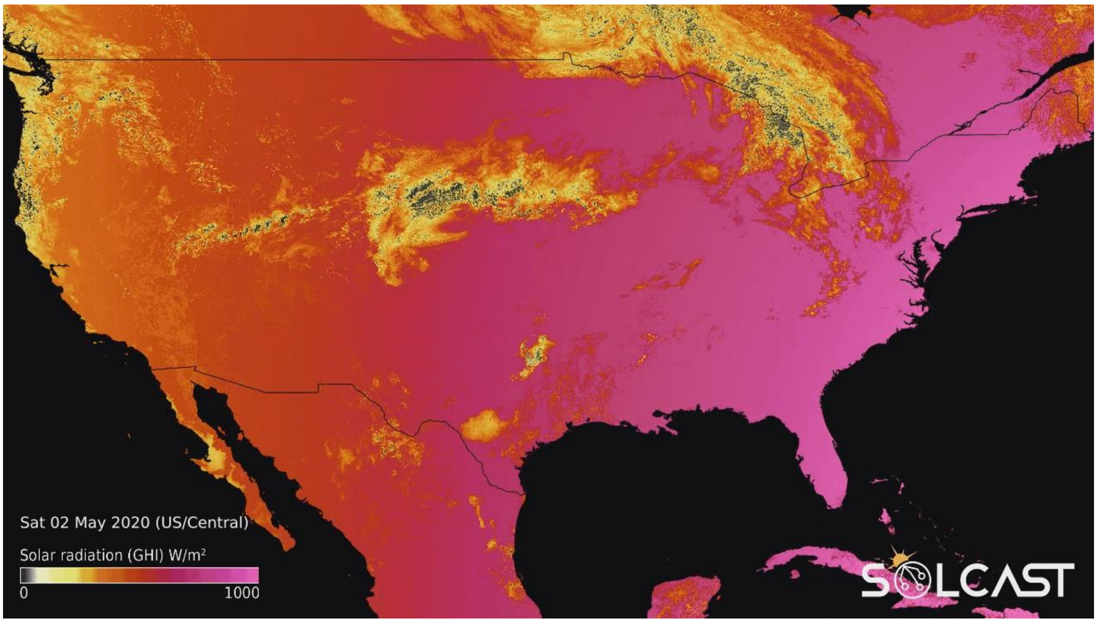
Solcast Live and Forecast Data

Verification and Accuracy Report

Last revision: 29th March, 2022



Caption text goes here. Right-click image to change picture.

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

A global, 12-month verification analysis of Solcast Live and Forecast Data is performed using surface measurements from 33 sites spanning a range of climate types and latitude zones. Day-ahead forecasts of daily total energy production for single-axis tracking sites exhibit mean MAPE errors of 5.9% averaged across a range of site climates, compared to 7.7% for a commonly-used global weather model (GFS). Intra-day forecasts of hourly energy production exhibit reduced MAPE of 2.9% and 4.1% at 1 and 2 hours ahead, respectively, due to satellite cloud tracking. For the Arid and Semi-Arid (precipitation < 750mm per year) locations at lower latitudes (<35 degrees) favoured by solar developers, Solcast forecasts exhibit errors approximately 1% lower than these global-mean values. Near real-time satellite-derived GHI estimated actuals exhibit a mean nRMSE error of compared to surface measurements of 3.8%, compared to 8.2% for GFS, for hourly average values.

2 LIVE AND FORECAST DATA PRODUCTS

Solcast operates a global cloud tracking and forecasting system, using near-real-time satellite imagery from 11 weather satellites, and weather data from 7 numerical weather prediction (NWP) models. These inputs are used to make irradiance and PV power data, which is distributed via the Solcast API (Application Programming Interface), which enables automated, synchronous data requests for any point on earth.

This paper focuses on live and forecast data, from -7 to +14 days from now. The live and forecast data products deliver PV power, irradiance, and weather data globally, at time granularities of 5, 10, 15, 30 and 60 minutes. Coverage is all non-polar continental areas and nearby islands, with spatial resolution of 2km and data updates every 5 to 15 minutes.

3 ACCURACY VERIFICATION METHODOLOGY

Accuracy is typically the most important product attribute amongst commercial users of the Solcast API. Customer evaluations or trials can produce definitive results; however they require a large amount of resource, expertise, and elapsed time to gather and analyse the data, and to quality control the surface measurements.

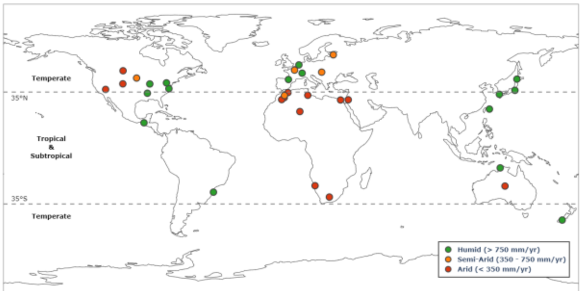
The purpose of this analysis is to enable users to estimate Solcast real-time and forecast accuracy for their site(s) prior to subscription or integration effort. It is based on 12 months of recent Solcast forecast and estimated actuals data, compared to surface measured actuals from high-quality measurement sites. Also included are error statistics, and benchmarks against common alternatives. Users focused on historical data or TMYs will find a separate, multi-year analysis on the Solcast website. This document analysis irradiance (GHI, DNI, GTI) and power production (fixed-tilt and single-axis tracking). For brevity, a selection of commonly used statistics are included, summarised by climate type. For a full range of statistics, at site level, and to request the raw data, please contact the Solcast team.

3.1 Measurement site selection

This analysis uses research grade, quality-controlled surface measurements from three networks: Baseline Surface Radiation Network (BSRN; global), the enerMENA Meteo Network (Middle East and North Africa), and the Bureau of Meteorology (BoM; Australia). The criteria for selecting these networks was (1) quality, characterised by robust calibration and maintenance standards and quality control of data - this is challenging for irradiance measurements, where calibration and maintenance often cause large errors; (2) recency, i.e. data is available for recent periods so that the analysis can use recent Solcast algorithmic configuration, and for relevance to newer customer sites; (3) availability, the measurements should be non-private, and readily available to users who may want to replicate the results; and (4) broad geographic and climate-type coverage, so that users can estimate accuracy for their own sites.

Sites with very high elevation (over 2000 metres) have been excluded due to limited applicability to solar energy applications. Otherwise, all sites with data available during the analysis period are included. Across these three networks a total of 33 sites are included. For climate type categorisation, two dimensions are used: (1) the latitude zone (Tropical/Subtropical being latitudes equatorward of 35 degrees, and Temperate being latitudes between 35 and 60 degrees north and south (no sites poleward of 60 degrees are included); and (2) site climate type based on annual average precipitation (Humid being sites of greater than 750mm per year, Arid sites are less than 350mm per year, and Semi-Arid sites are between 350mm and 750mm per year) based on the CPC Merged Analysis of Precipitation (CMAP) from NOAA.

Solcast does not perform site-specific adaptation of its satellite-derived data (i.e. measurements from a site are not used in the dataset for that site), and the period of this analysis is not used in any regional algorithmic calibrations.



Map of measurement sites included in the analysis. Of the 33 included sites, a total of 18 are designated Tropical/Subtropical (7 Humid, 1 Semi-Arid, 10 Arid). A total of 15 are designated Temperate (8 Humid, 4 Semi-Arid, and 3 Arid).

3.2 Benchmark data

Two commonly available benchmarks are included, to put accuracy performance in context, and to assess the marginal value of Solcast data. For day-ahead forecasts and real-time estimated actual Live data, the benchmark is the Global Forecast System (GFS) model from NOAA, a commonly-used global weather forecast model. For simplicity, GFS data is taken at +24 hours ahead, even for intraday comparisons, because global weather models such as GFS are not available until ~5 hours after real time. For intra-day forecasts (i.e. +1 and +2 hours), we also include a benchmark called Smart Persistence, which is created by persisting the last hour's clearness index (i.e. the level of cloudiness is persisted, but clear-sky irradiance updates). This means the Smart Persistence benchmark is aware of sun position, and does not, for example, persist irradiance values after sunset).

3.3 Method

During the 12-month period covering April 2020 to March 2021 inclusive, measurement data is 88% complete on average. All data were converted to hourly means prior to analysis. Forecast horizons available from the Solcast archive at the time were +1 hour, +2 hours and day-ahead only. A range of error statistics were calculated following "Metrics for Evaluating the Accuracy of Solar Power Forecasting" (NREL, 2013), including normalisation by AC capacity in the case of PV power. For irradiance statistics, normalisation uses standard test conditions (STC) irradiance of 1000W/m2, this enables sites with differing climates to be compared, and makes irradiance and PV power statistics more comparable.

Actuals power production data was derived by converting measured irradiance to power using Solcast's PV power model, which is based on the open source pvlib python tool with proprietary extensions, and adding noise in the irradiance to power conversion. The approach of simulating PV production rather than using actual plant data was chosen for the following reasons: (1) allows irradiance and power statistics to be available for the same sites; (2) ensures transparency and repeatability of results by users; (3) ensures a broad geographical spread of sites; (4) avoids conflation from plant availability and curtailment on accuracy statistics (in any case, users can specify real-time plant conditions in API requests). The plant specifications used are the Solcast defaults for 10MW Utility Scale sites, including an inverter loading ratio of 1.30. A noise model was used to model the error in the PV power model itself, in order to make accuracy results realistic. The noise model was trained on PV plant measurements and corresponding Solcast PV model estimates from ten sites globally where full plant specifications have been given to the PV model. PV Power and GTI, which use a Hay transposition model are not calculated unless a site has both GHI and DNI measurement.

This analysis excludes probabilistic statistics, because most users only require deterministic statistics. Users interested in probabilistic measures may note that forecast data from the Solcast API does include probabilistic 10th and 90th percentile values, which are dynamically generated using a combination of satellite cloud tracking and a range of NWP models.

Nocturnal zero values are included in this verification for hourly measures; users interested in daytime results only for hourly values can approximate results via doubling the error values listed here. For daily total energy values, there is no difference in statistics owing to inclusion of nocturnal values, so the statistics presented here remain directly applicable.

4 ACCURACY VERIFICATION RESULTS

4.1 Power production accuracy

The following tables show statistics for the Mean Absolute Percentage Error (MAPE), as defined in the above-mentioned NREL 2013 paper, of power production for single 10MW PV plants with horizontal single axis tracking. The statistics are grouped by climate type and latitude zone. For a complete range of statistics and PV system types, and results for specific measurement sites, please contact Solcast.

MAPE (%) errors of Day Ahead PV Power Forecasts

Measure: Daily total energy production (single-axis tracking)

Site type	Forecast	Day-ahead (+24hr) error (%)
Tropical / Subtropical, Arid & Semi-Arid (11 sites)	**Solcast**	**4.4% (2.1% to 5.9%)**
	Smart Persistence	7.6% (2.7% to 9.4%)
	GFS	5.8% (2.4% to 10.0%)
Tropical / Subtropical, Humid (7 sites)	**Solcast**	**7.6% (6.1% to 9.4%)**
	Smart Persistence	15.7% (7.9% to 20.6%)
	GFS	10.4% (5.3% to 13.2%)
Temperate, Arid & Semi-Arid (7 sites)	**Solcast**	**6.0% (4.4% to 7.1%)**
	Smart Persistence	11.1% (4.0% to 17.5%)
	GFS	8.1% (6.1% to 10.5%)
Temperate, Humid (8 sites)	**Solcast**	**6.5% (4.9% to 9.2%)**
	Smart Persistence	16.6% (12.1% to 21.2%)
	GFS	8.0% (6.1% to 10.5%)
Global average (all 33 sites)	**Solcast**	**5.9% (2.1% to 9.4%)**
	Smart Persistence	13.2% (2.7% to 21.2%)
	GFS	7.7% (2.4% to 13.2%)

MAPE (%) errors of Intraday & Day-Ahead PV Power Forecasts Measure: Hourly energy production (single-axis tracking), nocturnal zeros included				
Site type	**Forecast source**	**+1 hours ahead error (%)**	**+2 hours ahead error (%)**	**Day-ahead (+24hr) error (%)**
Tropical / Subtropical, Arid & Semi-Arid (11 sites)	**Solcast**	**2.3% (1.4% to 2.9%)**	**3.1% (1.9% to 3.8%)**	**3.9% (2.9% to 5.0%)**
	Smart Persistence	3.0% (2.0% to 3.7%)	4.1% (2.6% to 5.1%)	5.8% (3.8% to 8.6%)
	GFS	4.6% (3.7% to 5.9%)		
Tropical / Subtropical, Humid (7 sites)	**Solcast**	**3.0% (2.4% to 3.8%)**	**4.5% (3.2% to 5.6%)**	**6.1% (4.5% to 7.0%)**
	Smart Persistence	4.4% (3.0% to 5.3%)	5.8% (3.7% to 6.9%)	10.1% (5.3% to 12.3%)
	GFS	7.3% (4.6% to 8.5%)		
Temperate, Arid & Semi-Arid (7 sites)	**Solcast**	**3.8% (2.8% to 5.0%)**	**4.8% (3.5% to 6.1%)**	**5.3% (3.5% to 6.1%)**
	Smart Persistence	4.9% (3.4% to 6.3%)	6.0% (3.9% to 7.3%)	10.0% (5.8% to 11.4%)
	GFS	6.1% (3.3% to 7.8%)		
Temperate, Humid (8 sites)	**Solcast**	**2.9% (2.4% to 3.8%)**	**4.2% (3.4% to 6.2%)**	**5.5% (4.1% to 8.2%)**
	Smart Persistence	4.3% (3.3% to 5.5%)	5.5% (4.3% to 7.4%)	10.8% (7.8% to 14.0%)
	GFS	6.4% (4.8% to 9.2%)		
Global average (all 33 sites)	**Solcast**	**2.9% (1.4% to 5.0%)**	**4.1% (2.5% to 6.2%)**	**5.1% (2.9% to 8.2%)**
	Smart Persistence	4.0% (2.0% to 6.3%)	5.2% (2.6% to 7.4%)	9.0% (3.8% to 14.0%)
	GFS	6.0% (3.3% to 9.2%)		

4.2 Irradiance accuracy

The following tables show statistics for the normalised Root Mean Square Error (nRMSE), as defined in the above-mentioned NREL 2013 paper, of GHI irradiance. For a full range of statistics (including DNI and GTI), at site level, and to request the raw data, please contact the Solcast team.

nRMSE (%) errors of Intraday & Day-Ahead Irradiance Forecasts & Estimated Actuals

Measure: GHI hourly average, nocturnal zeros included

Site type	Forecast	Real-time	+1 hour	+2 hour	+24 hour
Tropical / Subtropical, Arid & Semi-Arid (11 sites)	Solcast	**3.5% (1.6% to 5.9%)**	**3.6% (1.7% to 6.0%)**	**4.8% (2.9% to 7.5%)**	**5.6% (3.4% to 9.7%)**
	Smart Persistence	N/A	4.8% (3.1% to 7.3%)	6.2% (4.1% to 8.8%)	8.8% (4.3% to 12.1%)
	GFS	6.5% (3.4% to 10%)			
Tropical / Subtropical, Humid (7 sites)	Solcast	**3.9% (3.1% to 5.0%)**	**4.3% (3.3% to 5.5%)**	**6.4% (4.8% to 7.7%)**	**8.1% (6.1% to 9.4%)**
	Smart Persistence	N/A	6.4% (4.5% to 7.3%)	8.1% (5.7% to 9.1%)	13.8% (8.2% to 16.5%)
	GFS	10.1% (7.2% to 11.9%)			
Temperate, Arid & Semi-Arid (7 sites)	Solcast	**4.7% (3.3% to 7.0%)**	**4.8% (3.5% to 7.2%)**	**5.9% (4.4% to 8.3%)**	**6.5% (4.8% to 7.8%)**
	Smart Persistence	N/A	6.2% (4.3% to 8.9%)	7.8% (5.4% to 10.2%)	12.1% (8.1% to 15.0%)
	GFS	8.2% (5.3% to 9.3%)			
Temperate, Humid (8 sites)	Solcast	**3.6% (2.6% to 4.3%)**	**3.9% (2.9% to 4.8%)**	**5.6% (4.1% to 7.8%)**	**7.0% (5.3% to 9.9%)**
	Smart Persistence	N/A	5.8% (3.8% to 7.2%)	7.5% (5.1% to 9.6%)	13.5% (8.7% to 16.7%)
	GFS	8.8% (6.2% to 11.7%)			
Global average (all 33 sites)	Solcast	**3.8% (1.6% to 7.0%)**	**4.1% (1.7% to 7.2%)**	**5.6% (2.9% to 8.3%)**	**6.7% (3.4% to 9.9%)**
	Smart Persistence	N/A	5.7% (3.1% to 8.9%)	7.3% (4.1% to 10.2%)	11.8% (4.3% to 16.7%)
	GFS	8.2% (3.4% to 11.9%)			