Applied Energy 217 (2018) 189–211

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Large-scale rooftop solar photovoltaic technical potential estimation using Random Forests

Dan Assouline[a,*], Nahid Mohajeri[b], Jean-Louis Scartezzini[a]

[a] Solar Energy and Building Physics Laboratory (LESO-PB), Ecole Polytechnique Federale de Lausanne, 1015 Lausanne, Switzerland
[b] Sustainable Urban Development Programme, Department for Continuing Education, University of Oxford, Rewley House, 1 Wellington Square, Oxford OX1 2JA, United Kingdom

HIGHLIGHTS

- Rooftop photovoltaic potential is quantified in $(200 \times 200)\,m^2$ pixels in Switzerland.
- Geographic Information Systems (GIS) tools are used for data processing.
- Random Forests are used to estimate solar and urban variables in the whole country.
- Prediction Intervals are computed to measure the uncertainty of the estimations.
- The rooftop PV production can cover 25% of Switzerland's demand in 2017.

ARTICLE INFO

2010 MSC:
00-01
99-00

Keywords:
Rooftop photovoltaics
Solar energy potential
Geographic Information Systems
Machine Learning
Random Forests

ABSTRACT

Photovoltaic (PV) panels are a very promising technology that answers part of the increasing need for global renewable energy production, particularly in urban areas. We present a novel methodology combining Geographic Information Systems (GIS), solar models and a Machine Learning (ML) algorithm, Random Forests, to estimate the potential for rooftop PV solar energy at the scale of a country. We use a hierarchical approach which divides the computation of the final potential into several steps. Each step is reached by estimating multiple variables of interest using widely available data, and combining these variables into potential values. The method for estimating each variable of interest is as follows: (1) collect all the data related to the variable, (2) train a Random Forest model based on the collected data and (3) use the model to predict the variables in unknown locations. The variables of interest include available area for PV installation on rooftops, shape, slope and direction of rooftops, global solar horizontal and tilted radiations, as well as shading factors over rooftops. The study focuses on Switzerland and provides the rooftop PV technical potential for each $(200 \times 200)\,[m^2]$ pixel of a grid covering the entire country. The methodology, however, is generalizable to any region for which similar data is available and could therefore be useful for researchers, energy service companies, stockholders and municipalities to assess the rooftop PV capacity of the region. Prediction Intervals are also provided for the different estimated variables, in order to measure the uncertainty of the estimations. The results show that Switzerland has a large potential for rooftop PV installations. More specifically, for roofs orientated at $\pm$ 90° from due south, the total estimated potential PV electricity production is about 16.29 TWh/year, which corresponds to 25.3% of the total electricity demand in 2017.

1. Introduction

Switzerland has ambitious goals for increasing its use of renewable energy and reducing CO_2 emissions. In particular, the Swiss Energy Strategy 2050 aims at phasing out of nuclear energy by 2035 and a possible 50–80% reduction in CO_2 emission by 2050. These aims can partly be reached through a great increase in the production of renewable energy and associated development of grid and storage capacity. One way to do so is building large solar farms in rural areas. Another, and complementary way is to increase the on-site solar-energy production in urban areas, that is, decentralized electricity production through photovoltaic (PV) technology on the building rooftops. Decentralized PV electricity production widely regarded as contributing favorably to environmental, economic, and social aspects of urban

* Corresponding author.
E-mail addresses: dan.assouline@epfl.ch (D. Assouline), nahid.mohajeri@epfl.ch (N. Mohajeri), jean-louis.scartezzini@epfl.ch (J.-L. Scartezzini).

https://doi.org/10.1016/j.apenergy.2018.02.118
Received 6 September 2017; Received in revised form 1 February 2018; Accepted 16 February 2018

Nomenclature

β	roof tilted angle
$\Delta i(n)$	impurity decrease between node n and children nodes n_L and n_R
δ	declination (monthly estimated)
γ	roof azimuth angle
η	PV panel efficiency
$\mathbb{1}$	indicator function
$\mathbb{E}$	expectation operator
$\mathbb{P}$	probability operator
$\mathbb{R}$	set of real numbers
$\mathbf{x}$	generic input vector (realization of $X_1,...,X_d$)
$\omega_i(\mathbf{x},T)$	weight stored by tree T for training sample $\mathbf{x}_i$ (within a RF)
ω_{sr}	sunrise hour angle on tilted surface
ω_{ss}	sunset hour angle on tilted surface
ω_s	sunset hour angle
ϕ	latitude
ρ	foreground's albedo
θ	incidence angle of the sun rays on a tilted plane
θ_Z	sun zenith angle
A	ratio of direct horizontal radiation to extra-terrestrial horizontal radiation
A_R^c	available roof area for PV installation over building cluster c
A_f^c	ground floor area for building cluster c
A_t^s	tilted area over roof surface c
B	number of trees in a Random Forest
b_j	number of buildings in pixel j
C_R^s	portion of ground floor area available for PV installation on the roof over building cluster c
C_R^s	portion of tilted area available for PV installation over roof surface s
d	generic number of features in training data
D_b	building density
D_p	population density
RMSE, E_R	Root Mean Square Error
F_{c_n}	spreading function distributing aspect and roof area values over a roof based on its type c_n
G_{Bt}	Direct Tilted Irradiance
G_B	Direct Horizontal Irradiance
G_{Dt}	Diffuse Tilted Irradiance
G_D	Diffuse Horizontal Irradiance
G_h	Global Horizontal Irradiance
G_{oh}	Extraterrestrial Horizontal Irradiance
G_{Rt}	Reflected Tilted Irradiance
G_t	Global Tilted Irradiance
h	total area spanned by the lateral sides for a hipped roof
h_R	available area for PV installation over the lateral sides for a hipped roof
$i(n)$	impurity of node n in a decision tree
I_{95}	95% Prediction Interval
$l(\mathbf{x},T)$	leaf reached by data point $\mathbf{x}$ when passed through tree T
m	number of variables to consider to split a node n in a decision tree
N	number of training points (size of training data)
N_{cells}	total number of 2×2 [m^2]raster cells over rooftops in pixel j
$N_{\text{sh,x}}$	total number of fully shaded 2×2 [m^2] raster cells over rooftops in pixel j, at time x
$N_{l(\mathbf{x},T)}$	number of training samples in leaf $l(\mathbf{x},T)$ in tree T
n_L	left child node of node n in a decision tree
N_{nL}	number of learning samples in left child node n_L in a decision tree
N_{nR}	number of learning samples in left child node n_R in a decision tree
N_n	number of learning samples in node n in a decision tree
N_{PV}	number of PV panels fitting within a roof polygon
n_R	right child node of node n in a decision tree
N_{test}	number of testing point (size of testing data)
$P_{j,\text{geo}}^{\text{OOSG}}$	geographical potential in pixel j, for OOSG and GEN zones
$P_{j,\text{geo}}^{\text{SON}}$	geographical potential in pixel j, for SON zone
$P_{j,\text{tech}}$	technical potential in pixel j
p_l^j	probability for a building cluster to have a roof with a main slope of β_l in pixel j
Q_α	α-quantile
q_m^j	probability for a building cluster to have a roof with a main aspect of γ_m in pixel j
R_b	beam irradiance factor
R_d	diffuse irradiance factor
c_n	roof type of class$n$
r_n^j	probability for a building cluster to have a roof of class c_n in pixel j
R_r	reflected irradiance factor
S_{hill}	hillshade average value over partially shaded cells
S_{sh}	daily ratio of fully shaded cells over rooftops
T	generic decision tree
$X_1,...,X_d$	generic input variables
Y	generic output variable
y	generic output value (label, realization of Y)
DHM	Digital Height Model
DOM	Digital Orthophoto Map
GEN	geneva canton zone
NRMSE, E_{NR}	Normalized Root Mean Square Error
OOB	Out-Of-Bag error
OOSG	Out-Of-SON-GEN zone, meaning the remaining zone of Switzerland once SON and GEN regions have been excluded
PI	Prediction Interval
PR	Performance Ratio
PV	photovoltaic
QRF	Quantile Regression Forest
RegBL	Registre des Batiments et Logements (Swiss building information data)
RF	Random Forest
SITG	Systeme d'Information du Territoire Genevois (Geneva Territory Information System - a precise GIS building database for Geneva canton)
SON	zone spanned by the Sonnendach project data
SVF	Sky View Factor
SVR	Support Vector Regression

Superscripts

c	aggregated per building cluster
j	averaged in pixel j
s	aggregated per roof surface

sustainability. Buildings rooftops within urban areas have always been interesting locations for PV system installation. However, developing an efficient method for finding and evaluating suitable roofs for the optimal placement of PV systems remains a challenge, particularly at a large scale. Taking the advantage of geospatial data and computational methods is the first step towards accurate solar potential estimation.

Several methods, at different scale of study, have been proposed for estimating the rooftop solar PV potential. While many studies focus on