Effect of Satellite Imagery and Machine Learning In An Open-Source Solar Power Forecasting Project

Jacob Bieker

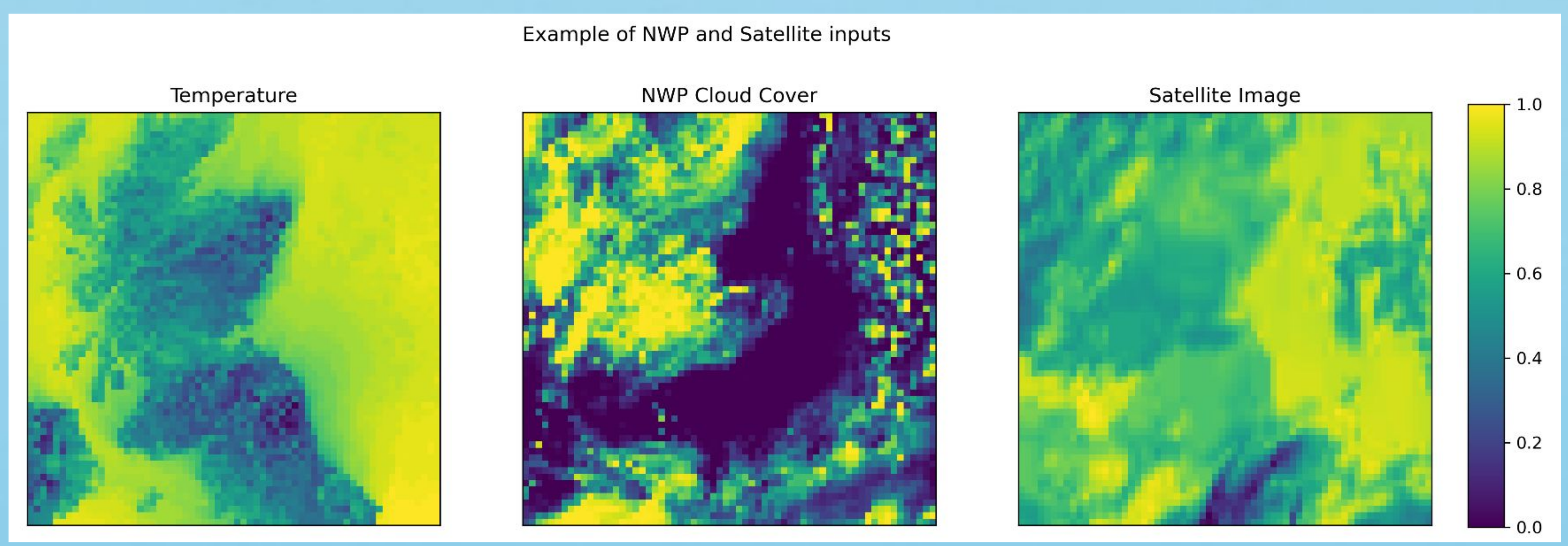
OPEN CLIMATE FIX

Abstract

Current solar PV generation forecasting primarily relies on Numerical Weather Predictions (NWPs) to predict PV generation hours and days into the future. In more recent years, satellite imagery has become more widely available and can provide a more up-to-date view of cloud movement than NWPs. In this work, we examine and quantify the impact of using recent satellite data for forecasting of PV generation in the UK from short to medium term horizons with a variety of machine learning models. We find that including satellite imagery improves short term solar forecast mean absolute error, but has a limited effect further into the future. We also detail and release a multi-decade archive of satellite imagery over Europe.

Methodology

We use EUMETSAT Rapid Scan Service 5-minutely satellite imagery, MetOffice hourly Numerical Weather Predictions, and live PV panel generation readings as inputs into machine learning models that then predict future PV generation. The models tested here include a single-branch 3D-CNN model [1] designed for short-term (2 hour) site-level forecasts using only NWP and satellite imagery with no PV history, an adapted form of MetNet-1 [2,3], which internally uses ConvLSTM and axial attention for up to 48 hour site-level forecasts, and two similar two-branched CNN-based models, called PVNet 1 and 2 [4], for 8-hour regional forecasts. Each model is trained on the same training and test sets for each experiment, but with certain modalities removed depending on the configuration.

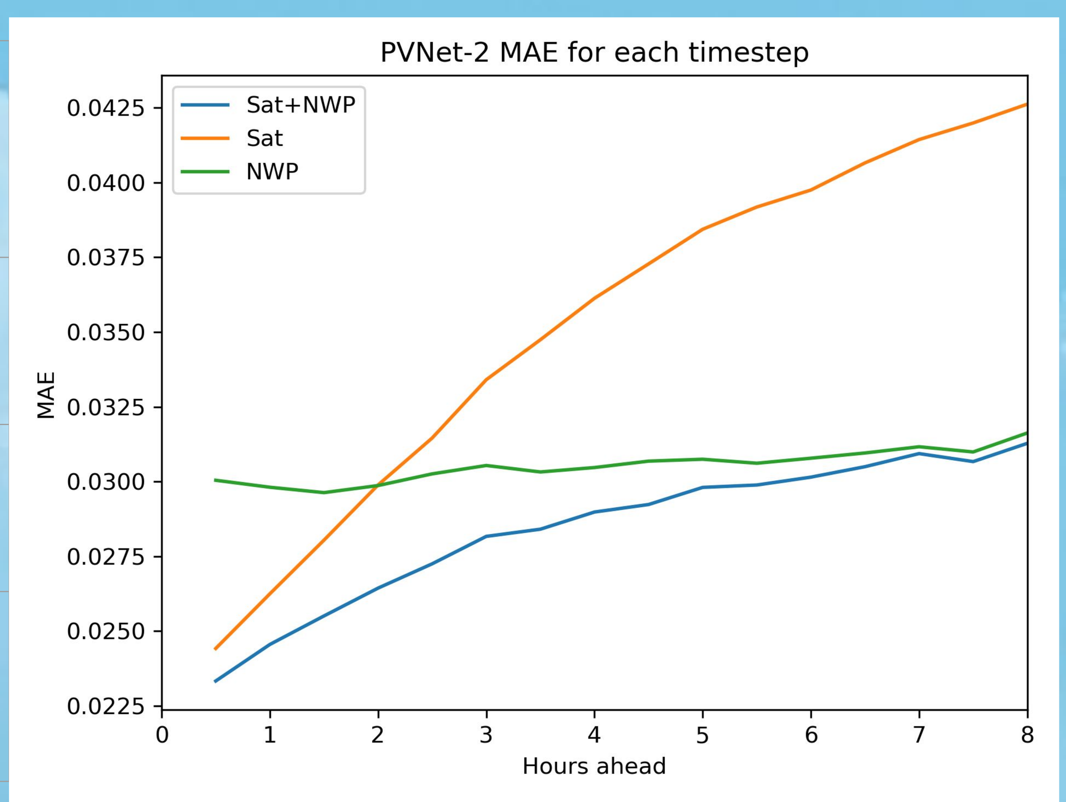


Results

We find that satellite imagery does improve forecasting accuracy on short time scales, although provides very little improvement when forecasting further into the future. In all cases, the combination of NWP and satellite imagery provides the lowest MAE, although the improvement ranges from 0.3% for 48 hour forecasts with MetNet-1, up to 7.3% improvement over just NWP for 8 hour forecasts with the PVNet-2.

Inputs	MetNet-1 (48 hr)	3DConv (2 hr)	PVNet-1 (8 hr)	PVNet-2 (8 hr)
NWP	1.003	1.032	1.029	1.073
Sat	1.109	1.009	1.052	1.243
NWP+Sat	**1.0**	**1.0**	**1.0**	**1.0**

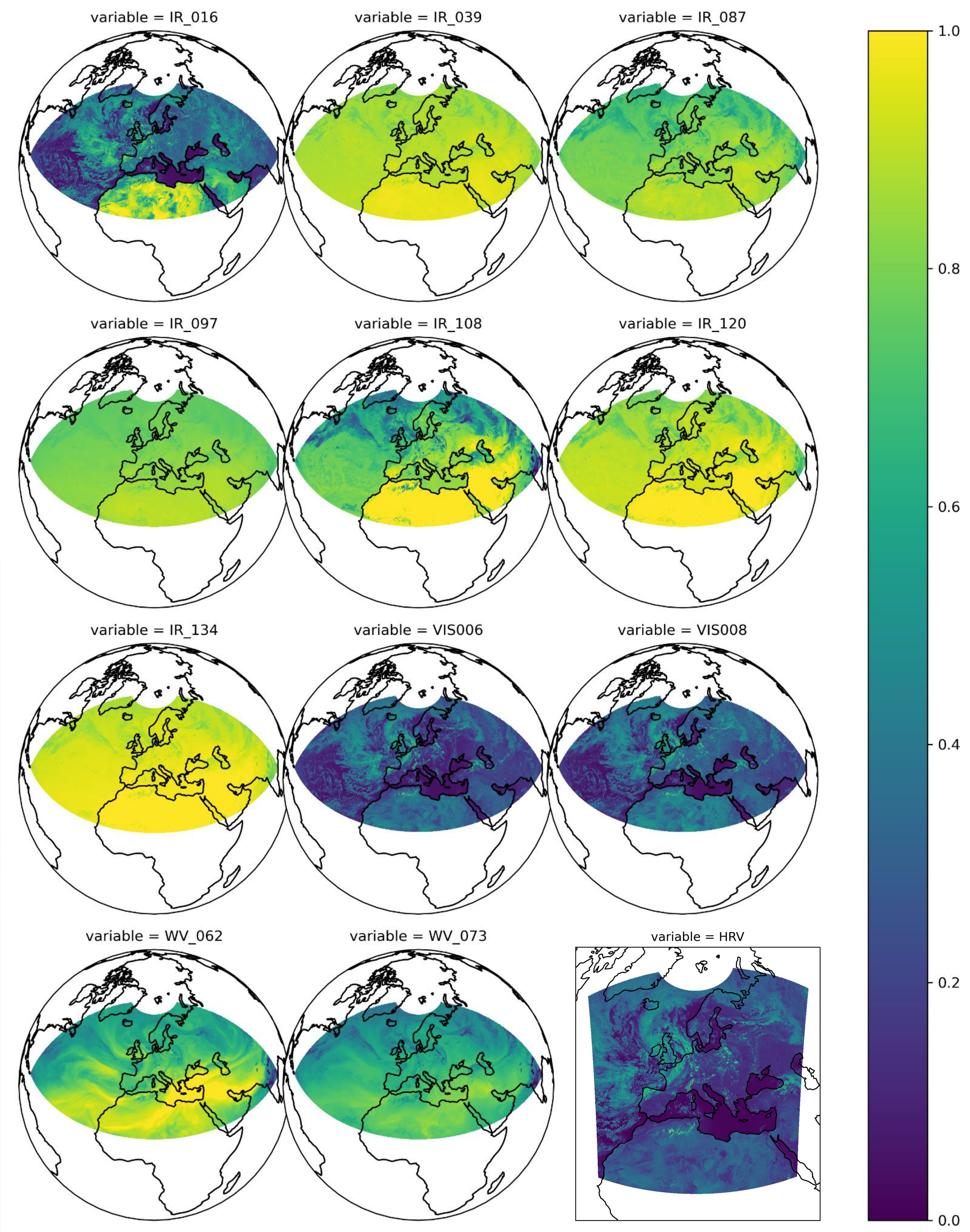
Average MAE across all forecast horizons normalized by the baseline NWP+Sat inputs



MAE per forecast horizon for PVNet-2 with various inputs

Satellite Dataset

As part of this research, we have also released nearly the entire archive of EUMETSAT 5-minutely Rapid Scan Service imagery on Google Public Datasets [5]. This dataset is comprised of all 12 spectral channels and the entire spatial extant from 2008 to the present, being updated roughly every week with new data. The data has been transformed from the original values by using Satpy [6] to calibrate them, and rescaling the values between 0 and 1. The current dataset is ~60 TB and freely accessible in Zarr format. Below is an example of the spatial coverage for each channel for noon on June 27th, 2020.



Discussion

Satellite imagery provides a useful input to ML models for PV forecasting that complements the information available from NWP forecasts and allows for more accurate short term forecasts, while not being as useful at longer time horizons. Using the imagery does bring up some challenges, primarily around the amount of data that is needed and increases the reliance of forecasts on more services that can go down for planned and unplanned maintenance.

To encourage more use of satellite imagery in solar PV forecasting, we have released an archive of European satellite imagery for the last several years, complimenting a previous release of UK PV data [7].

Next steps include examining if these results are consistent across other regions of the world, including places with more varied climates than the UK, such as India, or central Europe, as well as with more model architectures. The models and data used in this work is public, and we welcome contributions from the community, including expanding the experiments to new regions, adding newer models or supports for other satellites. All development is performed openly on GitHub [8].

References

[1] https://github.com/openclimatefix/pseudo-labeller
[2]Sønderby, Casper Kaae, et al. "MetNet: A Neural Weather Model for Precipitation Forecasting." (2020).
[3] https://github.com/openclimatefix/metnet
[4] https://github.com/openclimatefix/PVNet
[5]https://console.cloud.google.com/marketplace/product/bigquery-public-data/eumetsat-seviri-rss-hrv-uk
[6] https://github.com/pytroll/satpy
[7] https://huggingface.co/datasets/openclimatefix/uk_pv
[8] https://github.com/openclimatefix

Acknowledgements

We would like to thank EUMETSAT and the UK MetOffice for providing data used in this project, Nvidia for providing us with GPUs, and Google for hosting the public satellite dataset. We would also like to thank the help and collaboration of the rest of the OCF team, without which this project would not have been successful.