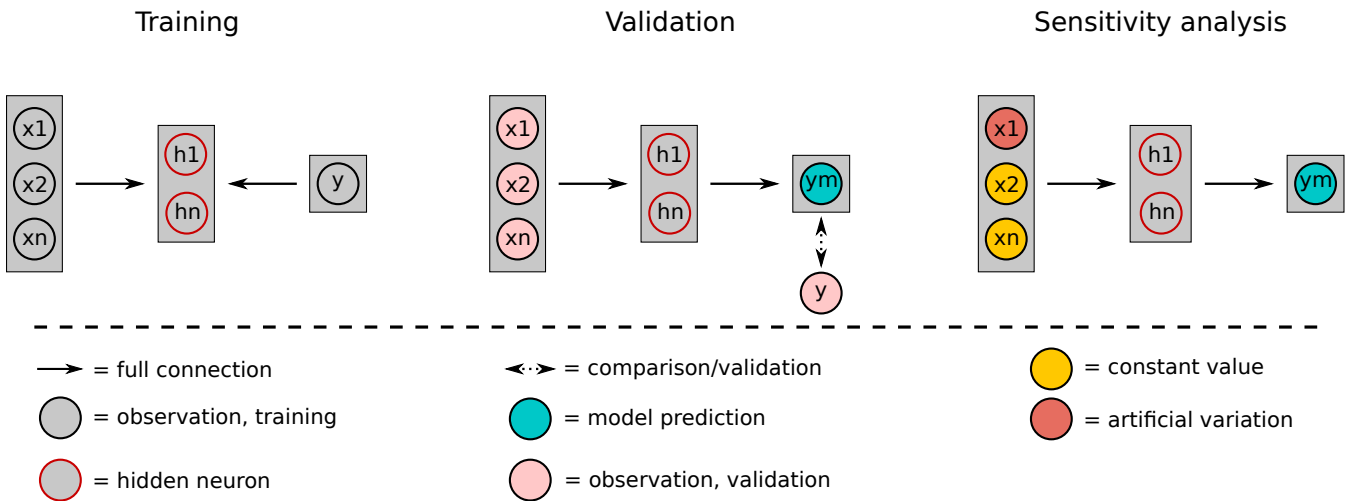


Figure 1. A schematic view on the general architecture and design of multilayer perceptron artificial neural networks. In this study, the ANN features a single hidden layer with 5 neurons.

typical sensitivity can be derived, as the predictor distribution is considered. Thus, in the context of this study, the sensitivity is defined as the mean difference between the predicted cloud property in the groups of low and high predictor values. Typically, the aerosol effect on e.g. CDR is described by the $\delta \log(CDR) / \delta \log(aerosol)$ relationship, where aerosol can be either AOD or AI (e.g. Costantino and Bréon, 2013). While this gives a regionally comparable estimate of the aerosol-cloud sensitivity, it does not explicitly consider the meteorological framework. As the temporal and spatial scales considered in this study are much larger than the actual processes, the calculated sensitivities represent the system scale, and may not match the magnitude of the sensitivities at the process scale (McComiskey et al., 2009; McComiskey and Feingold, 2012).

3 Results and discussion

3.1 Skill of the ANN in predicting cloud occurrence and properties

The skill of the ANNs to predict marine liquid-water cloud occurrence, as well as physical and optical properties is shown in Fig. 2 (blue boxes) and contrasted with the skill of a multiple linear regression using the identical set of predictors (red boxes) and a simple Pearson correlation between $\log(AI)$ and the cloud properties (black boxes). In the ANN, CLF is predicted with the highest accuracy. While for CDR the skill of the ANN is also > 0.5 for many regions, LWP and COT are predicted less accurately. The skill of the multiple linear regression is close to the skill of the ANN, but typically explains a few percent less of the cloud variability, possibly indicating a small contribution in model skill by nonlinear representations of relationships within the ANN. While the ANN is chosen in this study due to its slightly superior predictive capabilities, figure 2 suggests that other multi-variate methods would have been appropriate as well. It is shown that $\log(AI)$ alone typically explains less

than 20 % of the cloud property variability. As a much higher fraction of the cloud variability is explained in the multivariate approaches, the sensitivities derived from the ANN are likely to be more reliable.

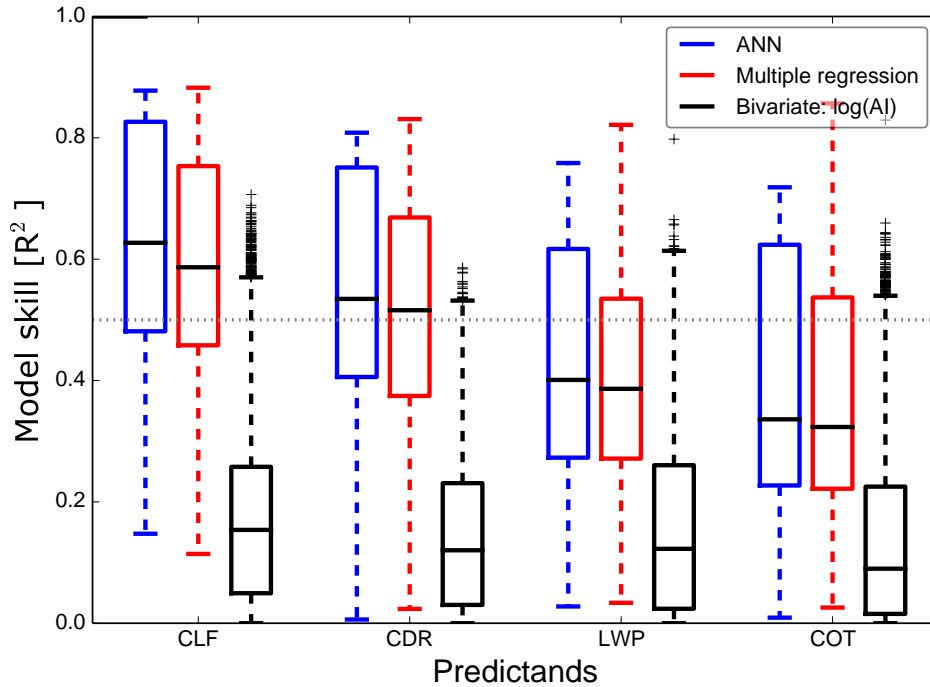


Figure 2. Predictand correlation with ANN test output, multiple linear regression (both multivariate) and $\log(\text{AI})$ (bivariate). The median is represented by the black horizontal line, framed by the interquartile range (boxes), whiskers expand the boxes by 1.5 interquartile ranges.

For all predictands there is a large spread in model skill, leading to distinct regional patterns as illustrated in Fig. 3. The skill of the ANNs is generally higher in the atmospherically stable regions off the western continental coastlines that are dominated by stratocumulus clouds. Less skilled ANNs can generally be found in the (sub-)tropic Pacific and the Indian Ocean. [These regions with comparatively low ANN skill may point to regions where the aerosol-cloud-climate system cannot be sufficiently explained with the choice of predictors used in this study and may thus represent regions of interest for future studies.](#)

The global spatial patterns of ANN skill are likely linked to the spatial patterns of the variability of the specific predictands (Fig. 4). A strong dependence on the variability can be noted for CLF and CDR (Fig. 4a and 4b), i.e. a higher variability enables the ANN to more skillfully represent the inherent relationships. This is sensible, as a higher predictand variability offers the ANN a stronger signal from which it can learn.

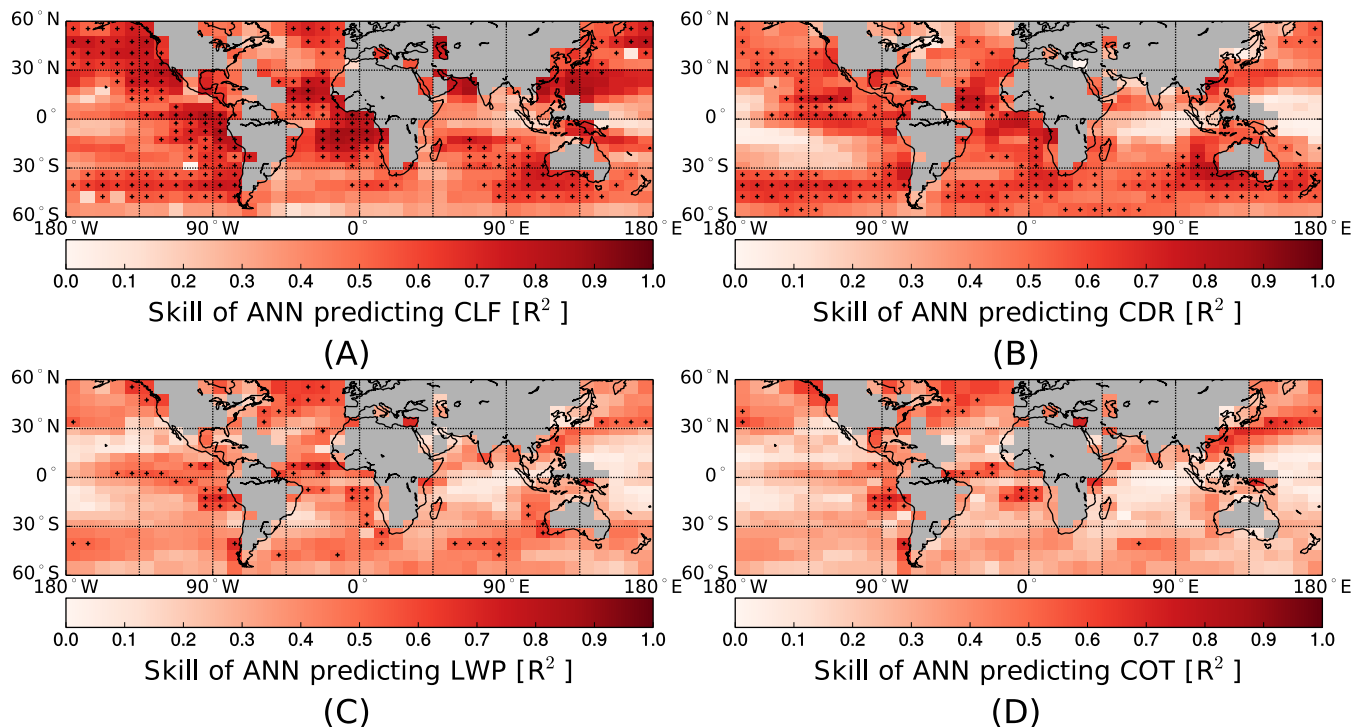


Figure 3. Global patterns of ANN skill [R^2] predicting a) cloud fraction, b) cloud droplet effective radius, c) cloud liquid water path and d) cloud optical thickness. As only ANNs with $R^2 > 0.5$ and rel. RMSE $<$ its global average are used to compute the sensitivities, these are marked by a '+'.

3.2 Determinants of cloud occurrence and properties

Sensitivities are analyzed in all ANNs with a skill of $R^2 > 0.5$ and with a rel. RMSE that is smaller than its global average. Figure 5 shows globally summarized mean and standard deviation of all predictor sensitivities for CLF (Fig. 5a), CDR (Fig. 5b), LWP (Fig. 5c) and COT (Fig. 5d). Positive sensitivities point towards a positive response to an increase in the specific predictor while holding the other predictors constant at their regional average values. CLF shows the greatest sensitivity to LTS, where an increase in LTS leads to a strong increase in CLF, underlining the importance of LTS found in earlier studies (e.g. Klein and Hartmann, 1993; Matsui et al., 2004; Andersen and Cermak, 2015). CLF is also positively related to relative humidity at all assessed pressure levels, with the strongest sensitivity at 950 hPa, where stratocumulus clouds and transitional clouds between stratocumulus and shallow cumulus are located (Gryspeerdt and Stier, 2012; Andersen and Cermak, 2015). While boundary layer height and aerosol are also positively connected to CLF, W sensitivity varies in sign. Sensitivities associated with W can generally be interpreted as the change in the predictand when W changes from updrafts to downdrafts. The most relevant pressure level in terms of W seems to be 700 hPa, with strong positive sensitivities, illustrating that the downdrafts at 700 hPa associated with stable conditions in the lower troposphere correspond to an increase in CLF. In terms of CDR sensitivities, LTS

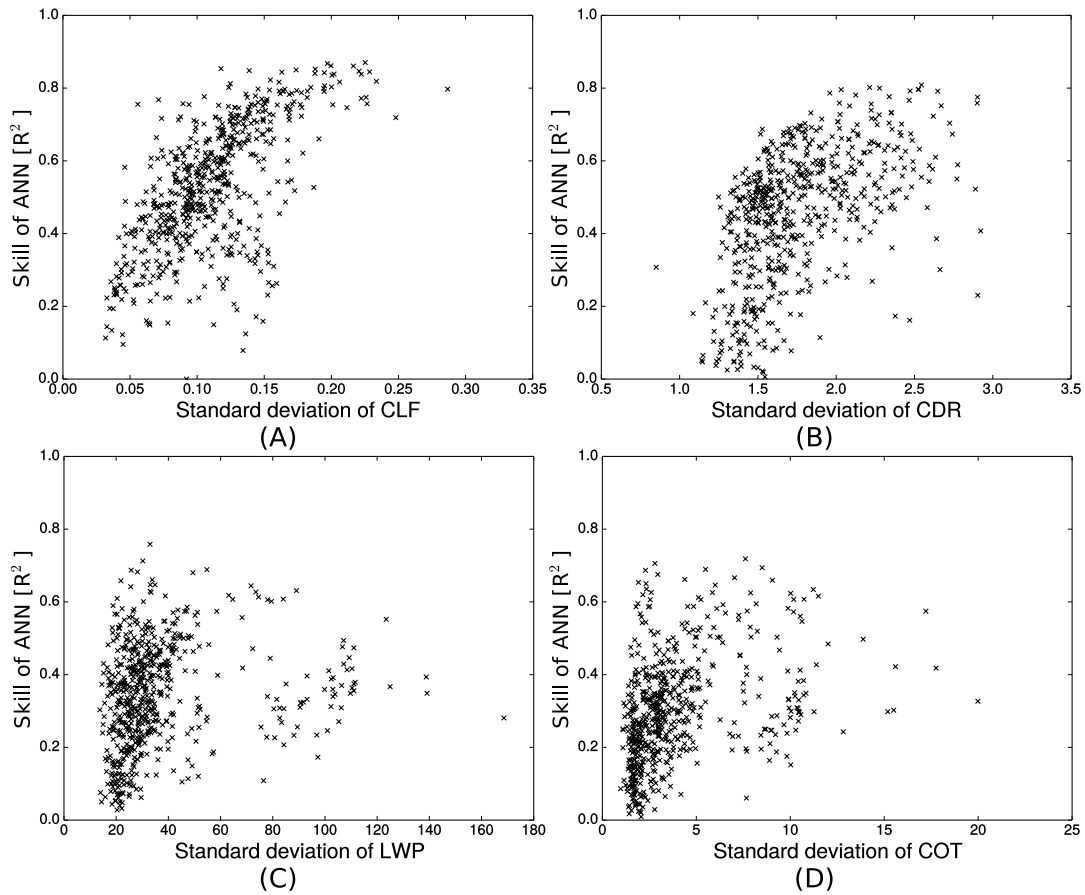


Figure 4. The skill of the ANN predicting a) CLF, b) CDR, c) COT and d) LWP as a function of the predictand variability (standard deviation). Each point illustrates the combination of skill and variability for a specific equal area region (pixel regions in figure 3).

also displays the strongest effect, with an increase in LTS connected to a distinct reduction in droplet size. RH at 850 hPa exerts the strongest positive CDR sensitivity, with many of the cloud tops located at this pressure level. AI has a notable sensitivity, showing a distinct negative association to CDR as previously assumed. Generally, updrafts favor larger CDR, with a stronger sensitivity at higher altitudes. Results of LWP and COT sensitivities are similar in terms of the signs and magnitudes of the individual sensitivities. Both are mainly determined by BLH and LTS, both positively associated with the respective cloud property. RH facilitates thicker clouds containing more liquid water, especially free tropospheric relative humidity at 700 hPa seems to have a positive impact on LWP and COT, as higher humidity levels at 700 hPa are likely to weaken drying effects of entraining air masses (Ackerman et al., 2004; Chen et al., 2012, 2014). While increases in aerosol lead to a negative LWP response, this does not lead to a similarly strong COT reduction. W is negatively related to both cloud properties, as situations with updrafts generally produce thicker clouds, the most relevant pressure level is at 850 hPa.

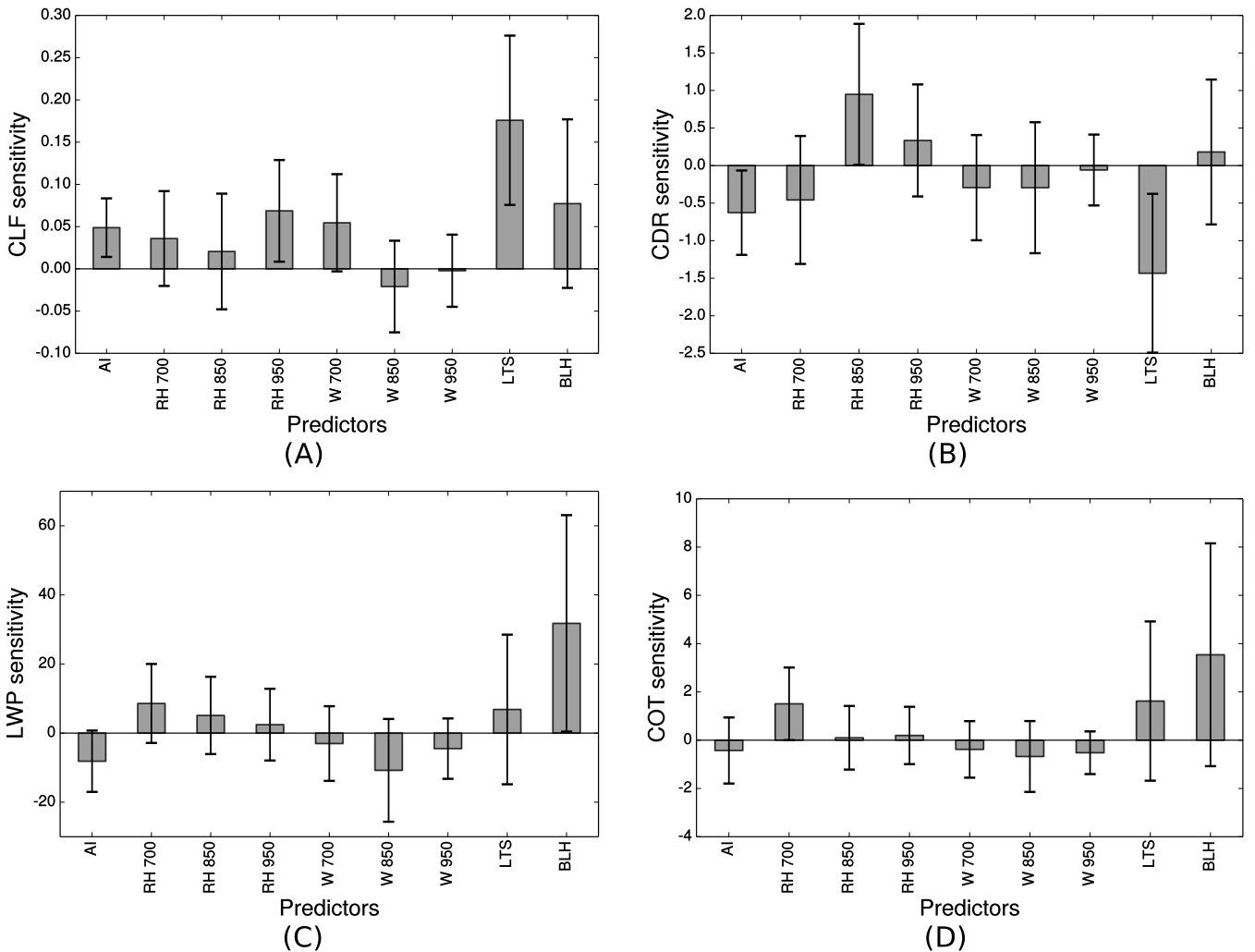


Figure 5. Global mean relative sensitivities as defined in section 2.3 of a) CLF, b) CDR, c) LWP and d) COT for all predictors of the ANNs (x axes). Error bars illustrate the regional variability of the sensitivities (global standard deviation).

The application of an individual ANN in every grid cell enables the analysis of regional patterns of the derived sensitivities. Panels of the left-hand column of Fig. 6 show regional patterns of CLF sensitivities, the panels of the right-hand column show regional patterns of CDR sensitivities. The range of the colorbars is identical within each column, so that both the overall magnitude as well as the spatial patterns of the sensitivities can be compared. LTS is the strongest determinant for CLF and is positively related to CLF everywhere on Earth, with especially strong sensitivities in atmospherically stable regions off the western coasts of the continents, where stratocumulus clouds are predominant (Klein and Hartmann, 1993; Russell et al., 2013). In these regions RH shows a strong positive CLF sensitivity at 950 hPa, pointing to the relevance of low-level humidity in these regions of low boundary layer clouds. Liquid water cloud fraction in the intertropical convergence zone is more sensitive to

5 RH at 850 hPa, reflecting the thicker boundary layer in this region. The most pronounced relation between the aerosol and CLF can be found at latitudes around 30°, especially over the Northwest Pacific.

CDR is markedly reduced by AI in the Northwest Pacific and the Southwest Atlantic and negatively associated with AI to a lesser degree in most other marine regions. The region close to the coastline of the Arabian Sea is an exception. Here, dust particles make up a significant portion of the aerosol species composition (Prospero, 1999; Kaufman et al., 2005), which may lead to larger droplet sizes when dust aerosols act as giant CCN (Levin et al., 2005; Barahona et al., 2010). The Southeast Atlantic features weakly positive sensitivities of CDR to changes in AI. While CDR is typically found to be negatively related to AI in the Southeast Atlantic (e.g. Costantino and Bréon, 2013), Andersen and Cermak (2015) found that AI and CDR can
5 be positively associated in very stable atmospheric conditions. Issues of sampling (few aerosol retrievals in high CLF regions), scale (highly aggregated data) or their combination might affect the observed CDR sensitivity to AI in this region. One should note that sensitivity maps were also produced using AOD as a proxy for cloud condensation nuclei instead of AI. While the overall results were very similar, changes in sensitivity of CDR to the aerosol quantities were observed in the Northeast Atlantic that is dominated by Saharan dust aerosols. Here, the difference between AOD and AI is substantial due to the abundance of
10 coarse dust particles. LTS is negatively associated with CDR, especially south of 30° and in the subtropical Atlantic as found by Matsui et al. (2006), as high LTS environments are connected with weaker updrafts and a shallower boundary layer, limiting cloud droplet growth. This excludes the Southeast Atlantic, where stable conditions may trap the humidity in the boundary layer (Johnson et al., 2004; Painemal et al., 2014; Andersen and Cermak, 2015). Similar effects may occur in the Southeast Pacific as well. RH features the strongest positive CDR sensitivity at 850 hPa with distinctly strong sensitivities in the subtropic
15 regions, where cloud tops are frequently located at this pressure level (Gryspeerd and Stier, 2012). Compared to these factors, W at 700 hPa seems to be a relevant determinant in very selected, mostly tropical regions only.

4 Summary and conclusions

The central aim of this study was to identify and analyze the main determinants of marine liquid-water clouds and their sensitivities at the system scale. Artificial neural networks were shown capable of predicting cloud patterns on a global scale
20 well, although ANN skill is dependent on the cloud property and its variability. Regions with a strong monthly variability such as the stratocumulus regions that feature a strong seasonal cycle are most skillfully represented.

Sensitivities were derived for all predictor-predictand combinations, revealing LTS to be the main determinant of monthly liquid-water cloud occurrence and properties. LTS is positively related to CLF on a global scale, with especially strong regional sensitivities in the subsidence regions and the mid-latitudes. In most of these regions, LTS features a strong negative sensitivity
25 towards CDR. One exception to this negative CDR-LTS relationship is the Southeast Atlantic, where high LTS conditions may trap humidity in the boundary layer, causing larger CDR and hence a positive CDR-LTS relationship (Johnson et al., 2004; Painemal et al., 2014; Andersen and Cermak, 2015). The sensitivity of cloud properties to changes in relative humidity is dependent on both region and pressure level. CLF in regions that feature predominantly stratocumulus clouds or other low-level clouds is most sensitive to RH at 950 hPa, whereas tropical regions with thicker boundary layers are more sensitive to

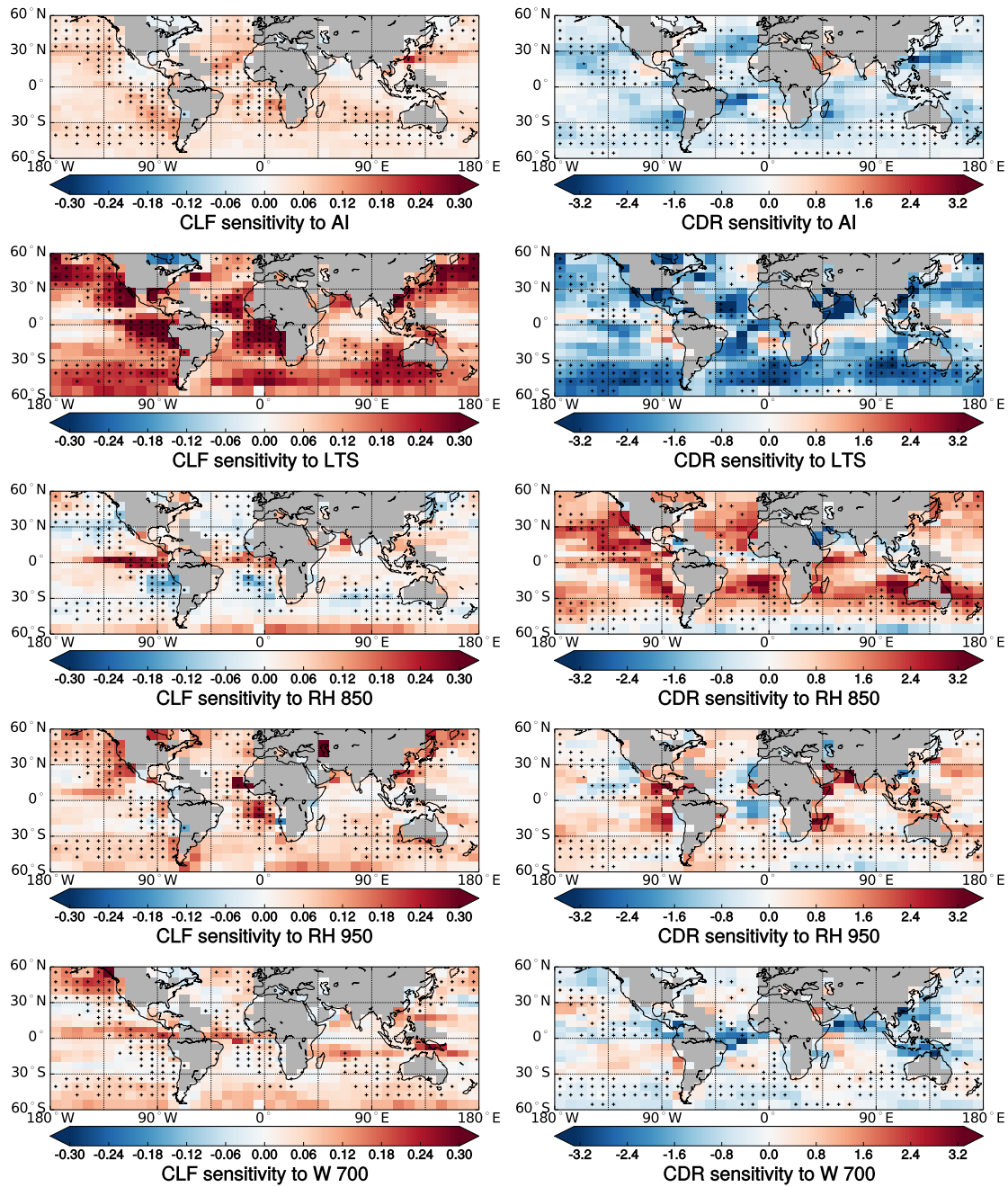


Figure 6. Global relative sensitivity patterns of selected CLF predictors in the left-hand column and CDR predictors in the right-hand column. Gray regions a) are over land or b) do not contain at least 2000 data points. Regions where the ANN test skill (R^2) is > 0.5 and rel. RMSE $<$ its global average are marked with a '+'.

30 RH at higher altitudes. CDR sensitivity to RH is stronger at higher pressure levels, where the cloud-tops are likely located. In addition to this, BLH is found to be a main determinant of LWP and COT. One should note though, that not all of the observed predictor-predictand sensitivities are necessarily a result of a direct physical relationship between the predictor and the predictand, but may in part be spurious due to cloud contamination of the satellite aerosol retrievals (Grandey et al., 2013), or due to the influence of confounding factors on both predictor and predictand (Gryspeerd et al., 2016). For example, CLF and AI/AOD are both positively related to RH, potentially contributing to the observed positive CLF sensitivity to AI/AOD. Gryspeerd et al. (2016) found that, by constraining potential aerosol-induced effects on CLF to situations where cloud droplet
5 number concentration is simultaneously increased, the MODIS log(AOD)-CLF relationship is reduced by about 80 %. Issues of this kind are addressed here by including information on all relevant confounding factors directly in the ANN, and - for comparison measure - when all other inputs are held constant at their grid-cell specific average, the log(AI)-CLF relationship is on average about 40 % weaker than the originally observed log(AI)-CLF relationship. While the decrease in the sensitivity is not quite as strong as in Gryspeerd et al. (2016), the results correspond well in the sense that bivariate determined aerosol-
10 cloud sensitivities as in Quaas et al. (2008) are likely to overestimate aerosol indirect effects significantly. While the influence of confounding factors is limited by the multivariate approach, effects concerning data quality (e.g. cloud contamination) are not accounted for and need to be considered, especially when interpreting the CLF sensitivity to AI.

The ramifications of the interactions between aerosols and cloud occurrence and properties seem to be represented well in the ANN, following the general understanding of ACI. Specific regions of interest arise, such as the Northwest Atlantic with
15 strong sensitivities to AI and regions that are affected by high dust loadings, with positive AI-CDR relationships and an above average positive AI-CLF sensitivity.

The results lead to the conclusion that on the system scale the aerosol may be viewed as a relevant determinant of marine liquid-water cloud fraction and microphysical properties, but only a secondary determinant for cloud optical thickness. On the scales considered here, lower tropospheric stability is the key controlling factor of cloud occurrence and droplet size, while
20 boundary layer height controls the liquid water path and thus optical thickness of the cloud. The results confirm findings of previous studies that analyzed determinants of cloud properties in a more isolated manner (e.g. Klein and Hartmann, 1993; Johnson et al., 2004; Matsui et al., 2006). The results give confidence that the combination of observational and reanalysis data sets in a multivariate statistical approach is able to capture the natural variability of cloud occurrence and properties, and that meteorological and aerosol effects similar to those found in other studies can be identified in this system. In the future, a focussed, cloud-regime specific ANN approach similar to Gryspeerd and Stier (2012) or Oreopoulos et al. (2016) could add to our system understanding. To address climate effects in a straight-forward manner, future research may also apply this study's approach to investigate the global determinants of cloud radiative effects.

5 5 Data availability

MODIS data used in this study were acquired as part of the NASA's Earth-Sun System Division and archived and distributed by the MODIS Adaptive Processing System (MODAPS). MODIS data were obtained from the Goddard Space Flight Cen-

ter (<http://ladsweb.nascom.nasa.gov/data/search.html>). ECMWF ERA-Interim data used in this study were obtained from the ECMWF data server (<http://apps.ecmwf.int/datasets/data/interim-full-moda/levtype=sfc/>).

- 10 *Author contributions.* J. Cermak had the initial idea and performed a precursor study. H. Andersen fully developed the method and the software, obtained and analyzed the data sets, conducted the original research and wrote the manuscript. H. Andersen, J. Cermak, J. Fuchs, R. Knutti and U. Lohmann contributed to study design and interpretation of findings.

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

- Acknowledgements.* Funding for this study was provided by Deutsche Forschungsgemeinschaft (DFG) in the project GEOPAC (grant CE
15 163/5-1). Constructive and helpful comments from three anonymous reviewers are gratefully acknowledged.

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, doi:10.1038/nature03137.1., 2004.
- Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, *Science*, 245, 1227–1230, doi:10.1126/science.245.4923.1227, 20 1989.
- Andersen, H. and Cermak, J.: How thermodynamic environments control stratocumulus microphysics and interactions with aerosols, *Environmental Research Letters*, 10, 24 004, doi:10.1088/1748-9326/10/2/024004, <http://dx.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-level marine clouds, *Journal of Geophysical Research: Atmospheres*, 121, 12 936–12 946, doi:10.1002/2016JD025614, <http://doi.wiley.com/10.1002/2016JD025614>, 2016.
- Andreae, M. O.: Correlation between cloud condensation nuclei concentration and aerosol optical thickness in remote and polluted regions, *Atmospheric Chemistry and Physics*, 9, 543–556, doi:10.5194/acp-9-543-2009, 2009.
- Barahona, D., West, R. E. L., Stier, P., Romakkaniemi, S., Kokkola, H., and Nenes, A.: Comprehensively accounting for the effect of giant CCN in cloud activation parameterizations, *Atmospheric Chemistry and Physics*, 10, 2467–2473, 2010.
- 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 Discussions*, 2017, 1–32, doi:10.5194/acp-2016-1130, <http://www.atmos-chem-phys-discuss.net/acp-2016-1130/>, 2017.
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X.-Y.: Clouds and Aerosols, in: *Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, [Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.
- Brenguier, J.-L., Pawlowska, H., Schüller, L., Preusker, R., Fischer, J., and Fouquart, Y.: Radiative Properties of Boundary Layer Clouds: Droplet Effective Radius versus Number Concentration, *Journal of the Atmospheric Sciences*, 57, 803–821, doi:10.1175/1520-0469(2000)057<0803:RPOBLC>2.0.CO;2, [http://journals.ametsoc.org/doi/abs/10.1175/1520-0469\(2000\)057%3C0803:RPOBLC%3E2.0.CO;2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(2000)057%3C0803:RPOBLC%3E2.0.CO;2), 2000.
- Bréon, F.-M., Tanré, D., and Generoso, S.: Aerosol effect on cloud droplet size monitored from satellite, *Science*, 295, 834–838, doi:10.1126/science.1066434, <http://www.ncbi.nlm.nih.gov/pubmed/11823636>, 2002.
- Cermak, J. and Knutti, R.: Beijing Olympics as an Aerosol Field Experiment, *Geophysical Research Letters*, 36, L10 806, doi:10.1029/2009GL038572, <http://www.iac.ethz.ch/people/knuttir/papers/cermak09grl.pdf>, 2009.
- Chen, Y.-C., Christensen, M. W., Xue, L., Sorooshian, A., Stephens, G. L., Rasmussen, R. M., and Seinfeld, J. H.: Occurrence of lower cloud albedo in ship tracks, *Atmospheric Chemistry and Physics*, 12, 8223–8235, doi:10.5194/acp-12-8223-2012, <http://www.atmos-chem-phys.net/12/8223/2012/>, 2012.
- 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, doi:10.1038/ngeo2214, 2014.

- Costantino, L. and Bréon, F.-M.: Aerosol indirect effect on warm clouds over South-East Atlantic, from co-located MODIS and CALIPSO observations, *Atmospheric Chemistry and Physics*, 13, 69–88, doi:10.5194/acp-13-69-2013, <http://www.atmos-chem-phys.net/13/69/2013/>, 2013.
- Cybenko, G.: Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals, and Systems*, 2, 303–314, doi:10.1007/BF02551274, <http://link.springer.com/10.1007/BF02551274>, 1989.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, doi:10.1002/qj.828, <http://doi.wiley.com/10.1002/qj.828>, 2011.
- Di Noia, A., Sellitto, P., Frate, F. D., and De Laat, J.: Global tropospheric ozone column retrievals from OMI data by means of neural networks, *Atmospheric Measurement Techniques*, 6, 895–915, doi:10.5194/amt-6-895-2013, www.atmos-meas-tech.net/6/895/2013/, 2013.
- Engström, A. and Ekman, A. M. L.: Impact of meteorological factors on the correlation between aerosol optical depth and cloud fraction, *Geophysical Research Letters*, 37, L18 814, doi:10.1029/2010GL044361, <http://doi.wiley.com/10.1029/2010GL044361>, 2010.
- Funahashi, K.-I.: On the approximate realization of continuous mappings by neural networks, *Neural Networks*, 2, 183–192, doi:10.1016/0893-6080(89)90003-8, <http://linkinghub.elsevier.com/retrieve/pii/0893608089900038>, 1989.
- Gardner, M. and Dorling, S.: Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences, *Atmospheric Environment*, 32, 2627–2636, doi:10.1016/S1352-2310(97)00447-0, 1998.
- Glassmeier, F. and Lohmann, U.: Constraining precipitation susceptibility of warm, ice- and mixed-phase clouds with microphysical equations, *Journal of the Atmospheric Sciences*, 73, 5003–5023, doi:10.1175/JAS-D-16-0008.1, <http://journals.ametsoc.org/doi/10.1175/JAS-D-16-0008.1>, 2016.
- Grandey, B. S., Stier, P., and Wagner, T. M.: Investigating relationships between aerosol optical depth and cloud fraction using satellite, aerosol reanalysis and general circulation model data, *Atmospheric Chemistry and Physics*, 13, 3177–3184, doi:10.5194/acp-13-3177-2013, 2013.
- Gryspeerd, E. and Stier, P.: Regime-based analysis of aerosol-cloud interactions, *Geophysical Research Letters*, 39, L21 802, doi:10.1029/2012GL053221, <http://doi.wiley.com/10.1029/2012GL053221>, 2012.
- Gryspeerd, E., Quaas, J., and Bellouin, N.: Constraining the aerosol influence on cloud fraction, *Journal of Geophysical Research: Atmospheres*, 121, 3566–3583, doi:10.1002/2015JD023744, 2016.
- 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, doi:10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2, 1992.
- Hartmann, H., Becker, S., King, L., and Jiang, T.: Forecasting water levels at the Yangtze River with neural networks, *Erdkunde*, 62, 231–243, doi:10.3112/erdkunde.2008.03.04, 2008.
- Hartmann, H., Snow, J. A., Su, B., and Jiang, T.: Seasonal predictions of precipitation in the Aksu-Tarim River basin for improved water resources management, *Global and Planetary Change*, 147, 86–96, doi:10.1016/j.gloplacha.2016.10.018, 2016.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359–366, doi:10.1016/0893-6080(89)90020-8, <http://linkinghub.elsevier.com/retrieve/pii/0893608089900208>, 1989.

- Johnson, B., Shine, K., and Forster, P.: The semi-direct aerosol effect: Impact of absorbing aerosols on marine stratocumulus, *Quarterly Journal of the Royal Meteorological Society*, 130, 1407–1422, doi:10.1256/qj.03.61, <http://doi.wiley.com/10.1256/qj.03.61>, 2004.
- 20 Kaufman, Y. J., Koren, I., Remer, L., Rosenfeld, D., and Rudich, Y.: The effect of smoke, dust, and pollution aerosol on shallow cloud development over the Atlantic Ocean, *Proceedings of the National Academy of Sciences of the United States of America*, 102, 11 207–11 212, doi:10.1073/pnas.05051911102, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1182178&tool=pmcentrez&rendertype=abstract>, 2005.
- Kecman, V. V. and Vojislav: *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*, MIT Press, <http://dl.acm.org/citation.cfm?id=559026>, 2001.
- 25 Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, *Journal of Climate*, 6, 1587–1606, doi:10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.
- Koren, I., Feingold, G., and Remer, L. A.: The invigoration of deep convective clouds over the Atlantic: Aerosol effect, meteorology or retrieval artifact?, *Atmospheric Chemistry and Physics*, 10, 8855–8872, doi:10.5194/acp-10-8855-2010, 2010.
- Koren, I., Altaratz, O., Remer, L. A., Feingold, G., Martins, J. V., and Heiblum, R. H.: Aerosol-induced intensification of rain from the tropics to the mid-latitudes, *Nature Geoscience*, 5, 118–122, doi:10.1038/ngeo1364, <http://www.nature.com/doi/10.1038/ngeo1364>, 2012.
- 30 Koren, I., Dagan, G., and Altaratz, O.: From aerosol-limited to invigoration of warm convective clouds, *Science*, 344, 1143–1146, doi:10.1126/science.1252595, <http://www.sciencemag.org/content/344/6188/1143.full>, 2014.
- Latham, J., Rasch, P., Chen, C.-C., Kettles, L., Gadian, A., Gettelman, A., Morrison, H., Bower, K., and Choulaton, T.: Global temperature stabilization via controlled albedo enhancement of low-level maritime clouds, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366, 3969–3987, doi:10.1098/rsta.2008.0137, <file:///Users/yangyang/YANGYANGXU/work/lynn{ }research/papersfromlynn/Lathametal.-2008-Philosophicaltransactions.SeriesA,Mathematical,physical,andengineeringsciences-Globaltemperaturestabilizationviacontrolledalbedoenhancementoflow>, 2008.
- 35 Le, J., El-Askary, H., Allali, M., and Struppa, D.: Application of recurrent neural networks for drought projections in California, *Atmospheric Research*, 188, 100–106, doi:10.1016/j.atmosres.2017.01.002, <http://linkinghub.elsevier.com/retrieve/pii/S0169809517300157>, 2017.
- Levin, Z., Teller, A., Ganor, E., and Yin, Y.: On the interactions of mineral dust, sea-salt particles, and clouds: A measurement and modeling study from the Mediterranean Israeli Dust Experiment campaign, *Journal of Geophysical Research: Atmospheres*, 110, 1–19, doi:10.1029/2005JD005810, 2005.
- 5 Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, *Atmospheric Measurement Techniques*, 6, 2989–3034, doi:10.5194/amt-6-2989-2013, <http://www.atmos-meas-tech.net/6/2989/2013/amt-6-2989-2013.html>, 2013.
- 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, D14 214, doi:10.1029/2007JD009763, <http://doi.wiley.com/10.1029/2007JD009763>, 2008.
- 10 Matsui, T., Masunaga, H., and Pielke Sr., R. A.: Impact of aerosols and atmospheric thermodynamics on cloud properties within the climate system, *Geophysical Research Letters*, 31, L06 109, doi:10.1029/2003GL019287, <http://doi.wiley.com/10.1029/2003GL019287>, 2004.
- Matsui, T., Masunaga, H., Kreidenweis, S. M., Pielke, R. A., Tao, W.-K., Chin, M., and Kaufman, Y. J.: Satellite-based assessment of marine low cloud variability associated with aerosol, atmospheric stability, and the diurnal cycle, *Journal of Geophysical Research: Atmospheres*, 15 111, D17 204, doi:10.1029/2005JD006097, <http://doi.wiley.com/10.1029/2005JD006097>, 2006.

- McComiskey, A. and Feingold, G.: The scale problem in quantifying aerosol indirect effects, *Atmospheric Chemistry and Physics*, 12, 1031–1049, doi:10.5194/acp-12-1031-2012, <http://www.atmos-chem-phys.net/12/1031/2012/acp-12-1031-2012.html>, 2012.
- McComiskey, A., Feingold, G., Frisch, a. S., Turner, D. D., Miller, M. a., Chiu, J. C., Min, Q., and Ogren, J. a.: An assessment of aerosol-cloud interactions in marine stratus clouds based on surface remote sensing, *Journal of Geophysical Research: Atmospheres*, 114, D09 203, doi:10.1029/2008JD011006, <http://doi.wiley.com/10.1029/2008JD011006>, 2009.
- 20 Nakajima, T., Higurashi, A., Kawamoto, K., and Penner, J. E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters, *Geophysical Research Letters*, 28, 1171–1174, doi:10.1029/2000GL012186, 2001.
- Norris, J. R. and Leovy, C. B.: Interannual Variability in Stratiform Cloudiness and Sea Surface Temperature, *Journal of Climate*, 7, 1915–1925, doi:10.1175/1520-0442(1994)007<1915:IVISCA>2.0.CO;2, [https://doi.org/10.1175/1520-0442\(1994\)007<1915:IVISCA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1915:IVISCA>2.0.CO;2), 1994.
- 25 Olden, J. D. and Jackson, D. A.: Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks, *Ecological Modelling*, 154, 135–150, doi:10.1016/S0304-3800(02)00064-9, 2002.
- Oreopoulos, L., Cho, N., Lee, D., and Kato, S.: Radiative effects of global MODIS cloud regimes, *Journal of Geophysical Research: Atmospheres*, 121, 2299–2317, doi:10.1002/2015JD024502, 2016.
- 30 Painemal, D., Kato, S., and Minnis, P.: 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, 11 288–11 302, doi:10.1002/2014JD022182, <http://doi.wiley.com/10.1002/2014JD022182>, 2014.
- Peters, K., Stier, P., Quaas, J., and Graßl, H.: Aerosol indirect effects from shipping emissions: sensitivity studies with the global aerosol-climate model ECHAM-HAM, *Atmospheric Chemistry and Physics*, 12, 5985–6007, doi:10.5194/acp-12-5985-2012, <http://www.atmos-chem-phys.net/12/5985/2012/>, 2012.
- 35 Pincus, R. and Baker, M. B.: Effect of precipitation on the albedo susceptibility of clouds in the marine boundary layer, *Nature*, 372, 250–252, doi:10.1038/372250a0, 1994.
- Prospero, J. M.: Long-range transport of mineral dust in the global atmosphere: impact of African dust on the environment of the southeastern United States, *Proceedings of the National Academy of Sciences of the United States of America*, 96, 3396–403, doi:10.1073/PNAS.96.7.3396, <http://www.ncbi.nlm.nih.gov/pubmed/10097049><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC34280>, 1999.
- 5 Quaas, J., Boucher, O., Bellouin, N., and Kinne, S.: Satellite-based estimate of the direct and indirect aerosol climate forcing, *Journal of Geophysical Research*, 113, D05 204, doi:10.1029/2007JD008962, 2008.
- Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristjánsson, J. E., Liu, X., Balkanski, Y., Donner, L. J., Ginoux, P. A., Stier, P., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevåg, A., Iversen, T., Seland, Ø., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J., Iacono, M. J., Kinne, S., and Schulz, M.: Aerosol indirect effects - general circulation model intercomparison and evaluation with satellite data, *Atmospheric Chemistry and Physics*, 9, 8697–8717, doi:10.5194/acp-9-8697-2009, 2009.
- 10 Quaas, J., Stevens, B., Stier, P., and Lohmann, U.: Interpreting the cloud cover - aerosol optical depth relationship found in satellite data using a general circulation model, *Atmospheric Chemistry and Physics*, 10, 6129–6135, <http://www.atmos-chem-phys.net/10/6129/2010/>, 2010.
- Russell, L. M., Sorooshian, A., Seinfeld, J. H., Albrecht, B. A., Nenes, A., Ahlm, L., Chen, Y.-C., Coggon, M., Craven, J. S., Flagan, R. C., Frossard, A. A., Jonsson, H., Jung, E., Lin, J. J., Metcalf, A. R., Modini, R., Mülmenstädt, J., Roberts, G., Shingler, T., Song, S., Wang, Z.,
- 15

- and Wonaschütz, A.: Eastern Pacific Emitted Aerosol Cloud Experiment, *Bulletin of the American Meteorological Society*, 94, 709–729, doi:10.1175/BAMS-D-12-00015.1, <http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-12-00015.1>, 2013.
- Shinozuka, Y., Clarke, A. D., Nenes, A., Jefferson, A., Wood, R., McNaughton, C. S., Ström, J., Tunved, P., Redemann, J., Thornhill, K. L., Moore, R. H., Latham, T. L., Lin, J. J., and Yoon, Y. J.: The relationship between cloud condensation nuclei (CCN) concentration and light extinction of dried particles: indications of underlying aerosol processes and implications for satellite-based CCN estimates, *Atmospheric Chemistry and Physics*, 15, 7585–7604, doi:10.5194/acp-15-7585-2015, <http://www.atmos-chem-phys.net/15/7585/2015/>, 2015.
- 445 Stevens, B. and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system, *Nature*, 461, 607–613, doi:10.1038/nature08281, <http://www.ncbi.nlm.nih.gov/pubmed/19794487>, 2009.
- Stier, P.: Limitations of passive remote sensing to constrain global cloud condensation nuclei, *Atmospheric Chemistry and Physics*, 16, 6595–6607, doi:10.5194/acp-16-6595-2016, <http://www.atmos-chem-phys.net/16/6595/2016/>, 2016.
- 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
450 concurrent meteorological analysis, *Journal of Geophysical Research: Atmospheres*, 115, D18 219, doi:10.1029/2010JD013948, 2010.
- Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, *Journal of the Atmospheric Sciences*, 34, 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.
- Werbos, P.: Backpropagation through time: what it does and how to do it, *Proceedings of the IEEE*, 78, 1550–1560, doi:10.1109/5.58337, <http://ieeexplore.ieee.org/document/58337/>, 1990.
- 455 Wood, R.: Stratocumulus Clouds, *Monthly Weather Review*, 140, 2373–2423, doi:10.1175/MWR-D-11-00121.1, 2012.