Impact of Meteorological Dataset Selection on Photovoltaic Performance Modelling: A Canadian Case Study

Summary

In this study, we evaluated the performance of six major irradiance products in Canada: three satellite-based models (CWEEDS, PSM4, NASA POWER) and three reanalysis datasets (ERA5, MERRA-2, and CaSR) against a network of 92 Canadian ground stations and 4 research-grade sites. Satellite models demonstrated ~25% median global horizontal irradiance (GHI) root mean square error (RMSE), while reanalysis error was ~33%. Direct normal irradiance (DNI) RMSE was higher across the board, varying between 40-80% by model and location. Using pvlib, we further quantified how regional biases in GHI and DNI propagated through the PV modeling chain. We simulated the energy yield of fixed-tilt and single-axis tracking (SAT) systems in four locations, using weather data from ground stations and CWEEDS, PSM4, and NASA POWER datasets. Preliminary results show that annual energy yield predictions vary by 0.5-9% in southern Canada, with deviations reaching up to 17% for SAT systems in the Arctic due to their high sensitivity to DNI.

Keywords: Solar resource, satellite, reanalysis, PV modelling, uncertainty propagation, error, Canada

Field of application for the findings: Solar Resources and Energy Meteorology

Addressed audience: solar resource researchers; PV modellers (industry and academic)

1. Introduction

Reliable solar irradiance data is essential for the monitoring and expansion of photovoltaic (PV) technologies, yet ground station sparsity remains a significant barrier to understanding irradiance data and PV model uncertainty, particularly in high-latitude regions. While satellite and reanalysis models offer increased data accessibility, their accuracy must be rigorously validated against ground truth measurements. This study evaluates the performance of six major solar data products in Canada against a network of 96 ground stations, and further quantifies how the satellite model choice impacts the outcomes of PV modelling in four locations.

2. Methodology

We first validated the performance of six weather models by comparing modelled irradiance to 4 research-grade reference stations in Canada and 92 non-reference ground stations, amounting to over 1200 station-years of data evaluated per model for GHI. A subset of the stations contain DNI data, resulting in an evaluation of 52 station-years of data per model with DNI data. All ground station data was quality controlled using the procedures outlined by Forstinger *et al.*, 2021. The ground stations examined in this work are shown in Figure 1. Table 1 summarizes the six weather data sources examined in this work: the Canadian Weather Canadian Weather Energy and Engineering Datasets (CWEEDS), NREL's Physical Solar Model v4 (PSM4), NASA POWER, ERA5, MERRA-2, and the Canadian Surface Reanalysis dataset (CaSR). RMSE and mean bias error (MBE) were calculated across all 96 ground stations and models, where applicable.

Following the validation analysis, we studied how irradiance errors propagate through the PV modelling chain for the 4 research-grade sites in Canada. We simulated the performance of two PV configurations: a fixed-tilt array and a horizontal single-axis tracker (SAT). To account for sky anisotropy, plane-of-array (POA) irradiance was calculated using the Perez sky diffuse model [Perez *et al.*, 1990], integrated with pvfactors. DC electrical yield was calculated using the California Energy Commission 6-parameter single-diode model [Dobos *et al.*, 2012]. Two parallel simulation tracks were executed: (1) the PV model was driven by high-quality ground measurements and (2) the identical PV model was driven by the three satellite datasets

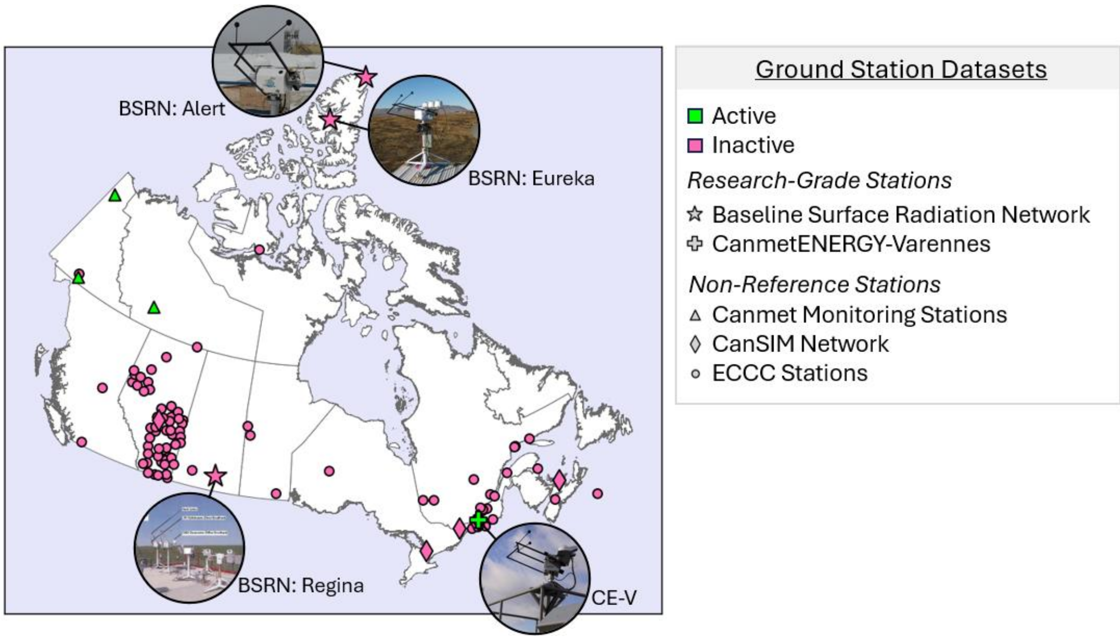

Figure 1: Ground stations analyzed in this work. Pictured sites are the research-grade stations.

Table 1. Summary of satellite and reanalysis models examined in this work.

Model Abbrev.	Model Type	Coverage	Temporal Resolution	Spatial Resolution	Years Covered
CWEEDS	Satellite	Canada	60-minute	649 sites. Underlying SolarAnywhere model: 0.1° x 0.1° (11 km)	1998-2023
PSM4	Satellite	USA & Americas, <60°N	30, 60-minute	4 km	1998-2023
		Polar region ≥60°N	60-minute	4 km	2013-2023
NASA POWER	Satellite	Global	60-minute	Solar variables: 1° x 1° Other: 0.625° x 0.5°	2001-present
ERA5	Reanalysis	Global	60-minute	0.25° x 0.25° (31 km)	1940-present
MERRA-2	Reanalysis	Global	60-minute	0.625° x 0.5°	1980-present
CaSR	Reanalysis	Americas, Greenland	60-minute	10 km	1980-2023

3. Results

3.1. Irradiance Model Validation

The results of the satellite and reanalysis dataset validation work are shown in Figure 2. Satellite models showed a notable increase in performance compared to reanalysis datasets, as anticipated, with ERA5 demonstrating the lowest error amongst studied reanalysis models.

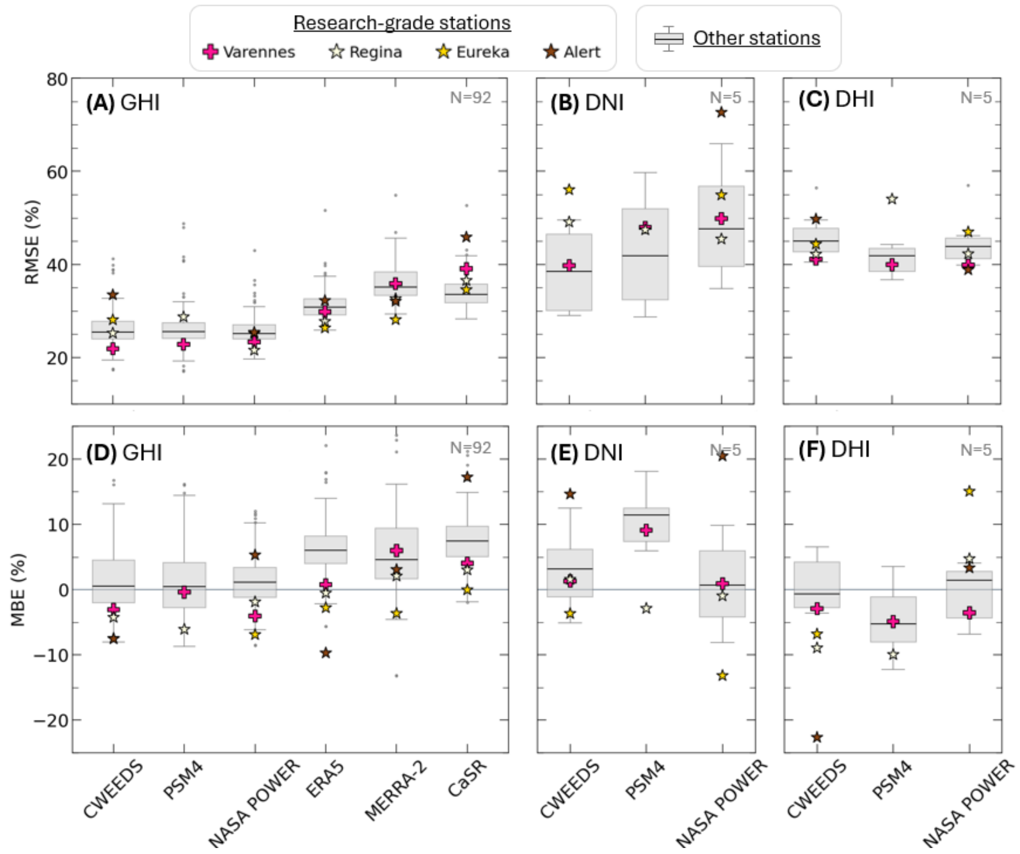

Figure 2: RMSE (a-c) and MBE (d-f) for GHI, DNI, and DHI irradiance components. Research-grade stations are shown as individual markers. The number of stations included in the grey boxplots is indicated in the top right corner of each subplot. Note that no overlap exists between the PSM4 model and data collection period of Eureka and Alert BSRN stations.

3.2. PV Modelling

Based on this validation, PV modelling was undertaken using the three satellite models and ERA5. Preliminary results indicate that energy yield deviations are significantly higher for SAT systems compared to fixed-tilt configurations in the Arctic locations of Eureka and Alert. This is likely attributed to the increased sensitivity of trackers to DNI. With DNI RMSE in the range of 55-82% in the Arctic, annual energy yield differences reach as high as 17% compared to ground-measured data in certain cases. In the southern locations of Regina and Varennes, these differences vary between 0.5-9%, depending on the model and PV system. Further results will be explored and demonstrated in the final conference proceeding.

4. Conclusion

These preliminary results underscore the critical importance of data quality and access in circumpolar regions, and quantify the impact of weather data selection when modelling PV systems. This work highlights a need for expanded high-quality ground station measurement across Canada's diverse climate zones to inform and improve the next generation of satellite and reanalysis models for high-latitude applications.

5. References

Forstinger, A. *et al*., "Expert quality control of solar radiation ground data sets," *Proceedings of SWC 2021: ISES Solar World Congress*, 1037-1048 (2021). 10.18086/swc.2021.38.02

Perez, R., Ineichen, P., Seals, R., Michalsky, J., and Stewart, R., "Modeling daylight availability and irradiance components from direct and global irradiance," *Solar Energy*, 44 (5), 271-289 (1990). https://doi.org/10.1016/0038-092X(90)90055-H

Dobos, A. P., "An Improved Coefficient Calculator for the California Energy Commission 6 Parameter Photovoltaic Module Model", *Journal of Solar Energy Engineering*, 134(2) (2012). https://doi.org/10.1115/1.4005759