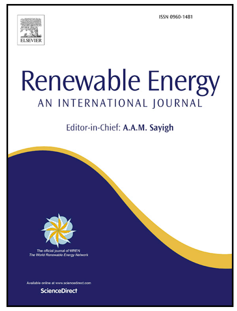
Accepted Manuscript



Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts

Hugo T.C. Pedro, Carlos F.M. Coimbra, Mathieu David, Philippe Lauret

PII: S0960-1481(18)30142-3

DOI: 10.1016/j.renene.2018.02.006

Reference: RENE 9737

To appear in: *Renewable Energy*

Received Date: 5 September 2017

Revised Date: 31 January 2018

Accepted Date: 1 February 2018

Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts

Hugo T.C. Pedro[a], Carlos F. M. Coimbra[a,*], Mathieu David[b], Philippe Lauret[b, a]

[a]*Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center for Energy Research*

University of California, San Diego, La Jolla, CA 92093, USA

[b]*Université de La Réunion; 15 avenue René Cassin, 97715 Saint-Denis Cedex 9, France*

[*]Corresponding author

Email address: ccoimbra@ucsd.edu (Carlos F. M. Coimbra)

Abstract

This work compares the performance of machine learning methods (*k*-nearest-neighbors (*k*NN) and gradient boosting (GB)) in intra-hour forecasting of global (GHI) and direct normal (DNI) irradiances. The models predict the GHI and DNI and the corresponding prediction intervals. The data used in this work include pyranometer measurements of GHI and DNI and sky images. Point forecasts are evaluated using bulk error metrics while the performance of the probabilistic forecasts are quantified using metrics such as Prediction Interval Coverage Probability (PICP), Prediction Interval Normalized Averaged Width (PINAW) and the Continuous Ranked Probability Score (CRPS). Graphical verification displays like reliability diagram and rank histogram are used to assess the probabilistic forecasts. Results show that the machine learning models achieve significant forecast improvements over the reference model. The reduction in the RMSE translates into forecasting skills ranging between 8% and 24%, and 10% and 30% for the GHI and DNI testing set, respectively. CRPS skill scores of 42% and 62% are obtained respectively for GHI and DNI probabilistic forecasts. Regarding the point forecasts, the GB method performs better than the *k*NN method when sky image features are included in the model. Conversely, for probabilistic forecasts the *k*NN exhibits rather good performance.

Keywords: Probabilistic solar forecasts; Global Irradiance; Direct irradiance; Machine Learning; Sky imagery

1. Introduction

Solar forecasts at various time horizons are needed in order to increase the share of solar energy into electricity grids. Indeed, the intermittent character of solar energy may result in imbalances between electricity supply and demand. This requires the power system to either procure additional reserves or adjust the output of conventional generators so as to ensure balance between supply and demand (Inman et al., 2013). It is therefore important that solar irradiance and the corresponding solar power output are accurately predicted so that the utility grid is able to take appropriate actions to manage intermittency.

In this study, we focus on intra-hour global horizontal irradiance (GHI) and direct normal irradiance (DNI) forecasts with forecasting time horizons ranging from 5 min up to 30 min. Targeting these two irradiance components is important: DNI is of particular interest to concentrating solar power plants (CSP) and installations that track the position of the sun; and both DNI and GHI can be used to estimate the plane-of-array irradiance on tilted/tracking PV panels. With respect to the time horizons, as mentioned by (Inman et al., 2013), intra-hour forecasts are relevant for optimal central plant operations, and is an enabling technology for optimal dispatch of ancillary resources and storage systems. Moreover, mitigation measures for large drops in solar irradiance, such as demand response, storage and intra-hour scheduling can only be maximized with accurate and reliable intra-our forecast (Inman et al., 2013).

Depending on the forecast horizon, different types of forecasting models are appropriate (Kostylev and Pavlovski, 2011). In the solar forecasting community, numerous works have been devoted to the development of models that generate point forecasts also called deterministic forecasts (Lauret et al., 2015; Voyant et al., 2017; Trapero et al., 2015; Huang et al., 2013). In addition, it must be noted that