Disentangling the Impacts of Saharan Dust Particles on Cloud Characteristics Through Spaceborne Observations and Machine-Learning

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Azusa Takeishi (azusa.takeishi@aero.obs-mip.fr),
Chien Wang, Eric Defer, Christelle Barthe, Pierre Tulet, Vincent Noel

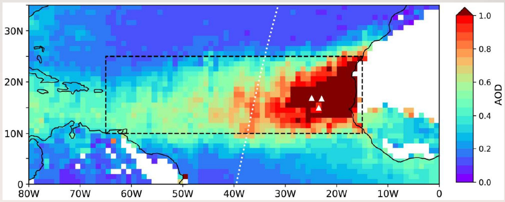
Laboratoire d'Aérologie, CNRS/UPS, Toulouse

cnes CENTRE NATIONAL D'ÉTUDES SPATIALES

Laero Laboratoire d'Aérologie

Introduction

Dust particles are known to have the ability to nucleate ice crystals (Cziczo et al., 2013). In mixed-phase clouds (-38℃ ≤ T ≤ 0℃), the existence and abundance of such ice-nucleating particles (INPs) are critical for determining the bulk properties; in general, glaciated clouds are optically thinner than clouds that consist of many small liquid droplets. It may also have some impacts on the timing and the amount of surface rainfall if clouds are precipitating. Even though several laboratory experiments and field campaigns have taken place to understand this dust-cloud relationship, it has been challenging to parameterize and incorporate the effects in numerical models (Kanji et al, 2017).

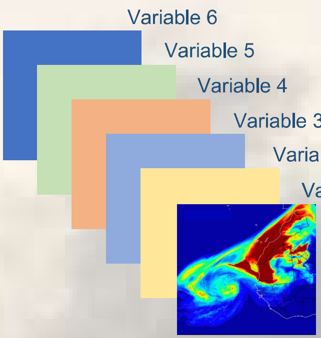


▲Monthly mean aerosol optical depth over the North Tropical Atlantic in June 2020

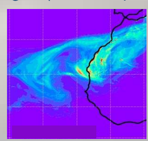
►Visibility lowered by a Saharan dust flow over Cape Verde, 2021

Saharan dust particles are consistently blown westwards by winds while interacting with radiation (the direct effect) and clouds (the indirect effect). This study addresses the following scientific question:
"***Can we obtain a quantitative relationship between Saharan dust abundance and cloud properties downwind by machine-learning the past observational data?***"

Methods



Input features + Target (= cloud property) × 20 years

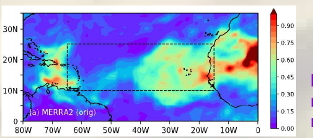


A machine is trained to learn the target values (i.e., a cloud property) as a complex function of input features (i.e., background dynamical & microphysical variables).

The target variables are **(1) cloud coverage [%], (2) cloud optical depth, (3) cloud phase (liquid/ice), and (4) cloud top height [km].** One machine is trained for each target variable with the same input features.

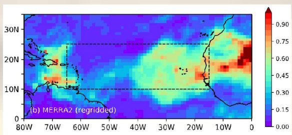
Data Preparation

Variables from different datasets need to be pre-processed so that (1) the resolution of the data gets unified to 1° × 1° (e.g., upper figures below) and (2) the area of missing data gets minimized (e.g., lower figures below).

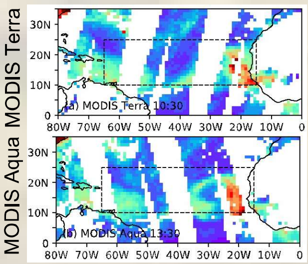


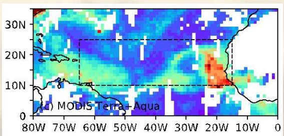
Original MERRA2 data (~50km resolution)

→ Nearest neighbor method



Processed MERRA2 data (1° resolution)

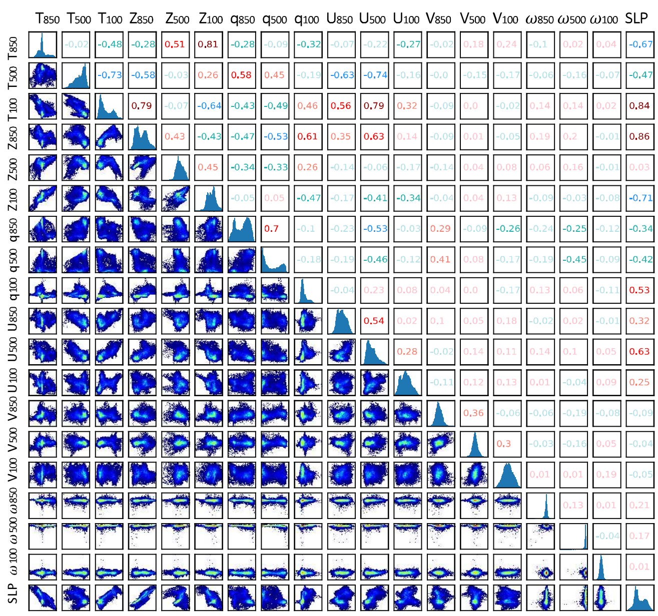




Terra + Aqua merged

*Although there is a 3h time difference, the area of missing data got minimized

Preliminary Results / Ongoing Work



▲ An example of the correlation analysis using the data from MERRA2 on June 1, 2000.

Minimizing the number of input features is a crucial step in machine-learning, as a resultant trained machine would require fewer variables for predictions (e.g., computationally less expensive, easier to be implemented in a model, etc.). In the example above, we can see strong relation between SLP and Z_{850mb}, for instance, and therefore one of them can be eliminated from input features. This analysis is being conducted on the 20-year data for June and July.

References

Cziczo, D. J., Froyd, K. D., Hoose, C., Jensen, E. J., Diao, M., Zondlo, M. A., Smith, J. B., Twohy, C. H., and Murphy, D. M. (2013). Clarifying the Dominant Sources and Mechanisms of Cirrus Cloud Formation, *Science*, 340(6138), 1320-1324, doi: 10.1126/science.1234145

Kanji, Z. A., Ladino, L. A., Wex, H., Boose, Y., Burkert-Kohn, M., Cziczo, D. J., and Krämer, M. (2017). Overview of Ice Nucleating Particles, *Meteorological Monographs*, 58, 1.1-1.33, doi: 10.1175/AMSMONOGRAPHS-D-16-0006.1