Mapping between shallow cumulus cloud field properties and three-dimensional surface solar irradiance

Jake J. Gristey (Jake.J.Gristey@noaa.gov) [1,2] | Graham Feingold[2] | Ian B. Glenn[3] | K. Sebastian Schmidt[4] | Hong Chen[4]

[1]Cooperative Institute for Research in Environmental Sciences [2]NOAA Chemical Sciences Laboratory [3]NASA Jet Propulsion Laboratory [4]Laboratory for Atmospheric and Space Physics

Check out my CIRES profile:

Motivation & Aim

- Ubiquitous shallow cumulus clouds exhibit highly 3D structure leading to complex variability in surface solar irradiance (SSI).
- We aim to use machine learning to establish direct relationships between the cloud properties and SSI beneath.
- Such direct relationships are of great scientific interest to understand the drivers of SSI variability, as well as of practical importance for efficient and accurate SSI predictions.

Cloud and Radiation Data

- Well resolved shallow cumulus cloud fields are generated from large eddy simulation (LES) and realistic associated SSI is calculated with Monte-Carlo 3D radiative transfer.

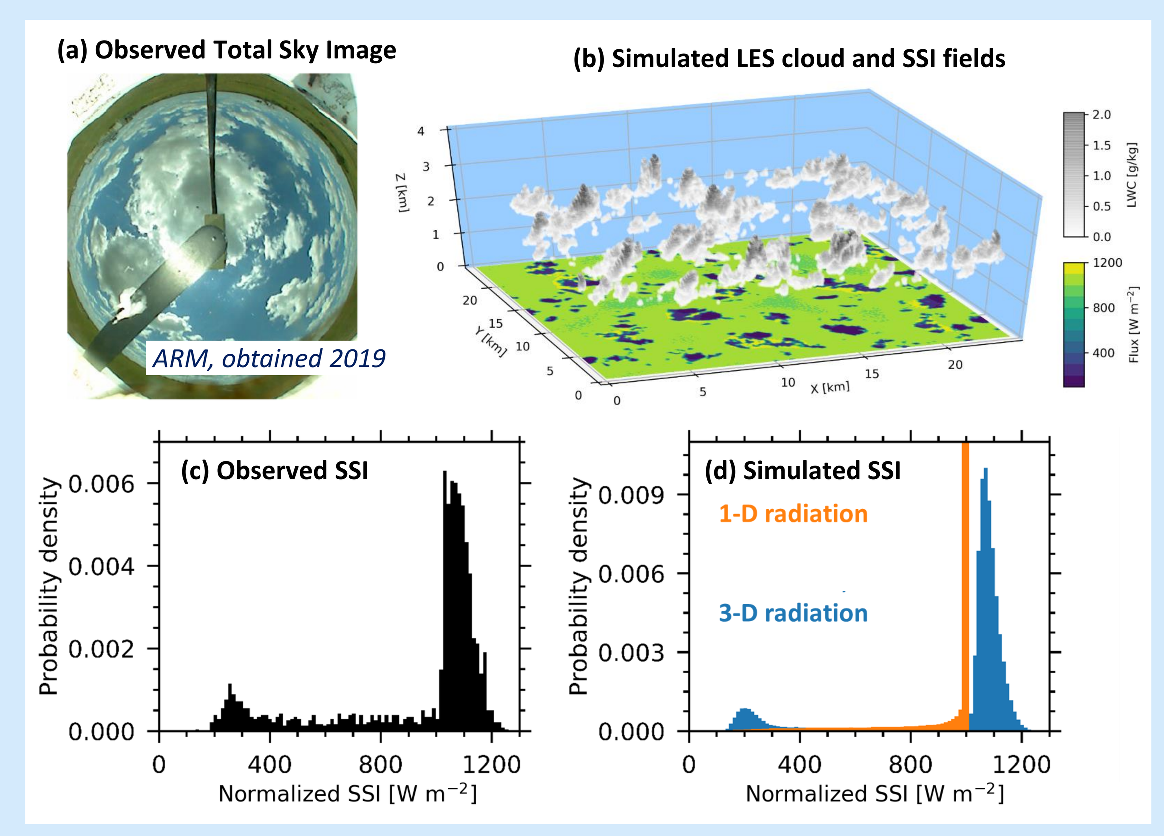

Fig. 1. Shallow cumulus clouds and associated SSI PDFs on the afternoon of 27 June 2015 in Oklahoma.

- The observed shape of the SSI probability density function (PDF) is only reproduced with 3D radiative transfer (Fig. 1).

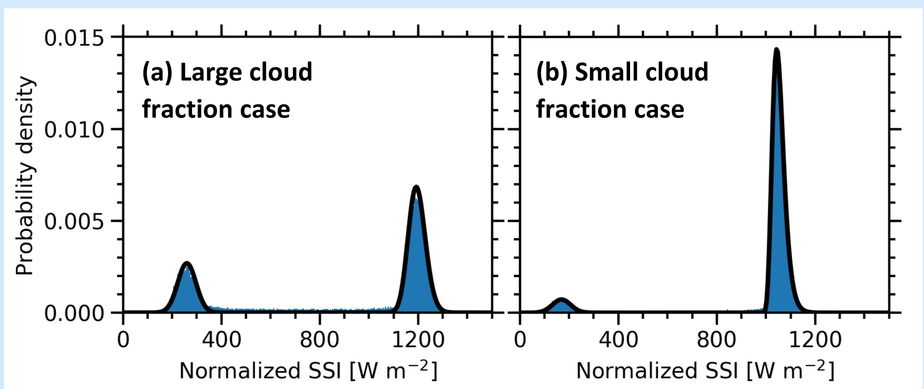

Fig. 2. Two representative examples of simulated SSI PDFs with fitted normal and lognormal distributions.

- Normal and lognormal distributions are fitted to the small and large irradiance modes of the 3D SSI PDFs, respectively (Fig. 2).
- 531 cloud fields and fitted SSI PDFs are selected.

Machine Learning Approaches

- Seek relationships between 6 cloud field properties and 7 SSI PDF fit parameters (Table 1).

Table 1. Machine learning inputs and outputs.

Cloud field properties (inputs)	SSI PDF fit parameters (outputs)
Mean cloud fraction: $\overline{f_C}$	Lognormal location parameter: θ
Dispersion in liquid water path: $D(LWP)$	Lognormal shape parameter: s
Mean drop number concentration: $\overline{N_C}$	Lognormal scale parameter: m
Mean projected cloud area: $\overline{A_C}$	Normal location parameter: μ
Mean distance to nearest cloud: $\overline{D_{C-NN}}$	Normal shape parameter: σ
Cosine of solar zenith angle: $cos(SZA)$	Weight of small irradiance mode: w_1
	Weight of large irradiance mode: w_2

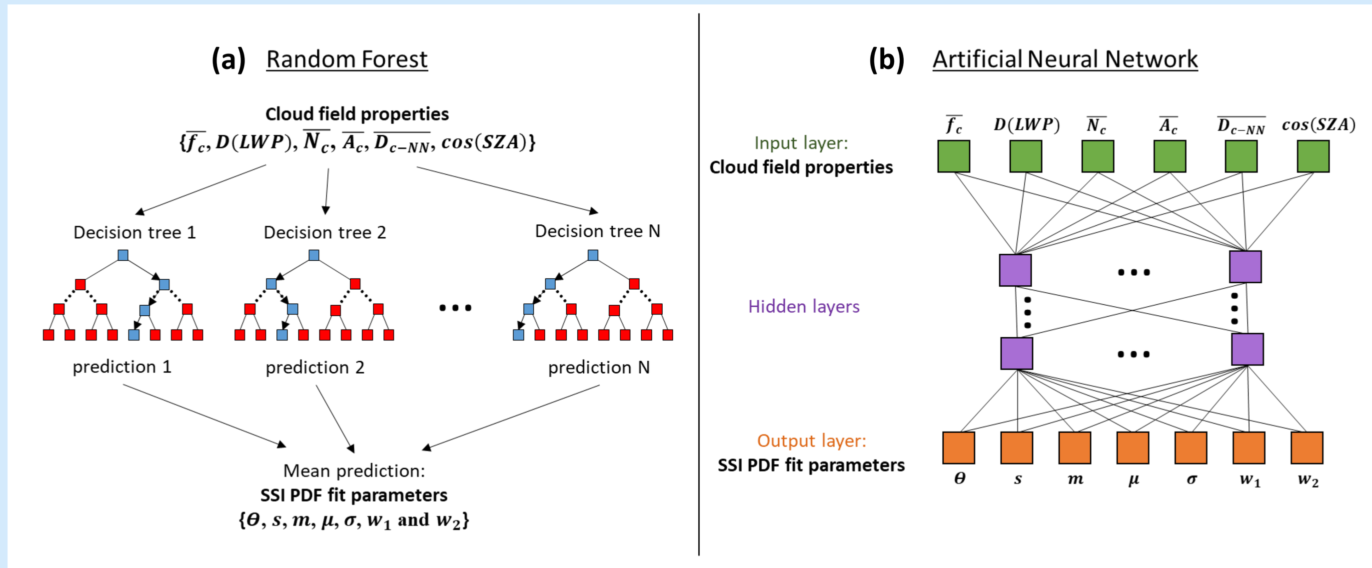

Fig. 3. Machine learning architectures employed.

- Two fundamentally different algorithms, a random forest (RF) and an artificial neural network (ANN), are used to build the relationships

Key Findings

1. Complex 3D surface solar irradiance simulated beneath continental shallow cumulus clouds is well predicted with just a handful of key cloud field properties using machine learning algorithms.
2. Two fundamentally different algorithms perform indistinguishably well, and both value cloud fraction and liquid water as the most important inputs although detailed cloud properties also matter.
3. Results have immediate use for surface energy assessments, with several other promising applications.

Prediction of Surface Solar Irradiance

- Predicted PDF fit parameters by the trained RF and ANN algorithms are used to reconstruct the SSI PDFs (Fig. 4).
- Predictions capture variations in the shape and size of both modes. For reference, note that 1D radiative transfer (Fig. 1d) does not capture even the bimodal shape, let alone the detailed variations.
- Both algorithms perform indistinguishably well, despite the different architectures (Fig. 3).

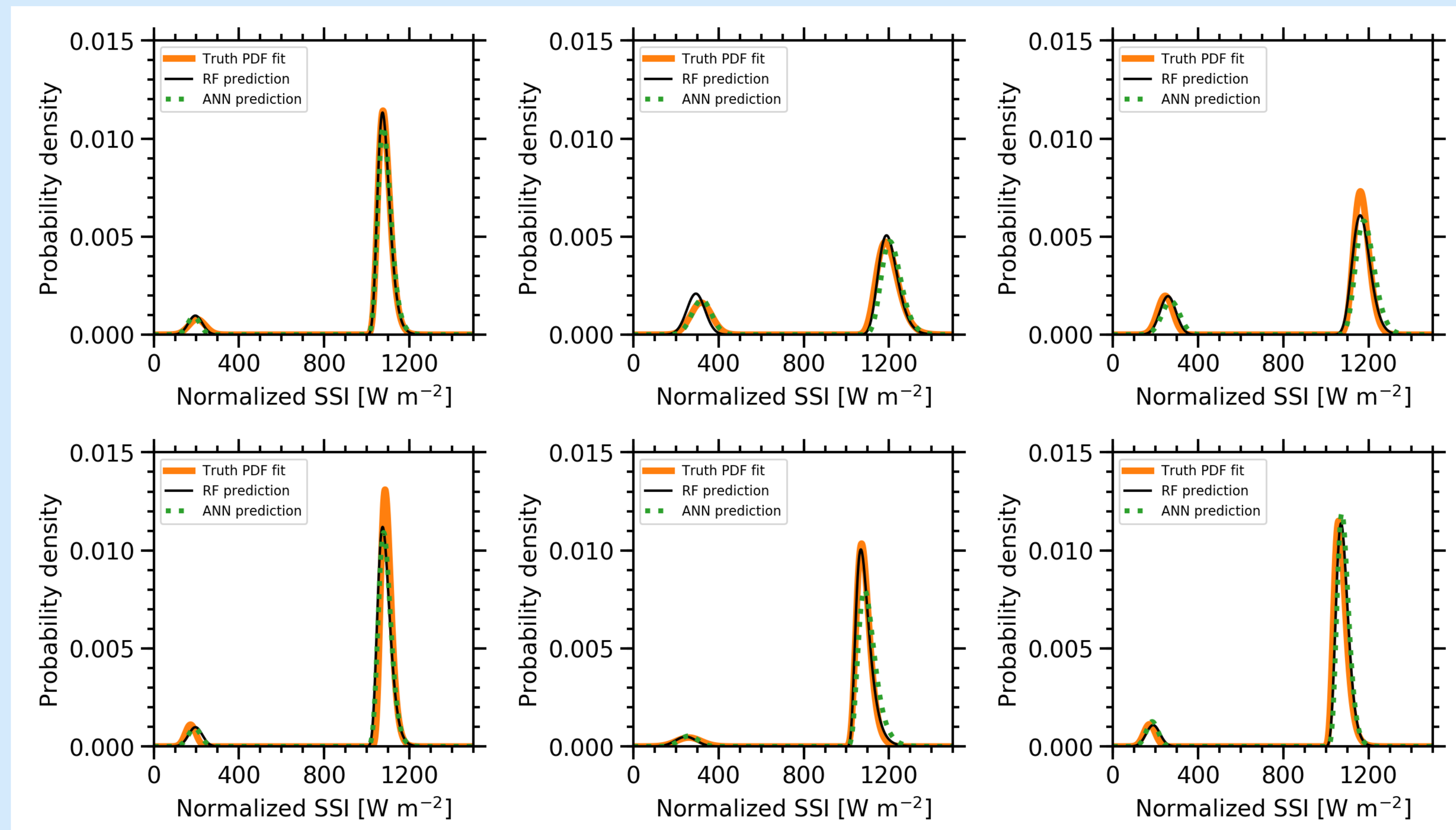

Fig. 4. Six random predictions of SSI PDFs on held-out test data.

Relative Importance of Cloud Properties

- The relative importance of the various input cloud field properties in arriving at the trained RF and ANN algorithms is quantified by impurity-based and permutation-based importance metrics (Table 2).

Table 2. Relative importance of the input cloud field properties

Input	RF impurity importance (%)	RF permutation importance (%)	ANN permutation-importance (%)
$\overline{f_C}$	63.0	74.2 ± 4.1	68.0 ± 4.0
$D(LWP)$	11.3	11.8 ± 0.7	9.6 ± 0.7
$\overline{N_C}$	7.5	4.5 ± 0.4	4.7 ± 0.4
$\overline{A_C}$	4.4	2.2 ± 0.2	9.1 ± 0.6
$\overline{D_{C-NN}}$	6.6	3.1 ± 0.3	3.2 ± 0.3
$cos(SZA)$	7.3	4.3 ± 0.3	5.3 ± 0.5

- For the RF, the impurity-based and permutation-based importance both indicate that $\overline{f_C}$ is the most valuable input followed by $D(LWP)$.
- The permutation based importance is generally consistent for the ANN, although it appears to make more use of $\overline{A_C}$ at the expense of $\overline{f_C}$.
- Even the lesser important inputs appear to provide predictive value. This demonstrates the complexity of the 3D SSI PDF; detailed cloud properties matter.

Applications and future work

- These results have direct application to modeling assessments of solar renewable energy, providing an accurate frequently distribution of SSI, and when combined with photovoltaic cell characteristics, a reliable estimate of energy output.
- One potential application is the parameterization of 3D radiative effects in LES, but requires future work to consider the spatial re-distribution of the PDF.
- Another potential application is bias correction of 3D radiative effects in weather prediction models, but also requires future work to find appropriate substitutes for the inputs because the detailed cloud properties such as $\overline{A_C}$ and $\overline{D_{C-NN}}$ are unlikely to be available.