Machine Learning Reveals Strong Grid-Scale Dependence in the Satellite Nd–LWP Relationship

Referee Comments

Point-by-point responses in blue, additions to manuscript are ***bold & italicized***

Dear Referees,

Thank you for your time and effort put into our manuscript. We appreciate that you provided a fair and insightful evaluation of this work and that your comments have led to changes in the manuscript that improved the clarity and accuracy of the analysis. Specifically, we have provided a detailed table of the machine learning model results for all 12 regions, which shows a robust N_d–LWP sensitivity and stable model results across regions. We have also followed the referees' suggestions by moving key figures from the Supplement to the manuscript and adding tables to better describe the dataset sources and aspects of the machine learning model so the reader can more easily access this information. Finally, we have made this integrated multi-satellite database publicly available on DataHub so the analysis can be reproduced, and new science conducted from it. Overall, the narrative has not changed, but we believe that with these changes the conclusions are now stronger.

Best regards,
Matt

Reviewer #1

General comments:

In this study, the authors investigate the Nd–LWP relationship (Nd: cloud droplet number concentration; LWP: liquid water path) retrieved from satellite observations at grid resolutions ranging from 10° to 0.05°. To reduce retrieval errors, they introduce a machine learning (ML) random forest model to estimate LWP using relevant cloud-controlling factors. After obtaining reliable ML results, the authors re-examine the Nd–LWP relationship and identify the main controlling factors that determine its characteristic shapes. They further apply this method to evaluate radiative forcing.

The reviewer is impressed by the methodology developed in this work, particularly the application of ML techniques to decompose the dominant controlling factors shaping the Nd–LWP relationship. The authors test their approach comprehensively across multiple grid resolutions (10° to 0.05°) and 12 different oceanic regions, successfully identifying general characteristics of the Nd–LWP relationship and its impact on radiative forcing. The reviewer finds this study innovative and believes it highlights a promising research direction for analyzing high-resolution satellite data. Therefore, the reviewer recommends publication of the paper, subject to minor revisions.

Many supplemental figures are shown in a separate file. In principle, the text should be readable without referring to the supplemental material. In this sense, it is better to place Figures S2 and S10 in the main text. Please reconsider the choice of the figures in the main text and those of the supplemental material.

>> We agree that the text should be understandable without needing to refer to the supplementary material and have therefore moved Figures S2 and S10 to the main text, as you suggested, since they are referenced multiple times throughout the manuscript.

Specific comments:

L75–76: Please explain the names of the filters, Q06 and G18. Do they refer to specific papers?

>> Thank you for raising this point. The terminology describing the filter names is the same as that used in Gryspeerdt et al. (2022). We have clarified the naming conventions explicitly in the text as:

- Q06: Includes all filters from the All composite plus $\tau_c > 4$ and $R_e > 4$ µm. *This filter is called Q06 because it uses the same set of constraints as those used in Quaas (2006).*
- G18: Includes all properties from the Q06 composite plus 5-km CF > 0.9, solar zenith angle (θ_{solar}) < 65°, satellite zenith angle ($\theta_{satellite}$) < 55°, and sunglint pixel index (SPI) < 30°. *This filter is called G18 because it uses the same set of constraints as those used in Grosvenor et al. (2018).*

References

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z. B., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409–453, https://doi.org/10.1029/2017rg000593, 2018.

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

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

L140–141: CEP is located in the Eastern Pacific. The naming of this oceanic region could be improved. Why was no region in the Western Pacific selected? For example, the Western Pacific near the equator at around 160°E. It is a typical convective area.

>> We agree that the naming for CEP could be improved and have therefore added the word "East" to the acronym CEP, so it now stands for "Central East Pacific," to avoid confusion. We did not include a CWP ("Central West Pacific") region—located near Indonesia—because this area is dominated by deep convective clouds, leaving few reliable satellite-retrieved samples of low-level boundary layer clouds suitable for analysis.

L149, "Precipitation rates are also large in the tropics (compared to the subtropics)." According to Fig. 1e, the precipitation rate is larger in ST compared to TR.

>> We are not sure where the confusion arises regarding this referee comment. According to Fig. 1e, the precipitation rate is larger in the TR region compared to ST (ST: 0.03 ± 0.17; TR: 0.10 ± 0.48). Thus, precipitation in the TR regions is indeed higher. It is possible that the reviewer may have mistakenly referred to panel d instead of panel e for the precipitation rate. Therefore, the text remains unmodified.

L237, "precipitation can also decrease LWP": What type of case is considered when precipitation leads to a decrease in LWP?

>> We have further clarified this point in the text (and have also moved content related to precipitation to section 5.1 for increased readability).

Precipitation (probability and intensity) and N_d are closely associated with LWP, typically increasing as LWP increases in warm clouds. While LWP and precipitation generally increase together as clouds deepen, in more developed or heavily drizzling systems, efficient rainout *processes can deplete cloud liquid water, leading to a reduction in LWP and a bidirectional response in dLWP/d*N_d (e.g., in CloudSat observations of L'Ecuyer et al. 2008, Lebsock et al. 2008 and Chen et al., 2014).

L304, Section 5.3: According to Fig. 6, relative humidity above PBL is not an important factor determining LWP. The discussion in Section 5.3 seems redundant.
>> We appreciate the reviewer's comment. However, this finding—that free-tropospheric relative humidity is *not* an important factor controlling LWP—is precisely the key point we wish to emphasize. Previous studies (e.g., Ackerman et al., 2004; Chen et al., 2014; Gryspeerdt et al., 2019) have highlighted above-PBL humidity as a major driver of the negative LWP–N_n adjustment, yet our analysis shows this influence to be weak or even opposite in sign under most conditions. Demonstrating the lack of a strong RH control, despite its presumed importance, is a meaningful and novel result. To make this clearer, we have revised Section 5.3 to explicitly highlight this interpretation: *This analysis is included to explicitly demonstrate that, contrary to prior expectations, free-tropospheric humidity exerts only a weak influence on LWP.*

Section 5.4 is understandable, but it can be improved in terms of readability. Please relate each cloud and radiative effect listed in Table 2 to the mathematical expression in the text.
>> Thank you for raising this point. We now describe and bolden each term—the **Twomey**, **liquid water path**, and **cloud fraction** radiative effects—in the text following the presentation of Equation 5. We also describe the radiative scaling term included in Table 2 as follows:

*The three terms in parentheses correspond to the **Twomey, liquid water path**, and **cloud fraction** radiative effects, respectively. These are multiplied by a radiative scaling factor defined as* $(-1) \times CF \times \overline{F^{\downarrow}} \times \frac{dlnN_d}{dlnAI} \times \overline{\Delta lnAI}$ *where the negative sign indicates that an increase in albedo (from higher* N_d*) reduces the net downward shortwave flux.*

As a final note, during the revision of Table 2 we identified a typo in Equation 1, which previously omitted the minus sign. The negative sign indicates that an increase in planetary albedo reduces the net downward (absorbed) shortwave flux, consistent with the convention that a positive ΔF represents a warming (increase in absorbed energy).

L454: What does "the rapid adjustments" refer to here?
>> This is an insightful question. As "rapid" adjustments refer to fast cloud and atmospheric changes occurring before significant surface temperature climate responses manifest, our study does not actually need to use this climate-related term; therefore, the word "rapid" adjustments have been removed from the manuscript.

Reviewer #2
General comments:
In the article “Machine Learning Reveals Strong Grid-Scale Dependence in the Satellite N_d – LWP Relationship”, the authors employ a machine learning model to investigate the relationship between cloud droplet number concentration and liquid water path, and their connection to aerosol-cloud interactions, such as the Twomey effect.
With their random forest model, the authors provide highly interesting results for distinct changes of the N_d-LWP relationship with grid resolution and regional effects. As such, the article provides an efficient and innovative approach to quantify effects of aerosols on cloud processes, offering exciting opportunities for Earth system models.
Overall, the authors present their findings clearly and concisely, allowing readers to easily follow their approach. Hence, I regard this article with its findings on aerosol-cloud interactions and the introduced machine learning approach as a valuable contribution to the scientific community and future research. While I recommend this article for publication, I have some minor comments where additional clarification would be appreciated before publication.

Specific Comments:
L. 51-52: “We have generated a series of collocated global datasets at a series of spatial resolutions from 10° × 10° down to 0.05° × 0.05°”. Could you specify your resolutions in this section? The information can be found in Section 3.1, but it would be helpful to include a list of resolutions here.
>> Good point. We now list all six spatial resolutions up front in our study.

L. 74-77: The naming of the filters (Q06, G18) does not seem intuitive to me. What do Q06 and G18 stand for? Please add either a reference or introduce the acronyms.
>> Thank you for raising this point. The terminology describing the filter names is the same as that used in Gryspeerdt et al. (2022). We have clarified the naming conventions explicitly in the text as:

- Q06: Includes all filters from the All composite plus $\tau_c > 4$ and $R_e > 4$ µm. *This filter is called Q06 because it uses the same set of constraints as those used in Quaas (2006).*
- G18: Includes all properties from the Q06 composite plus 5-km CF > 0.9, solar zenith angle (θ_{solar}) < 65°, satellite zenith angle ($\theta_{satellite}$) < 55°, and sunglint pixel index (SPI) < 30°. *This filter is called G18 because it uses the same set of constraints as those used in Grosvenor et al. (2018).*

References
Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C.,

Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z. B., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409–453, https://doi.org/10.1029/2017rg000593, 2018.

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

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

L. 126: “following 13 predictor variables”: Instead of only naming all variables, you could help the reader by providing an overview table for included predictors and their respective sources.

>> Thank you for this suggestion. We have added a table of each predictor variable and its respective source to the supplement file.

Table S1. Overview of predictor variables and their respective data sources.

Predictor Variable	Source / Dataset
Planetary Boundary Layer Height (PBLH)	Reanalysis data
Lifted Condensation Level (LCL)	Calculated from reanalysis data
Relative Humidity above PBL Height (rhAbovePBL)	Calculated from reanalysis data
Estimated Inversion Strength (EIS)	Calculated from reanalysis data
Surface Temperature Advection (Tadv)	Calculated from reanalysis data
Surface Latent Heat Flux (LH)	Reanalysis data
Total Column Water Vapor (tqv)	Reanalysis data
10-m Surface Wind Speed (ws10)	Reanalysis data
Surface Precipitation	AMSR-E satellite retrieval
Cloud Top Height (CTH)	MODIS satellite retrieval
Cloud Fraction (CF)	MODIS satellite retrieval
Cloud Albedo (A_{cld})	CERES satellite retrieval
Cloud Droplet Number Concentration (N_d)	Calculated from MODIS satellite retrieval

L. 131-132: "Each tree is trained on approximately 60% of the training dataset with replacement, utilizing the remaining 40% as out-of-bag observations to test tree performance". Please describe how you split the dataset (random, temporal, spatial). Did you use the same dataset for validation and test? Ideally, you would have three datasets to ensure evaluating on an independent test set the model has not seen before.

>> We thank the reviewer for noting this important detail. We have clarified the following in the manuscript based on your comment:

The full dataset was randomly partitioned into three independent subsets: 65% for training, 25% for testing, and 10% for validation. Randomized sampling partitions, as opposed to sequential (e.g., yearly) splits, did not have a significant impact on the model's outcomes. Within the training set, each decision tree in the random forest is trained on approximately 60% of the training data (sampled with replacement), while the remaining 40% serves as out-of-bag data for internal performance evaluation. The validation set was used for tuning model hyperparameters described in Table S2. After hyperparameter tuning, the model was retrained using both the training and validation data (75% total) to optimize performance, ensuring that the test set remained unseen and provided an unbiased assessment of model accuracy.

L. 133-134:" evaluated using different hyperparameter values, such as the number of trees and the minimum number of samples per leaf". It would be great to have an overview table for all hyperparameters.

>> A new table of the hyperparameters and their description used in the ML model have been added to the supplement file for the manuscript.

Table S2. Overview of random forest hyperparameters and associated values used in this study.

Hyperparameter	Value	Description
Number of trees	100	Total number of decision trees in the random forest ensemble. A larger number improves stability and accuracy but increases computational cost.
Minimum leaf size	7	Minimum number of samples required to form a terminal leaf node. Controls model complexity; smaller values allow deeper trees that may capture finer variability but risk overfitting.
Sample fraction	0.6	Fraction of the training data randomly sampled (with replacement) to train each tree, defining the bootstrap sample size and influencing model diversity.

L. 217: I found it a bit difficult to follow the results in this section. They mostly relate to different regional characteristics (i.e., California), but I miss a more general evaluation of the Nd-LWP relationship and a comparison across regions. Do you have the same number of samples in all regions? If not, it would be interesting to compare the findings between regions in connection to their robustness.

>> We appreciate your comment and agree that the manuscript would benefit from including an ML comparison across all 12 regions. To better evaluate the robustness of the ML-derived N_d-LWP relationship across regions, we have added a new summary table (Table S3) to the Supplement (also shown below).

Table S3. Random forest predictions of LWP for each region using the 0.1°-resolution dataset. Shown are the number of samples (N$samples$), the linear least-squares fit of dLWP/dNd for non-raining and raining conditions, the Pearson correlation coefficient (R^2), the mean percentage error (MPE), and the top three predictor variables ranked by importance (from highest to lowest).

Region	$N_{samples}$	dLWP/dlnN_d (non-raining)	dLWP/dlnN_d (raining)	R^2	MPE (%)	Importance Order
CAL	2.26e+07	-0.003	0.24	0.75	22.0	Pr,A_{cld},N_d
PER	2.32e+07	-0.06	0.33	0.71	26.0	Pr,A_{cld},CTH
NAM	2.34e+07	0.04	0.24	0.74	20.7	Pr,A_{cld},CTH
AUS	2.03e+07	-0.08	0.29	0.74	27.7	Pr,CTH,N_d
CEA	1.04e+07	-0.02	0.29	0.71	31.2	Pr,CTH,N_d
WEI	5.19e+06	-0.06	0.15	0.68	32.5	Pr,CTH,N_d
CEP	1.32e+07	-0.03	0.16	0.75	29.5	Pr,CTH,N_d
ENA	1.42e+07	0.08	0.31	0.73	32.7	Pr,A_{cld},CTH
WNP	9.64e+06	0.24	0.33	0.78	26.0	Pr,A_{cld},C_f
CNP	1.16e+07	0.12	0.30	0.77	25.2	Pr,A_{cld},TQV
ESA	6.77e+06	0.07	0.29	0.80	23.3	Pr,A_{cld},C_f
ESI	1.02e+07	0.10	0.34	0.75	29.5	Pr,A_{cld},CTH

The following has been added to the main manuscript in section 5

Section 5: *The model also shows robust and stable performance in terms of r^2, mean percentage error, and the ranking of variable importance across all 12 regions in our study (Table S3).*

Section 5.1: *The random forest model performs consistently well across all 12 regions, exhibiting a similar pattern of small positive or negative dLWP/dN_d sensitivities for non-precipitating clouds and larger positive sensitivities for raining clouds (Table S3).*

Figure 1: Add ENA and WNP in the figure caption.
>> Done

Figure 7: Is this also for California, or averaged over all regions?
>> The caption has been clarified that this is for the California region.