Approximating downward short-wave radiation flux using all-sky optical imagery using machine learning trained on DASIO dataset

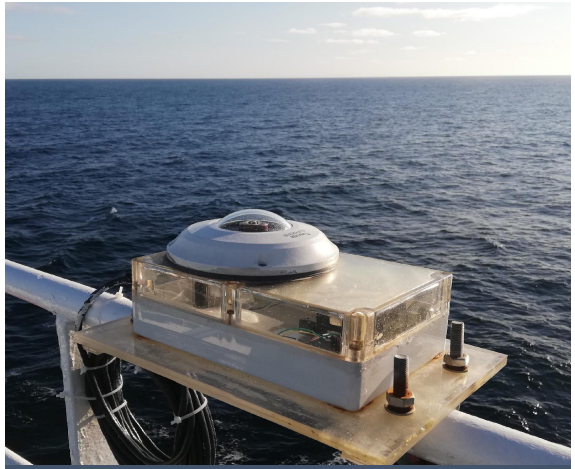
(display material)

Vasilisa Koshkina, Mikhail Krinitskiy, Nikita Anikin, Mikhail Borisov, Natalia Stepanova, and Alexander Osadchiev

Equipment

radiometer
KIPP&ZONEN CNR-1

more than 7’000 €



cloud-camera
«SAIL CLOUD V.2»

less than 1’500 €

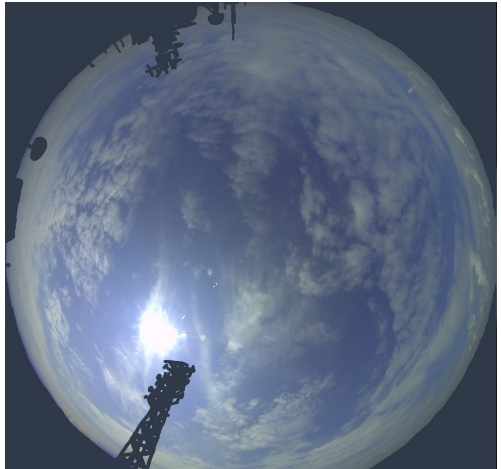
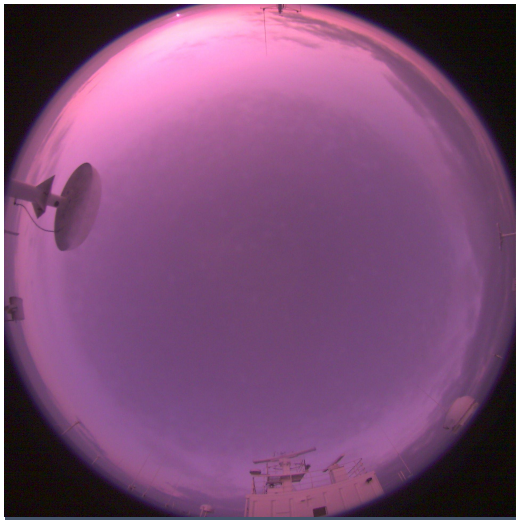


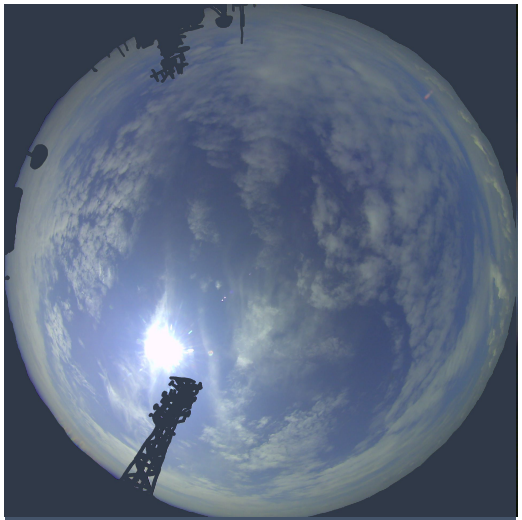
photo with mask

Pre-processing



"Bad" photo at <5 W/m^2

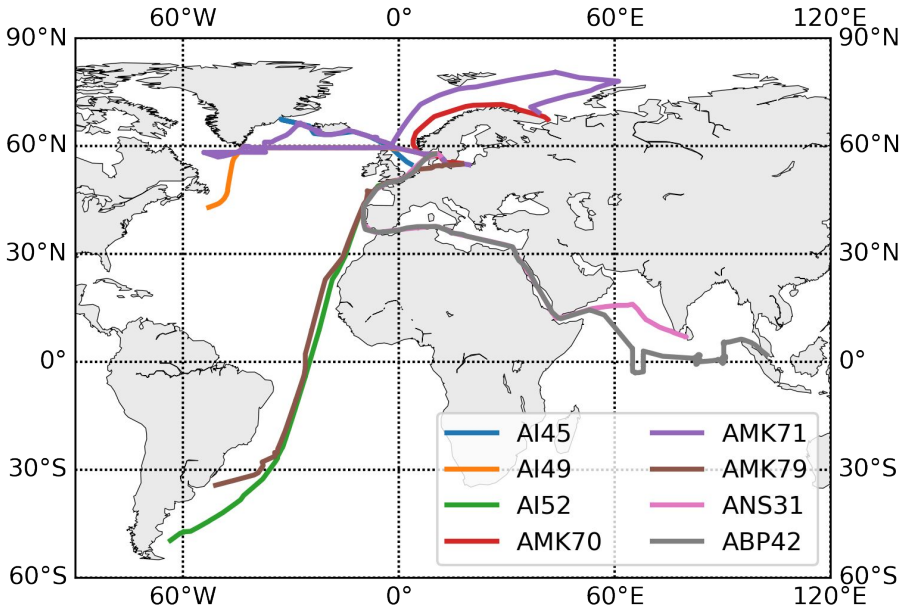
Mask



"Good" photo with mask

Data

- over 2'000'000 all-sky photo from ocean surface:
  The Dataset of All-Sky Imagery over the Ocean (DASIO)[1]
- resolution 1920*1920
- auto white balance
- auto brightness adjustment
- rare images considered outliers
- RGB-channels



Map of cruises of the research vessels

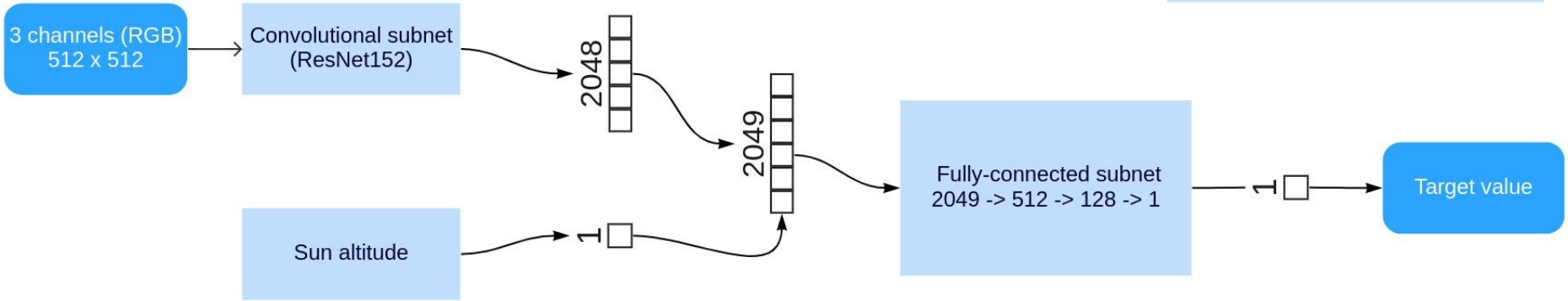
[1]Krinitskiy et al. "On the Generalization Ability of Data-Driven Models in the Problem of Total Cloud Cover Retrieval." Remote Sensing. 2021; 13(2):326.

Methods

- **Classic ML models** (applied to pre-processed data):
  - linear models: *Linear Regression*
  - ensemble models: *Random Forest* and *Gradient Boosting*

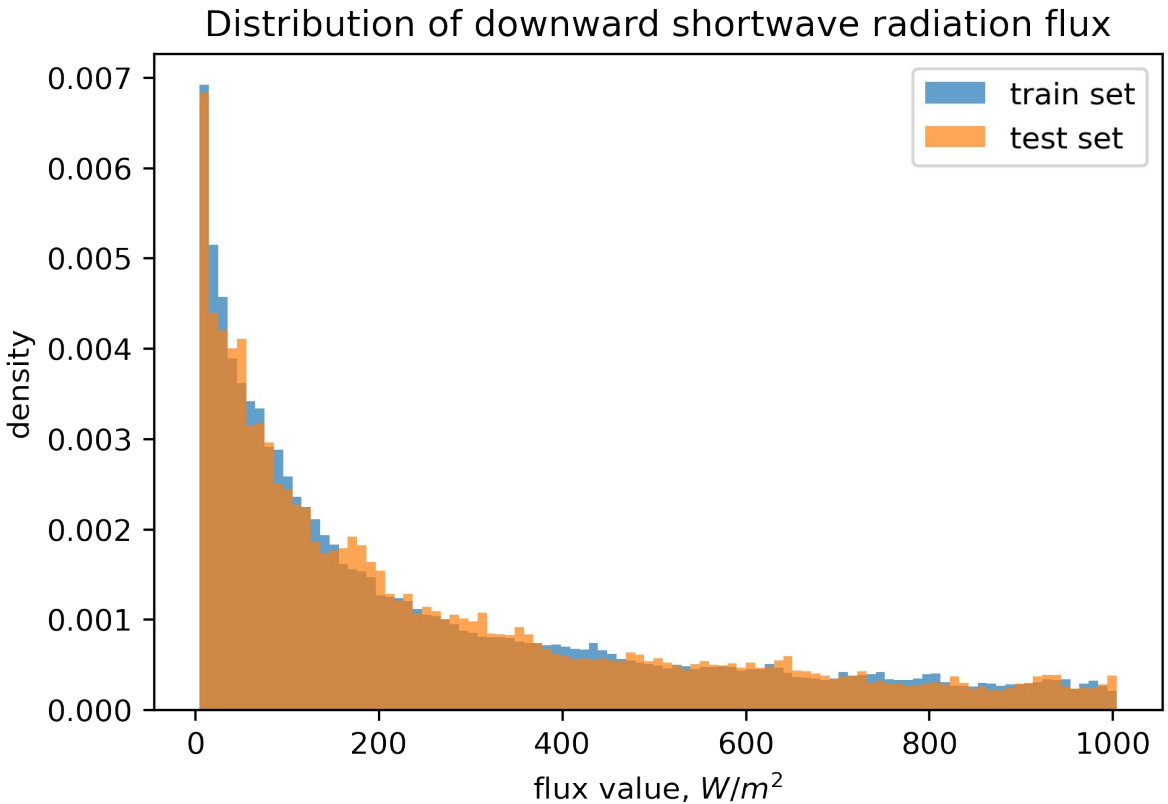
- **End-to-end ML approach** (applied directly to photo):
  - *Convolutional Neural Network (CNN)* with heavy images augmentation (brightness, gaussian noise)

163 features for classic approach
max, min values
mean, variance
skewness, kurtosis
percentile set
sun altitude

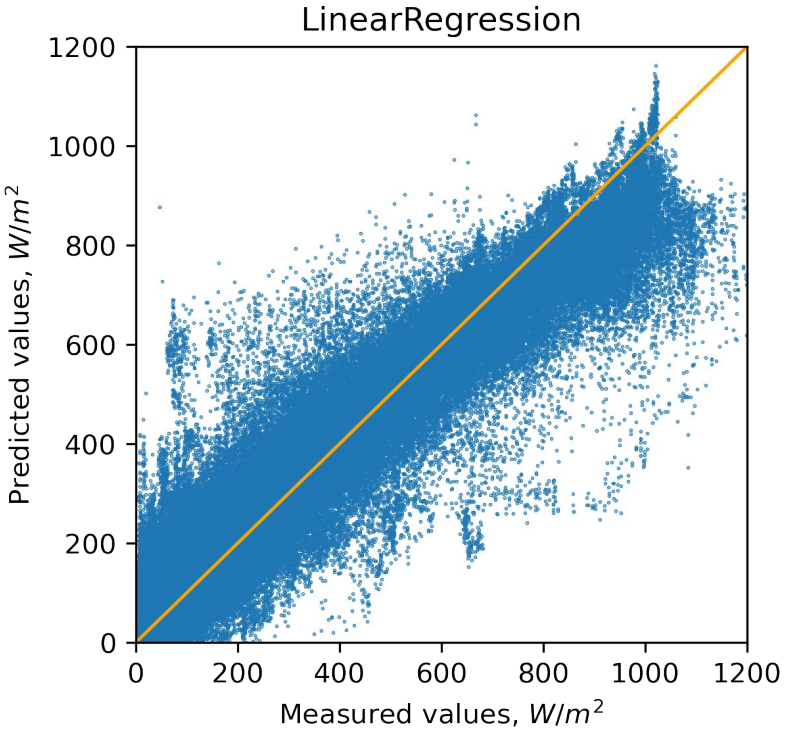


Structure of CNN

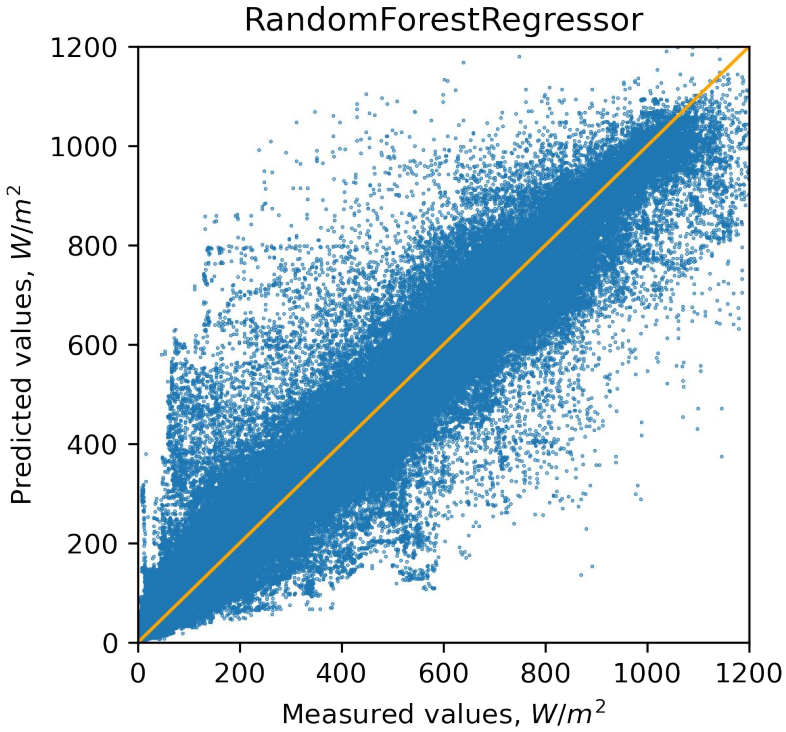
Target values distribution



Value mapping diagrams

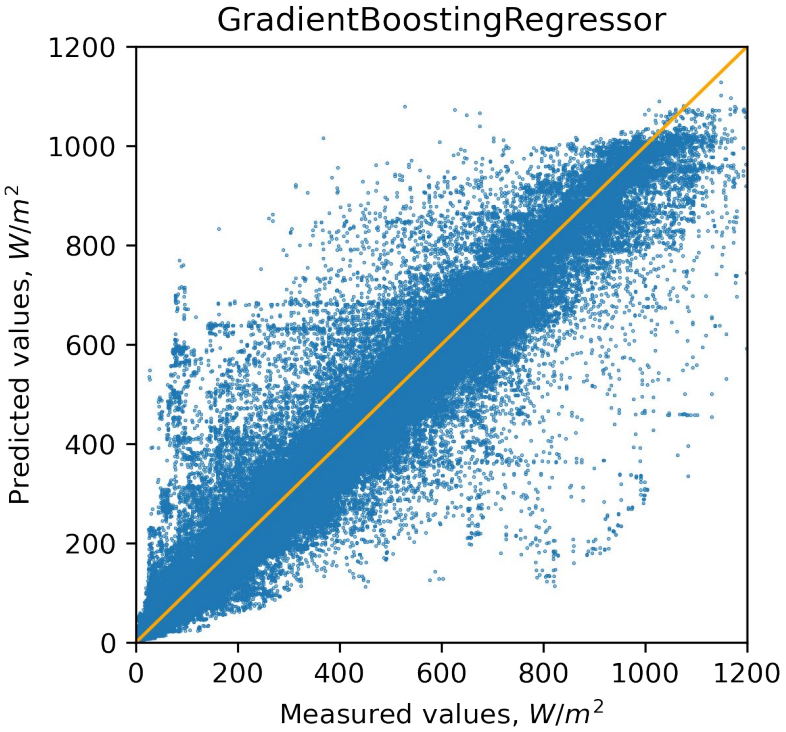


RMSE = 84 W/m²

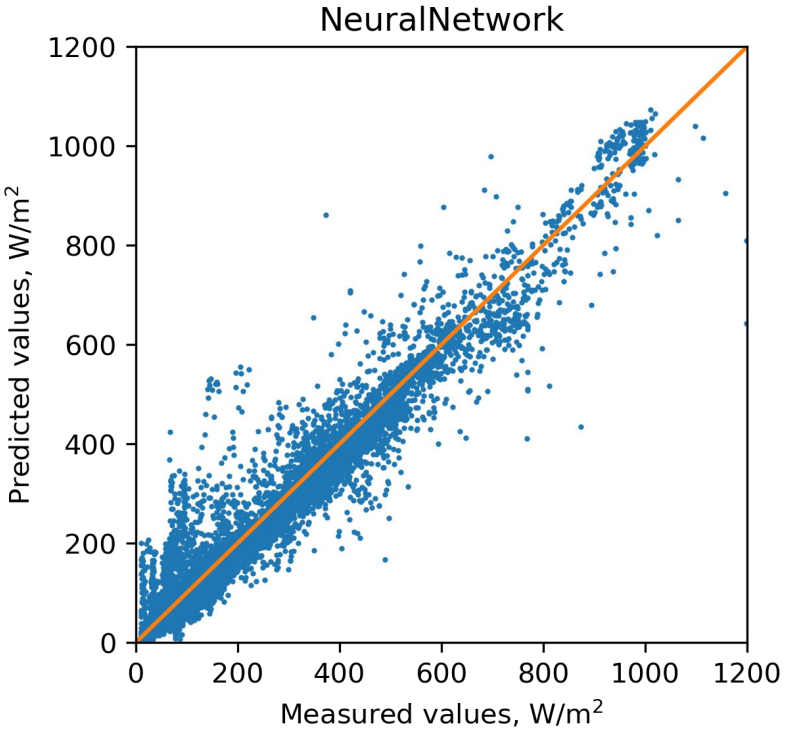


RMSE = 62.1 W/m²

Value mapping diagrams

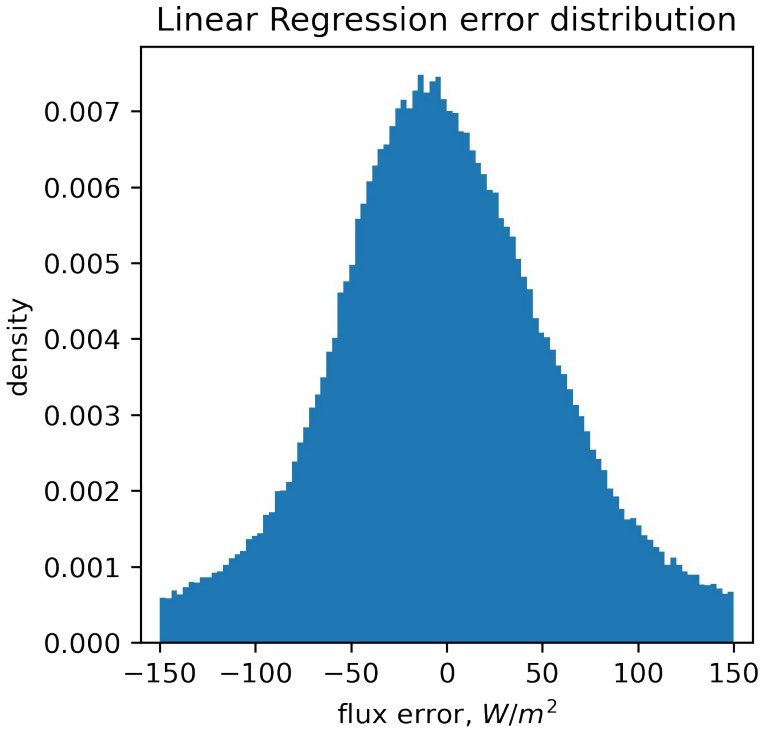


RMSE = 53.5 W/m^2

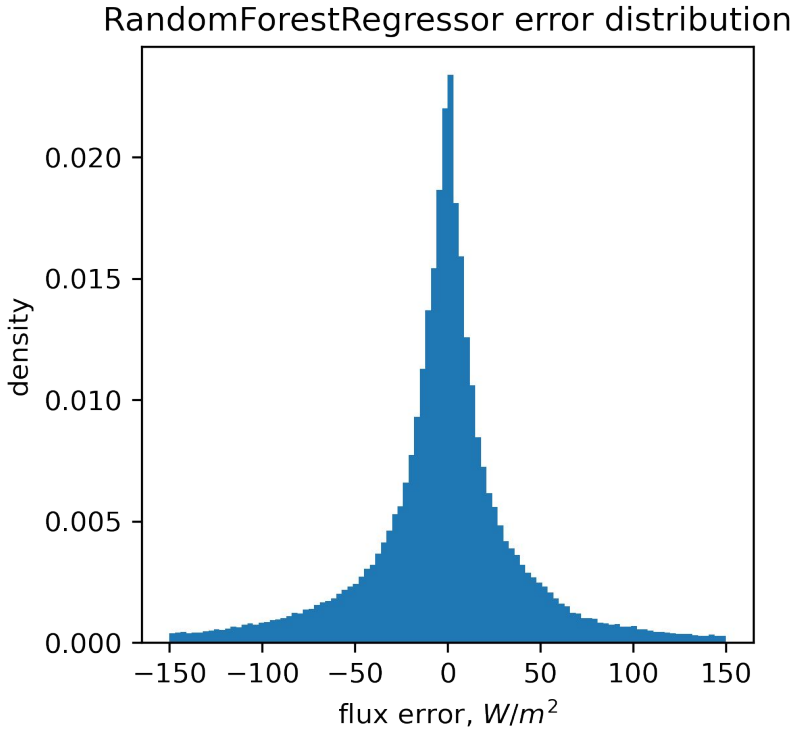


RMSE = 39.2 W/m^2

Error distribution

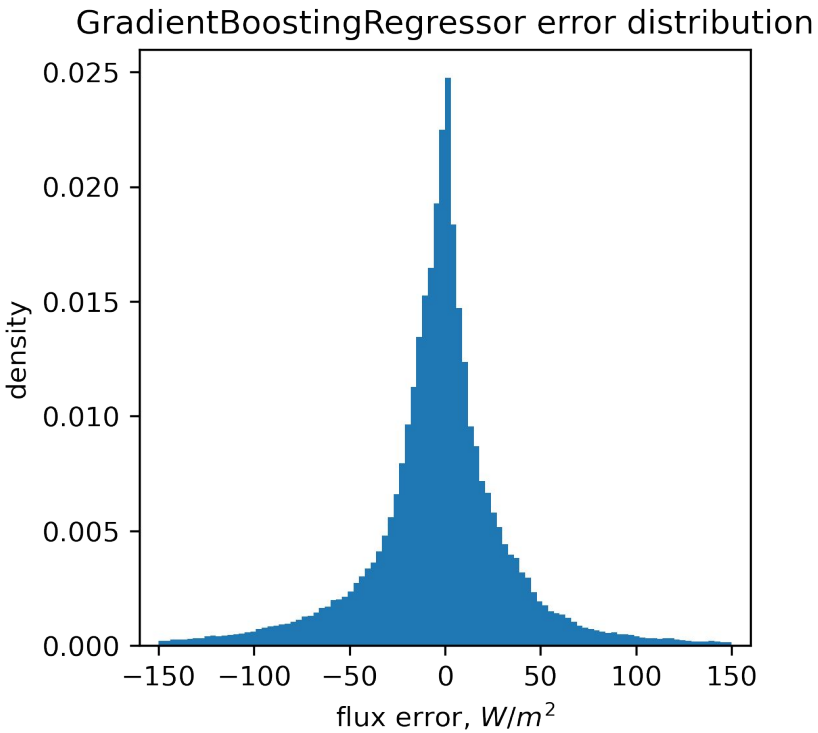


RMSE = 84 W/m^2

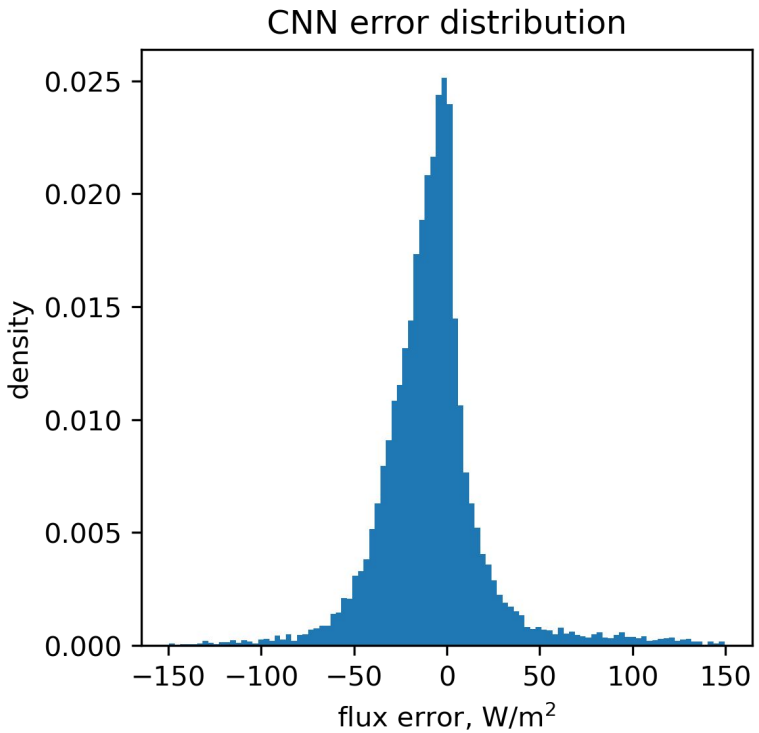


RMSE = 62.1 W/m^2

Error distribution



RMSE = 53.5 W/m^2



RMSE = 39.2 W/m^2

Error distribution