Intra-hour Solar Forecasting for Photovoltaic Systems Integration in Weak Electric Grids

David Cañadillas Ramallo

Director	Ricardo Guerrero Lemus
Codirectores	Benjamín González Díaz
	Jan Peter Hartmut Kleissl
Coordinador	Fernando Luis Rosa González

Doctorado en Ingeniería Industrial, Informática y Medio Ambiente

Línea 1: Ingeniería industrial, tecnología electrónica y comunicaciones

Universidad de La Laguna

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Agradecimientos

Antes de entrar en materia, me permitirán que reserve este pequeño espacio para agradecer a todas las personas que me han apoyado a lo largo de estos años de doctorado y durante la redacción de la tesis, y sin las cuales, esta tesis no hubiese sido posible de una forma u otra.

En primer lugar, quiero agradecer a mi director Ricardo Guerrero, por haber confiado en mí desde el principio y por haberme apoyado siempre. Gran parte de la culpa de que yo acabara haciendo una tesis doctoral fue suya en primer lugar, y si bien es verdad que cuando empecé, yo no estaba muy convencido, ahora al final, no me arrepiento de haber emprendido este camino. También le agradezco el haberme hecho partícipe de los múltiples proyectos que se han surgido durante el desarrollo de la tesis, sin los cuales, ésta carecería del carácter práctico que yo tanto valoro, y estos años no hubiesen sido tan entretenidos. Por todo eso y mucho más, gracias Ricardo.

También quiero agradecer a mi codirector Benjamín González, por haber estado siempre ahí, y por haberme ayudado incondicionalmente siempre que se lo he pedido. Valoro mucho el haber tenido a alguien con quien hablar cuando lo he necesitado, y especialmente, el haberme ayudado a que todos estos años se hicieran más llevaderos. También le agradezco enormemente el haber sido mi guía en la selva hostil que para mí es la burocracia universitaria, pero por la que él sabe transitar como pocos. Y por supuesto, no olvido las veces que me has invitado a café del bueno. Muchas gracias Benji.

Me gustaría también agradecer a mi otro codirector, Jan Kleissl, por haberme acogido con tanta calidez en su grupo de investigación en la Universidad de San Diego, durante los tres meses de estancia que estuve allí, y por haberse involucrado de una manera activa en la investigación una vez estuve de vuelta, así como aceptar mi propuesta de ser mi codirector. Thank you, Jan!

Por supuesto, y más allá de los agradecimientos estrictamente académicos, quiero agradecer a mis padres, Mary y Carlos, por haberme educado como lo hicieron, y por haberme dado siempre su apoyo y comprensión en todas las etapas de mi vida. Ellos son los principales responsables de que hoy esté aquí escribiendo una tesis doctoral, ya que siempre que han podido no han dudado en ayudarme con lo que necesitase, ya fuera cuidando a mi perra Cuba cuando me iba de viaje para un congreso; o ayudándome con alguna herramienta cuando tenía que hacer un prototipo para un proyecto. Así que me gustaría que sintieran que parte de esta tesis es suya también. Muchas gracias por todo.

También quiero agradecer a mi hermano Fernando, que siempre ha sido un ejemplo para mí en muchas facetas de la vida y del que admiro su capacidad de seguir mirando la vida con la felicidad de un niño, por mucho que los años pasen; y al resto de familia en el extranjero (Amanda y Joan), que, aunque debido a las circunstancias no hemos podido

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

compartir mucho tiempo juntos últimamente, siento su apoyo desde el otro lado del mundo.

Por último, me gustaría hacer una pequeña ronda de agradecimientos a todos mis amigos y a todas las personas con las que he compartido partes de mi vida durante estos años. Quiero agradecer a mi grupo de amigos con los que siempre he podido contar cuando los he necesitado, especialmente en los momentos en los que la motivación era escasa, ya sea para tener conversaciones acerca de cómo me iba en la tesis, o para todo lo contrario, y hablar de cualquier cosa que no fuera el trabajo. Y también han aguantado mis múltiples trayectos desde y hacia la locura, y eso es algo que valoro enormemente. Así que a todos, y con especial cariño a Meso, Paula, Dani y Miriam; y al resto de la gente de los grupos de Telegram, muchas gracias.

También quiero agradecer a Adri y Alber, y al resto del grupo de escalada, por sacarme del abandono físico al que el doctorado me tenía relegado y hacerme "mover el trasero de la silla". No saben lo importante que ha sido para mí esos momentos donde la única preocupación era cómo hacer una vía y llegar hasta el final.

Me gustaría hacer una mención especial, por la empatía añadida que existe, a mis amigos "pioneros del doctorado", en especial a Diego, por enseñarme los caminos del doctorado y por hacerme saber que hay vida después del mismo; y a Silvia, porque nunca ha dudado en ofrecer su ayuda y consejo, y por haber luchado incluso más que yo por mis derechos.

También quiero agradecer a mis compañeros en la universidad, en especial a Pablo, por esos momentos de pausa y café, necesarios a lo largo de la mañana, donde surgían las conversaciones más inesperadas; y especialmente también a Alfredo, con quien he tenido el placer de coincidir en el último año, y que me ha aconsejado mucho con respecto a la tesis, y con quien he tenido conversaciones fascinantes acerca del mundo de las energías renovables, al que ambos pertenecemos.

Finalmente, quisiera agradecer a las personas que me han acompañado durante este largo camino y que me han dado parte de su tiempo, de su cariño y de su amor, y a las que tengo muy presentes igualmente en mi mente y corazón.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Preface

Forecasting the future, as one may expect, is a complicated task. Even more complicated, when some of the most complex natural systems on the planet (atmospheric, climatic) and human-crafted systems (electric) are involved. Back in 2016, I was an Energy Engineer who just got his Master's degree in Renewable energy, and I had the opportunity to start my PhD. "What would the research be about?", I asked. "Well, the idea is to predict the movement of passing clouds over the photovoltaic plants using cameras, in order to anticipate the variations that clouds cause to the power output of photovoltaic plants". Predict. Clouds. At that moment, I had a romantic idea of clouds, and remembered how, as a kid, I stare at them trying to identify animal or human shapes for brief moments until they suddenly change and disappear. It never crossed my mind that their movement should be predicted for any purpose. But I have to say that it seemed quite challenging, and I liked the idea of starting something new. So the challenge was accepted.

When I first started this journey, I did not know anything about forecasting or the techniques used to do so. In fact, I have little to no knowledge about many of the disciplines involved in the task: atmospheric dynamics, cloud formation, photography, image processing and computer vision, time-series analysis and prediction, etc. Moreover, this was a brand new line of research for the group itself, and the beginning was especially difficult, since I did have neither the theoretical background nor support required. However, the definition of challenge is "*something that needs great mental or physical effort in order to be done successfully and therefore tests a person's ability*". And that was exactly what this research has meant to me.

It is safe to say that today, I have completely change the way I see the clouds, and after almost five years analyzing cloud images, I will probably never be able to stare at the clouds with the joy of a kid ever again. But despite that little sacrifice, throughout the five years I have been doing this work, I have learned so many things that I feel like I have a completely new skill set in my possession right now. Who would have told me a few years ago that I would be able to code scripts, or understand the underlying nature of optimization algorithms or deep learning models? During these years, I have also had the pleasure to meet a lot of interesting and relevant people, I have had the privilege of working with some of the most important utilities in the power sector, I have visited amazing places... So I can only be grateful for all the experiences this PhD has brought me, and, although the road has been bumpy sometimes, I am very happy to finally reach the final stop of my PhD, which for sure will mean the beginning of a new phase in my life.

Although the three papers presented in this dissertation show part of the research made and the results obtained during the PhD, they are not a fair reflection of all the work that has been done during the five years. It should be considered that part of the research has been done in the framework of private projects (meaning that some processes take

more time since more actors are involved), and a big part of the latest research, dealing with Deep Learning models for image processing and time-series forecasting, is yet to be published. Nevertheless, all the research done during these years, and all the derived research to come, will be published in the coming years without doubts.

As of today, I can safely say that our research group now possesses a very specialized knowledge in several areas such as solar forecasting, photovoltaic and renewable energies in general, integration of variable renewable sources in electric grids, image processing techniques, machine and deep learning models, and optimization algorithms applied to complex energy problems, etc. that were not available before, at least in the Canary Islands and in the energy context.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Abstract

Environmental concerns and decreases in costs are shooting up the deployment of solar photovoltaics (PV) all over the world. As more photovoltaic capacity is added to the global electric mix, the intrinsic variable nature of solar resource, particularly variability in the short-term due to the clouds, is posing several challenges on the operation of electric grids. A series of multipurpose measures can be adopted to deal with these challenges.

First, solar resource and photovoltaic potential should be assessed in order to foreseen the possible future scenarios when photovoltaic energy will dominate the electricity mix. It is more likely that regions with good solar resource will experience a higher deployment, and therefore, they will be more vulnerable to the effects of solar variability. National and regional policies also play an important role on the installation rates of renewable energies, and even on the performance of the PV plants, where lack of incentives could lead to bad management of operation and maintenance. Policies should be carefully planned according to national or regional targets and to the overall technological and cost situation.

Second, the impacts derived from the variability of solar resource should be addressed. As these impacts of variability on electric grids depend on several factors, such as the voltage level the PV plants are connected or the relative size and robustness of the electric system, different approaches are considered for different situations. For instance, in the case of small and islanded electric systems (also called weak systems), with small numbers of high inertia generators and no interconnections, short-term variability of a large share of PV on the electric mix may lead to frequency deviations. In order to prevent these frequency deviations, intra-hour solar forecasting is an interesting approach, where the variations of solar PV in the near future are predicted, and the rest of the elements and devices in the electric system may be accommodated to deal with the ramps caused by the solar variability. The production of good intra-hour solar forecasts seems to be tied to the use of sky-imagers, since traditional time-series analysis does not capture the stochastic nature of clouds. However, costs of commercial sky-imagers compared to the quality of the produced forecasts are too high to be considered as an attractive solution for most system operators. Therefore, there is a necessity of reducing the costs of the technology, while improving its overall prediction quality, to enable a massive use of these types or forecasting systems.

On the other side, PV systems installed at the distribution level have a larger impact on the voltage of the feeders. Smart global controls for the PV inverters, taking advantage of new features such as reactive power compensation or active power curtailment, can be enabled to address these voltage fluctuations. With the improvement of the communication infrastructures, global control strategies can take into consideration all the PV systems in a distribution feeder, and they could solve an optimization problem to

select the best set of parameters for each PV inverter in the grid at any time. However, with high numbers of PV systems deployed in the distribution feeders, the complexity of the optimization problem increases dramatically, and thus, fast optimization algorithms or metaheuristics are required to find good solutions in acceptable times. Thus, the study and implementation of fast metaheuristics capable of solving these high complexity control problems should be on the agenda of distribution system operators for the next years.

In this thesis, the problems presented above have been tackled from different perspectives, and as a result, the three papers conforming this compendium present several solutions to deal with the variability of solar PV on weak electric systems. First, an extensive PV potential assessment using big-data techniques on the Canary Islands was performed, where the importance and effects of national and regional policies on the PV installation rates and overall performance was highlighted. Second, the early prototypes of the low-cost sky-imagers designed and built on the framework of this thesis to make intra-hour solar forecasts were deployed in three different locations, and the results from each experience and the overall forecasting ability (and extra functionalities of sky-imagers) are presented. Finally, an optimized global control for voltage regulation in distribution grids is proposed, using a fast metaheuristic to find the best sets of parameters for PV inverters in a distribution grid with a high PV penetration.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Table of contents

List of figures

List of Acronyms

ACO	Ant Colony Optimization
AI	Artificial Intelligence
ANN	Artificial Neural Networks
AR	Auto-Regressive
ARIMA	Auto-Regressive Integrated Moving Average
ARIMAX	Auto-Regressive Integrated Moving Average with eXogenous variables
ARMA	Auto-Regressive Moving Average
ARMAX	Mixed Auto-Regressive Moving Average with eXogenous variables
BD	Big Data
BOS	Balance of System
CBDR-RC	Common but Differentiated Responsibilities and Respective Capabilities
CBH	Cloud Base Height
CNMC	Comisión Nacional del Mercado y la Competencia
CNN	Convolutional Neural Networks
CSI	Clear Sky Index
DG	Distributed Generation
DHI	Diffuse Horizontal Irradiance
DL	Deep Learning
DNI	Direct Normal Irradiance
DSO	Distribution System Operator
EC	Evolutionary Computation
EDA	Estimation of Distribution Algorithms
EI	Effectiveness Index
ESS	Energy Storage Systems
EU	European Union
FiT	Feed in Tariffs
GA	Genetic Algorithms
GBT	Gradient Boosted Trees
GHG	Greenhouse Gases
GHI	Global Horizontal Irradiance
GIS	Geographic Information Systems
GRASP	Greedy Randomized Adaptive Search Procedure
GTI	Global Tilted Irradiance
HDR	High Dynamic Range
IEA	International Energy Agency
JBSA	Joint Base San Antonio
k-NN	k-Nearest Neighbors
KSI	Kolmogorov-Smirnov Integral
LCOE	Levelized Cost of Electricity
LVR	Line Voltage Regulators
MA	Moving Average
MAE	Mean Absolute Error

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
          UNIVERSIDAD DE LA LAGUNA

          María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
          UNIVERSIDAD DE LA LAGUNA

MAPE	Mean Absolute Percentage Error
MBE	Mean Bias Error
ML	Machine Learning
MLP	Multilayer Perceptrons
NARMAX	Non-linear Auto-Regressive Moving Average with eXogenous variables
NDC	Nationally Determined Contributions
NREL	National Renewable Energy Laboratory
nRMSE	Normalized Root Mean Squared Error
NWP	Numerical Weather Predictions
OLTC	On-Load Tap Changers
PCC	Point of Common Coupling
PF	Power Factor
PNACC	Plan Nacional de Adaptación al Cambio Climático
PNIEC	Plan Nacional Integrado de Energía y Clima
POA	Plane Of Array
PPA	Power Purchase Agreement
PSO	Particle Swarm Optimization
PV	Photovoltaics
R&D	Research and Development
RF	Random Forests
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Networks
RTM	Radiative Transfer Models
SA	Simulated Annealing
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
SARIMAX	Seasonal Auto-Regressive Integrated Moving Average with eXogenous variables
SI	Swarm Intelligence
SIFT	Scale Invariant Feature Transformation
SP	Smart Persistence
ss	Skill Score
STATCOM	Static Synchronous Compensators
SVC	Static Var Compensators
SVM	Support Vector Machines
TIC	Total Installed Costs
USA	United States of America
UTSA	University of San Antonio, Texas
VRE	Variable Renewable Energy
VUF	Voltage Unbalance Factor
VV	Volt-var
VW	Volt-Watt

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
             UNIVERSIDAD DE LA LAGUNA

             María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
             UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1.   Introduction

1.1   Global Energy Outlook and Energy Transition

During the last decades, the world has been undergoing an important energy transition. We are witnessing how energy systems, which have relied on the use of fossil fuels as main sources of energy for centuries, are today transforming to more sustainable, renewable and electrified ones. Obviously, this transition is not casual, and at least two reasons can be identified as main drivers of this transformation: environmental concerns due to climate change and global warming, and cost reductions in renewable energy.

On the one hand, the effects derived from **climate change** are becoming more evident every year, and the threats related to them are starting to equally concern the population in general, and the governments in particular. Numerous reports and data have been published showing the changes in measured variables (evidencing increases in annual surface temperature, sea level, or ocean heat content for example) and analyzing the possible impacts (health related, economics, agricultural...) of climate change and global warming [1], [2]. Irrefutable proofs in data are starting to convince every country of the importance of the problem, but the approaches selected to solve the problem ranges in a wide spectrum of decisions.

Although the political focus tends to have a limited temporal sight based on electoral periods, the scientific community have been warning about the hazardous effects of climate change for a long time [3]. The majority of the developed countries are taking the problem as seriously as it deserves (at least at a political level), with the European Union (EU) being the leader in this transition. As of 2019, 28 countries had issued "climate emergency" declarations, in many cases, accompanied by national plans and strategies for transitioning to more renewable and sustainable energy systems [4]. There are other countries that refuse to embrace the recommended measures to improve the situation, for a variety of reasons, sometimes due to ideological reasons (for instance the United States of America (USA) withdrawing from the Paris Agreement during Trump's election [5]), others due to economic issues (some developing countries do not have the economic means to achieve the transition).

1

There is also an ongoing argument for the "historical guilt" of the emissions. In this sense, developing countries, such as India, claim that more developed countries are responsible for the majority of the historic greenhouse gases (GHG) emissions, and the adoption of climate change politics in terms of emission reductions, may affect their developing capabilities. China, which is the current major GHG emitter, has historically emitted half of the cumulative emissions of the USA [6]. However, other countries and regions see this transition as an opportunity to achieve a green and sustainable development, and get things done the right way. In any case, the "historical guilt" debate has been approached since the first international climate treaty, back in 1992, where a Common but Differentiated Responsibilities and Respective Capabilities (CBDR-RC) principle was included, but enforcing the pledges of each country has been proved to be an impossible task. Probably this was the reason for the new modality of the Paris Agreement in 2015, in which each signing country committed to certain self-assigned objectives, in a framework of common and overarching climate goals to limit global temperature rise to 1.5 ºC [7].

Despite that not all the countries are on the same boat regarding the way to approach the issue, it is true that large efforts have been made to reach an agreement for the year 2050. In the same Paris Agreement, signing countries commit to make efforts in the reduction of GHG emissions through their Nationally Determined Contributions (NDC). Although, in many cases, the NDC goals are less ambitious than they supposed to be to achieve the 1.5 ºC limit goal. In terms of renewable electricity, current NDC targets will only cover 40% of the renewable energy needed to meet the climate goals.

On the bright side of things, the European Commission presented at the end of 2019 its European Green Deal, which is a set of policies to achieve carbon neutrality in Europe by the end of 2050, which aligns with the objectives of the Paris Agreement. The objective is to mobilize at least EUR 1 trillion to support sustainable investments over the next decade, as well as EUR 100 billion over the period 2021-2027, in what is called the "Just Transition Mechanism", to support the regions more vulnerable to the profound effects of the transition [8], [9]. However greater efforts should be made by the international community to address the problem, since the projection of the effect of current measures are showing to be clearly insufficient.

The second major reason provoking the energy transition is the **decrease in costs of all renewable energy technologies**, mainly in the power sector. During the decade 2010-2020, electricity costs from renewable have fallen sharply, principally driven by improvements in technology, an increase in the competitiveness of supply chains, economies of scale and acquired manufacturer experience [10]. These changes are different depending on the technology analyzed, but the decrease in costs is common to all technologies (with the exception maybe of hydropower).

2

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1: Introduction

In the energy sector, two indicators are commonly used to measure the costs related to electricity projects: the Levelized Cost of Electricity (LCOE) and the Total Installed Costs (TIC). The LCOE measures the cost required to produce a kWh of electricity accounting for all the costs of the installation through its entire useful life, and it is expressed in EUR/kWh, EUR/MWh, USD/kWh or USD/MWh (currency per unit of energy). The TIC, also known as Capital Costs, includes all the fixed, one-time payments required for producing electricity with a determined installation (including land, equipment, installation, etc.). In the electric sector, it is usually expressed as a ratio between the cost and the capacity of the installation. Thus, its unit is EUR/kW or USD/kW (currency per unit of capacity).

Costs related data along with average capacity factor for the wind and solar power technologies can be seen at a glance in Figure 1 [10]. The technology experiencing the largest cost reductions is solar photovoltaics (PV) by far. Although the global weighted average TIC for all the major renewable technologies have been falling continuously during the last decade, solar PV is the technology with the largest decrease in costs, with a fall from USD 4,702/kW in 2010 to USD 995/kW in 2019 (79%). Solar PV has become in the last few years the technology with the lowest TIC. Regarding wind power, onshore wind power maintains a steady decrease of installation costs from USD 1949/kW to USD 1473/kW (24%); and offshore wind follows the trend with a decrease from USD 4650/kW to USD 3800/kW (18%). Concentrated Solar Power (CSP), although with much higher costs in absolute terms, has also experienced a decline in costs in the last ten years from USD 8967/kW to USD 5774/kW (36%).
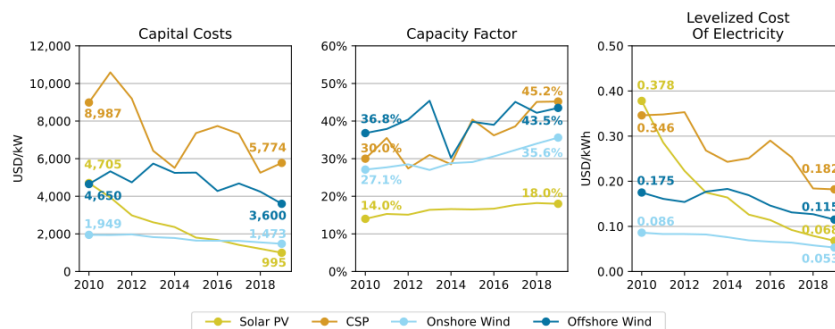
Figure 1. *Evolution of capital costs, capacity factors and LCOEs for solar and wind technologies over the last decade.*

Along with the decrease in TIC for all the technologies, a steady increase of the Capacity Factor is also observed, which of course, explained the generalized fall of LCOE for all technologies. According to [10], between 2010 and 2019, in terms of global weighted-average LCOE for utility scale plants, the costs of solar PV fell 82% (from USD 0.378/kWh to USD 0.068/kWh). The ranking is completed by the CSP, which also experienced an important fall in costs of 47% (from USD 0.35/kWh to USD 0.07/kWh),

3

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

onshore wind power with a 39% (from USD 0.086/kWh to USD 0.053/kWh) and offshore wind with 29% (from USD 0.161/kWh to USD 0.115/kWh).

With this context of generalized cost decreases, some renewable technologies are starting to out-compete the cheapest fossil fuel-based alternatives in many global regions. The value in solar PV is particularly significant, since in 2010, the LCOE of solar PV electricity was about 8 times higher than the cheapest coal-fired power generation. And the trends in PV and wind power show no sign of change in the next few years. Even during 2019, year-to-year costs continued decreasing. Some PV plants concurrent in Power Purchase Agreement (PPA) and auctions could have an average LCOE of USD 0.039/kWh, which represents a 42% reduction over the weighted average. According to [11], there are unsubsidized utility scale PV plants achieving LCOEs as low as USD 0.031/kWh. Some studies analyzing the costs of over 2000 GW of coal-fired power generation suggest that, taking into account these costs for PV, 1200 GW of coal-fired power plants could have higher marginal operating costs than the average price of new utility-scale solar PV [12].

On a regional scale, there are two main factors affecting the cost of renewable energy: the availability of renewable resources and its climatic conditions, and the state of the markets in the region. On the one hand, the cost of electricity by any renewable technology is subject to the weather and climatic conditions of the location (where higher renewable resources will have a huge impact on the overall costs of the installation). Regarding this issue, it could be argued that the best locations for renewable energy installations are already in use in some regions, and probably it is true. However, there are still large areas with excellent renewable resources not being used, and also, repowering of old plants with advanced technology is also on the roadmaps for the next few years, once the older plants fulfill their life cycle.

On the other hand, the markets for each technology are different depending on the region of the world. There are countries with well-established markets (EU, USA, China…) that are more competitive than others (developing countries for example). However, there are multiple factors influencing the final cost of the electricity beyond the competitiveness of the market or the quality of supply chains, and for example, India became in 2019 the country with the lowest TIC for solar PV with an extraordinary cost of USD 618/kW [10]. It is possible that proximity to China (the main global supplier in the PV sector) and low salaries are the reasons behind this number.

On a side note, the year 2020 has been a difficult year for the whole world. The COVID-19 pandemic has shaken the foundations of the contemporary, completely interconnected and globalized society, triggering a disruptive situation worldwide that has put a lot of pressure on the economy of the majority of countries. While the economic effects of the pandemic are still to be determined, it is foreseen a big economic shock

4

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

comparable to that of the Influenza pandemic in 1918 or the financial crisis of 2008. Despite that, preliminary reports indicate that investments in renewable energy have not suffered that much in comparison with other sectors. In fact, some entities are considering this awful situation as an opportunity to build new social and economic dynamics, centering the efforts in a sustainable development that spread widely through all sectors, including economic, social, politic and environmental. In this context, renewable and sustainable energies may play an essential role in the recovery of the pandemic and building the foundations of a new future [13].

In any case, the transition to more sustainable energy systems is tied to the use of renewable energy, modern clean energy vectors (such as hydrogen) and other carbon neutral technologies. There are several reasons guiding the world to an energy change: environmental concerns, security of energy supply, resilience of energy and electric systems, economic benefits, etc. Although the energy transition has begun, we are still far from the objectives required even for the most conservative scenarios to prevent global warming. There is still a long way to travel before we achieve sustainable energy production on a global scale, but the expectations are positive, and this is a travel that the whole world must do together.

1.2   State of the Energy Transition

Once the principal causes of the energy transition have been presented and reviewed in the previous section, the current state of the transition at different levels and from different points of view is analyzed in this section.

1.2.1   Global situation

As of the beginning of 2021, there is still a large path ahead to reach the goals for a more sustainable world and to achieve the transition of energy systems, described in the previous section. As of 2019 (last data available [14]), global primary energy consumption by source looked as shown in Figure 2. A quick glance evidences that global primary energy consumption is still heavily dominated by fossil fuels, with oil being the most consumed energy source (which is mainly explained by the use of oil products as combustible for transportation) followed by coal and natural gas. Besides hydropower, it is extremely difficult to distinguish other renewable sources in the graph. Wind and solar energy combined only add up to 3.29% of the total primary energy consumed. Percentage-wise (bottom of the figure), fossil-fuels have a share of 84.32% of all the primary energy

consumed in the world. Nuclear[1] and renewables complete the picture with 4.27% and 11.41% shares respectively.

Figure 2. *Global primary energy consumption by source. Percentages in the legends refer to the last year in the graph (2019) [14].*

Regarding the final energy consumption by sectors (Figure 3), the transportation sector has the largest share of final energy consumption with 29.09%. Other industries and residential consumptions follow in the ranking with 28.57% and 21.22% respectively. This lines up with the fact that the two largest shares of final consumption by source are oil products (40.64 %) and electricity (19.31%). The renewable share in Figure 3 does not include renewable energy contributions to electricity generation.

Figure 3. *Global total final energy consumption by sector (left) and by source (right). Percentages in the legends refer to the last year in the graph (2019) [14].*

---

[1] Nuclear is listed separately since it is a near zero-carbon energy source.

6

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1: Introduction

Regarding electricity generation (Figure 4), the picture is quite similar, with fossil fuels almost completely dominating the scene with 63.31% of the share. In this case, coal and not oil, is the predominant energy source for electricity generation. The share of renewable energy is considerably higher (26.24%), but this is not a surprise considering that the main use of renewable sources is electricity production. Nuclear energy completes the picture with a 10.44% of the total electricity generated.

Figure 4. *Global electricity generation by source. Percentages in the legends refer to the last year in the graph (2019) [15].*

In Figure 5, global CO_2 and GHG emissions are depicted, along with the annual CO_2 percentage change. It can be observed that global emissions consistently increase, and the growth rate remains almost unchanged. Slight reductions observed are linked to global crisis events such as the oil crisis in 1979 or the financial crisis of the early 90's and 2007. It is expected that after the COVID-19 crisis, a slight reduction in the GHG emissions will be experienced, but it could be mostly circumstantial, and the recovery period that will follow, will probably bring GHG emissions increases if production and energy systems remain unchanged.

7
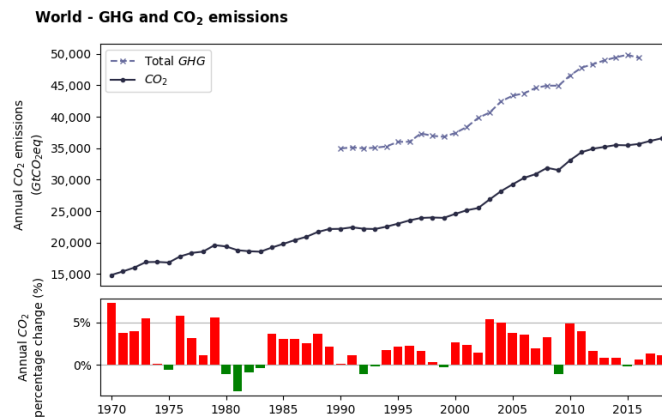
Figure 5. *Global GHG and CO_2 emissions (top) and annual CO_2 percentage change (bottom) [16].*

1.2.2 Spain

In Spain, the tale is almost the same. We are still far from the objectives agreed in the Paris Agreement in terms of reduction of GHG emissions and energy transition. These objectives will only be achieved if our energy systems are completely transformed. However, the current state and rate of change of our energy systems is not high enough to reach the targets set from that agreement. Figure 6 shows the primary energy consumption in Spain by source, in which a picture heavily dominated by fossil fuels can be observed. In the last years, Spain has been transitioning away from coal uses, which can be evidenced by the 3.67% of the total primary energy accounting for that fossil fuel, the lowest in the whole historical series.
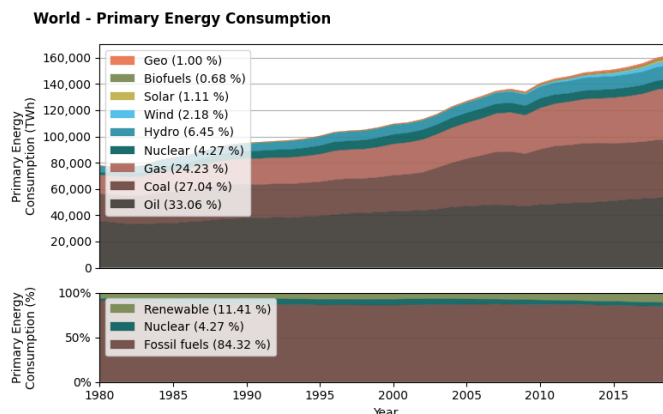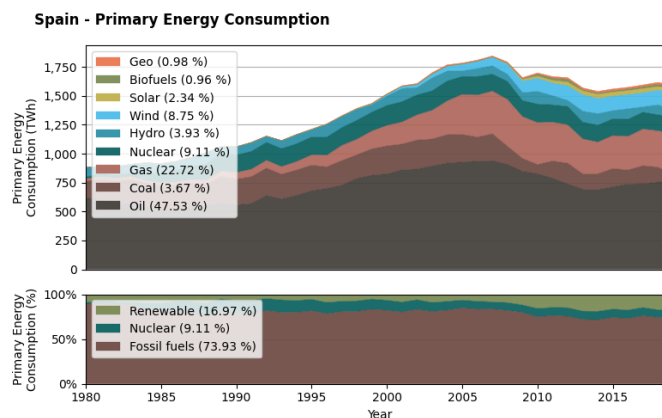

Figure 6. *Spain primary energy consumption by source. Percentages in the legends refer to the last year in the graph (2019) [14].*

8

During the year 2020, the Spanish government presented two national plans which come to address the measures needed to be adopted in the next decade to accomplish the objectives and pledges of the Paris Agreement. These plans were the Integrated National Plan of Energy and Climate (*Plan Nacional Integrado de Energía y Clima*, PNIEC [17]) and the National Plan of Adaptation to Climate Change (*Plan Nacional de Adaptación al Cambio Climático*, PNACC [18]). The PNIEC fixed a set of goals to be achieved during the next decade (2021-2030), from which, the most relevant are the following:

– a 23% GHG emissions reduction from the levels of 1990 (commitment of the Paris Agreement)

– 42% of final energy use coming from renewable sources

– 39.5% of energy efficiency increase

– 74% of renewable sources in electricity generation

With these objectives in mind, the following question can be asked: how is the current situation in Spain, and how big is the effort Spain has to make in order to fulfill these objectives? As of 2018, the share of final energy coming from renewable energy sources is anecdotal (0.4%, without including electricity generation), although the main contributions are made in the electricity sector. Figure 7 shows the final energy consumption by sector and source in Spain. Figure 8 shows the electricity generation by source in Spain in the last decades. Electricity generation coming from renewables in 2019 was 37.54%. This means that the share of final energy consumption coming from renewables adds up to about 9.3%, quite far from the 42% objective for the year 2031.
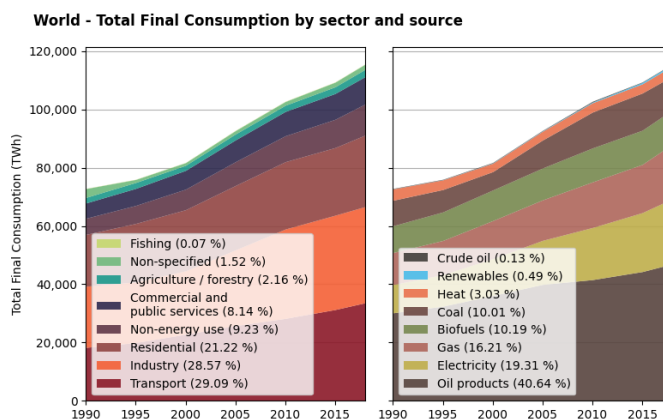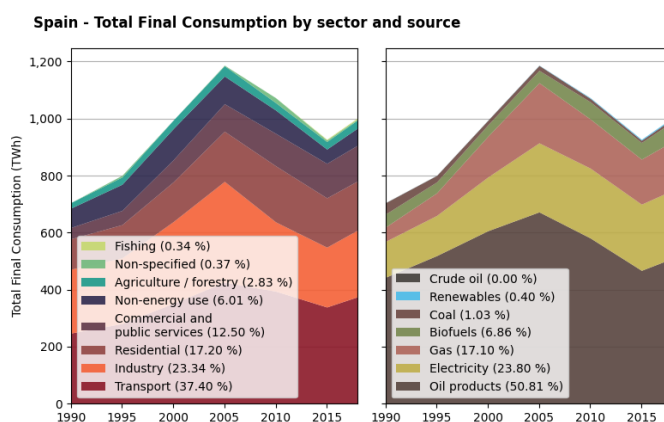

Figure 7. *Spain total final energy consumption by sector (left) and by source (right). Percentages in the legends refer to the last year in the graph (2019) [15].*
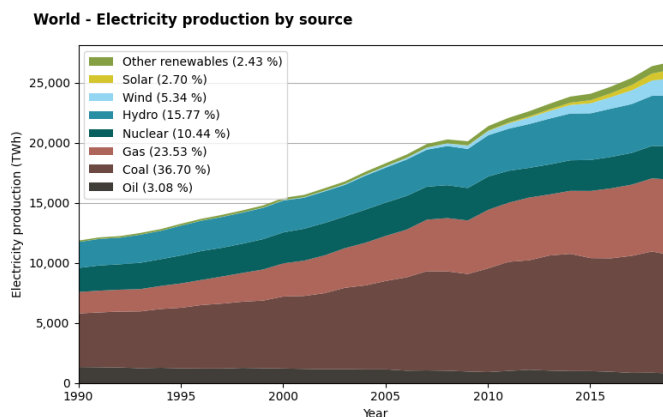
9

Figure 8. *Spain electricity generation by source. Percentages in the legends refer to the last year in the graph (2019) [15].*

Regarding GHG emissions, the effects of the 2007 financial crisis are evident in Figure 9, after which, deceleration of the economy led to several consecutive years of emission reductions. There is a compound effect of this event with the growth of renewable energy in the electricity sector, but it is difficult to distinguish the effective influence of each situation. Nevertheless, the trend is positive if compared with global emissions. Introduction of new renewable energy, electrification of transport sector, carbon capture technologies and introduction of new green fuels will be determinant in the next years to reach the pledges of the Paris Agreement, which are indicated by the green area in the figure.
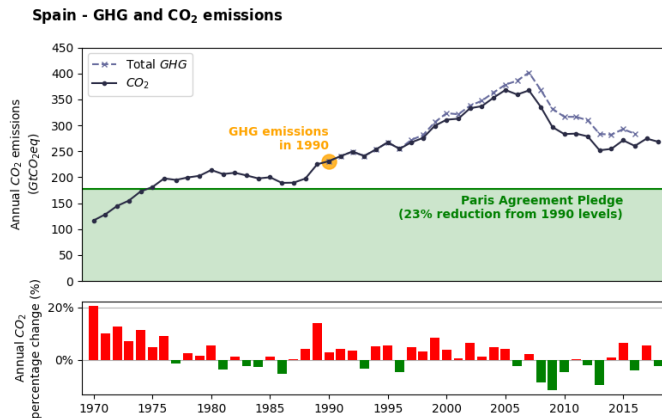
Figure 9. *Global GHG and CO2 emissions (top) and annual CO2 percentage change (bottom) [16].*

10

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1: Introduction

Considering the current situation, these objectives are still far from accomplishment. In order to fulfill all the objectives, Spain faces a series of important challenges in the next decade:

– Renewable energy electricity generation must increase 97.1% in the next decade, from the levels of 2019 (37.54%) in order to achieve the objective of the PNIEC (74%).

– Total final energy from renewable sources should increase by a factor of 4.5 (from 9.3 to 42%)

– GHG emissions should decrease 37% to reach the 23% reduction from 1990 emission levels.

Spain confronts huge challenges in the next few years, in which national strategies to advance in the energy transition and recovery actions from the COVID crisis must be balanced and intertwined. The latent economic crisis expected after the pandemic will not contribute to reach the objectives either, but the government should link the measures adopted for the recovery of the health crisis to the sustainable development and climate emergency situation.

1.2.3 Canary Islands

In the Canary Islands, there are specific concerns on the population, moved by the particularities of the archipelago (with a large number of protected natural areas and biosphere reserves) and it could be argued that the population is more aware of the necessity of an energy transition than in other territories of the country. Proof of these are the massive popular mobilizations that took place against the oil explorations on the archipelago in 2014 [19]. Despite that, final energy consumption is clearly dominated by fossil-fuels, and the share of renewable energy is insignificant, accounting only for a 2.1%, which mostly comes from electric generation, as it can be seen in Figure 10. The main reason for this low contribution of renewable energy to the final energy mix is that most of the energy consumed is done in the transport sector, which supposes a 75.14% of the final energy consumption in the Canary Islands (including air, maritime and land transportation). Almost all of the energy demanded by this sector is supplied by fossil fuels.

11
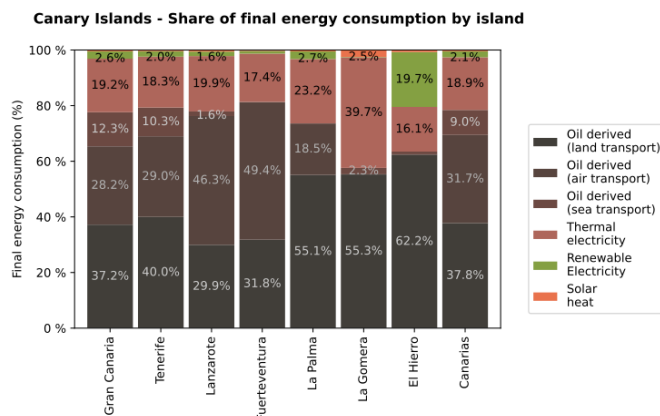
Figure 10. *Share of final energy consumption by island and source (2018) [20].*

It is expected that during the next decades, the electrification of vehicles and the introduction of new green fuels (such as hydrogen) will lead to a decrease in the fossil fuel consumption associated to transportation, but the challenge for the energy transition in this sector is huge, and will remain dominated by hydrocarbons for the next decades if current trends in electric vehicle adoption do not change abruptly.
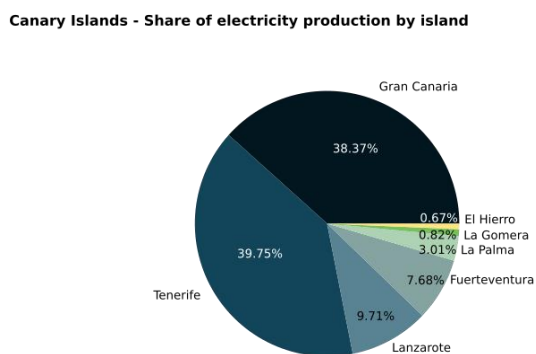
Figure 11. *Share of electricity production by island [21]*

Considering the renewable shares in the final energy mix, the only island achieving significant figures in the use of renewable energy is the island of El Hierro, which is mainly due to the wind-hydro plant for electricity production. However, there is a demographic explanation for this large difference. In the Canary Islands, there is a huge imbalance among the populations and therefore, energy consumptions, of each island. In the Archipelago, there are two islands (Tenerife and Gran Canaria) considerably larger and

12

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1: Introduction

more populated than the rest of them. Tenerife and Gran Canaria combined accounts for 82% of the total population and 78% of the electricity consumption, which can be observed in Figure 11. El Hierro only has 11,147 inhabitants (0.5% of the total), and the recent installation of the 11 MW wind-hydro plant has put the island leading the renewable transition in the islands.
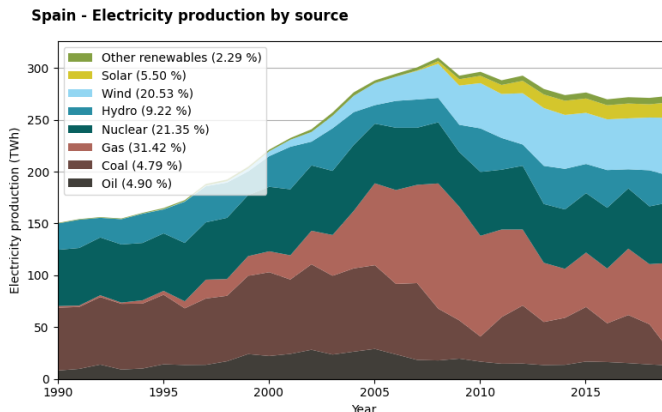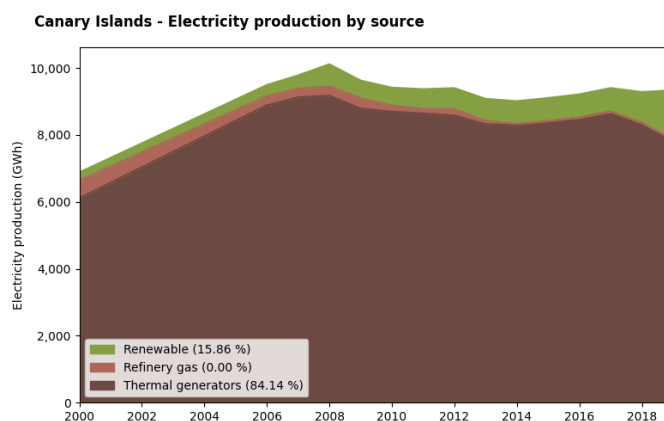

Figure 12. Canary Islands electricity generation by source. Percentages in the legends refer to the last year in the graph (2019) [21].

As it can be seen in Figure 12, electricity generation in the Canary Islands is mainly dominated by fossil fuel-fired thermal generators, using mainly fuel-oil and diesel as fuels due to the lack of natural gas and coal supplies in the Archipelago. During the last decade, the idea of including natural gas in the energy mix has been hovering over the head of politicians, with the projection of a regasification station on Tenerife Island, which would bring a less polluting fuel for producing electricity. But the idea seemed to fade out for several reasons. A relative rejection of the population, the decrease in costs of other renewable energy alternatives and the future projections of cleaner energy vectors such as hydrogen that could be included in the future energy mixes have highlighted the necessity of thinking a different strategy of transitioning away from fossil fuels. Nonetheless, natural gas may still play an important role during the first stages of the energy transition in the Canary Islands, either by displacing the consumption of more pollutant fuels (such as fuel oil or diesel oil) or by mixing it with renewable hydrogen to fuel the thermal generators.

Renewable energy share has grown steadily during the last years, adding up to 15.86% of the total electricity generated in the year 2019. The renewable capacity in the Archipelago has also grown during the last decade, following global trends, with some discontinuities due to regional and national regulations preventing the deployment of

13

more renewable installations[2]. Figure 13 shows the share of the different technologies involved in electricity generation by island. The island of El Hierro, with its wind-hydro power plant, is the only one capable of achieving substantial figures in renewable energy, with a 66.7% of the total annual electricity produced coming from renewable sources in the year 2019 [21]. The largest shares of renewable electricity are coming from wind power, which lines up with the fact that wind capacity is more than double the solar PV capacity in the archipelago. Renewable energy generation in Tenerife adds up to 18.4% and Gran Canaria is 15.4%.
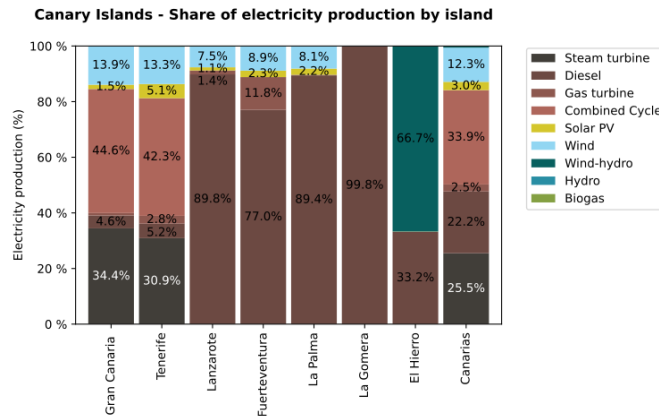
Figure 13. *Canary Islands electricity generation by island and by source (2019) [21].*

1.3   Technical Considerations of Electric Systems

From all forms of energy consumption, electricity is the most convenient for its ease and safety of use, quick availability and relative low cost. However, electric systems are among the most complex engineering systems, and therefore, transitioning from a fossil fuel-based electricity production to a renewable energy-based one is not a task exempt of difficulties and barriers. There is a very important difference between both approaches: contrary to the "on demand" availability of fossil fuels, renewable energy production is only possible when there is renewable resource available if no energy storage is considered. This is particularly true for the two major technologies dominating the scene in the last years: wind and solar power. This fact increases the variability of these energy sources, and that is why they are called Variable Renewable Energy (VRE). VRE sources (also called Non-Manageable Renewable Energy), such as solar and wind power, pose new

---

[2] A more detailed analysis of the impacts of policies is given in Section 2.1.5

14

difficulties in an already complex system, as the electric is, and operating the electric grid becomes a more challenging task.

As a simple graphic analogy, one could think of the electric systems as an old weighing scale, with the generation on one side and the demand on the other. In order to achieve static equilibrium, both plates should have equal "*masses*". Electric systems behave similarly. They must maintain a real-time and delicate equilibrium between demand and generation at all times in order to control the frequency of the system, and therefore ensure safety, quality and security of supply.

Continuing with the analogy, load would be the unknown "*mass*" we are trying to measure with generation units, of which we know their properties, and we use as reference. Traditionally, the main sources of uncertainty and variability have been associated with the demand side (residential customers and industries make use of the electricity whenever they want, and this has a somehow stochastic nature), and therefore, forecasts have been made to reduce the effect of this uncertainty on the electric system. Nowadays, when aggregated, demand or load curves for determined electric systems can be confidently predicted taking into account exogenous variables such as the time and day of the year, week-days or not, temperature, etc. Load forecasting is a problem that has been around over a century [22], and it has been tackled with multiple approaches for decades. In the aggregated case, it may be considered as a solved one.

However, with the introduction of VRE sources in the generation side, the electric system has a new source of variability and uncertainty in its components. That is (and finishing with the analogy), we are trying to measure the load "*mass*" without clear references, because our reference weights may change at any moment, generating high uncertainty. And, in the electric system, these uncertainties are considerably worrying, since they affect the overall generation-load equilibrium, meaning that a miscalculation on the possible outputs could lead to catastrophic effects in the system. If VRE generation has an important share in an electric system, measures should be taken to reduce the uncertainty and ensure a safe operation.

At this point, it is worth defining the terms variability and uncertainty. In electric systems, basically all the components, such as loads, power lines, and generator availability and performance; have a degree of variability and uncertainty [23]. In a general sense, variability accounts for the changes in a determined time-series over time; while uncertainty refers to the lack of accuracy in the predictions of the changes [24].

Solar PV generation for example, is a very low inertia energy source, meaning that changes in the output are fast (with no conservation of energy momentum) and the changes in outputs may be in the range of the milliseconds for single PV cells, and close to the minute for PV plants depending on its size. That is, solar PV has high variability. Thinking ahead, in scenarios with 100% renewable energy, if for example, a heavily solar

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
            UNIVERSIDAD DE LA LAGUNA

            María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
            UNIVERSIDAD DE LA LAGUNA

PV-based system confronts a partly cloudy day, the variations in solar output could be quite important and fast, leading to deviations in the frequency higher than supported by the system, resulting in malfunctions of the system itself that could be followed by disconnection of some generation groups and ultimately, to complete blackouts.

Even with low shares of PV in relatively weak electric grids, the impacts of renewable variability are starting to be noticeable. For instance, Figure 14 shows the frequency deviations in the electric system in Tenerife (Canary Islands) due to the renewable generation. During this event, the PV contribution was around 15% of the total electricity produced. It can be observed how the ramps in PV generation (red line) have a major influence on the frequency of the system (green line, right axis) with a positive correlation. Thus if such frequency excursions occur with low renewable shares, it is likely that these kind of events would be more frequent as renewable energies become more dominant in the electric mix.
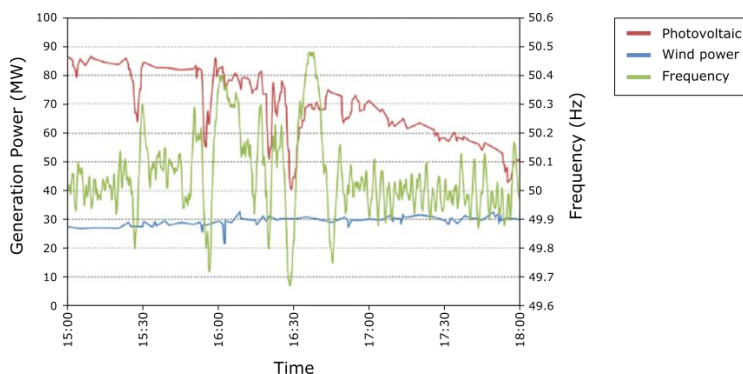
Figure 14. Frequency fluctuations due to a photovoltaic ramp in the island of Tenerife in 2017 (own elaboration).

Although the variability and uncertainty of VRE generation increase the response requirements from conventional generators, they do not increase overall capacity requirements [25]. That means that peak load with VRE is never higher than peak loads without them. However, VRE does impact the system operation in terms of regulation, load following and unit scheduling.

Along with the variability and uncertainty of VRE, there is another factor related to them that heavily impacts the operation of the electric grid. Modularity of some VRE technologies (particularly solar PV) makes it possible to install them in a wide range of configurations, with capacities ranging from few kW to several MW. They are easy to operate and safe to install on the rooftops of any house with minimal maintenance requirements and costs. The electric grid was originally conceived to be operated radially (where power flows uni-directionally from large generators to the end-users), but wide

16

spreading of small scale VRE connected to distribution grids have completely changed the operating conditions of the electric system. The small-scale installations on distribution grids are considered to work as Distributed Generation (DG), and it has shifted the operation paradigm from a centralized to a decentralized approach, with generators of many sizes on every level of the grid, and where power flows in lines are multi-directional [26].

Some of the effects and impacts that variability in the generation due to VRE may have in the electric grid (both from a technical and economic point of view) are voltage instability, frequency instability, bidirectional power flows or increased integration costs[3].The severity of these impacts will depend on the time scale and the voltage level of the grid at which the events are produced, as well as the robustness (or weakness) of the grid itself. In this sense, robustness includes all the characteristics of the grid that help support its stability, such as size, number of high inertia generators, interconnections, etc.

In order to mitigate these impacts, there is a great deal of strategies that could be adopted, such as forecasting the output of renewable energies, using batteries to store surplus energy from VRE and serve it when it is needed, including other devices (supercapacitors, flywheels) to stabilize the grid, using electrolyzers as variable loads to synthetize hydrogen with surplus of renewable energy or using advanced control strategies for the electronic devices. While in the coming years new technologies will be added to the system to cope with the variability of VRE, forecasting of renewable energy (and even other parameters of the grid) will remain to be of great interest, since it deals with the other part of the problem, that is, the uncertainty, and minimizing the uncertainty of the systems could lead to economic benefits in terms of optimized operation of power plants or extending the useful life of some devices.

The use of advanced control strategies for power electronic equipment, such as solar PV inverters, is another attractive option to mitigate the impacts, particularly voltage related problems at the distribution level. The use of reactive power compensation or limiting the amount of active power injected to the feeder also enables the penetration of higher shares of renewable sources in distribution grids. Additionally, this measure is considerably cheaper than other alternatives, since no investment in equipment is needed, and it only requires the configuration of the inverters.

Besides the improvement in control strategies once the events have occurred, any attempt to reduce the uncertainty in every part of the electric systems is welcomed. That is where solar forecasting (and any kind of forecasting actually) can play an important role. As mentioned in previous paragraphs, there are forecasting problems that may be considered as solved (aggregated load, for instance), at least at some time horizons. But

---

[3] The impacts of PV systems on electric grids are further discussed in Section 2.3.

when one unravels the data and starts working with non-aggregated time series, the problem becomes a non-trivial one and a very tough task. This is particularly true for solar and wind forecasting, since their stochastic behavior makes them difficult to predict. For that reason, improvements in the technology and methods are required in order to provide effective and useful services to the electric system. Some of the effective applications of forecasting in the electric systems include the following[4]:

– Ensuring grid stability and quality of electric supply

– Optimization of grid parameters

– Optimization of system costs

– Extending the useful life of certain devices (on-load tap changers, batteries…)

– Optimization of market interactions

In conclusion, although the inclusion of other technologies will also be required in the future to achieve 100% renewable energy electric systems, VRE forecasting is a relatively inexpensive tool that will help to cope with the variability, and if integrated in the system operation, will lead to improved performance of the grid, minimized operational costs, increased safety and security of supply and safer operation of electric systems.

1.4 Motivation of the thesis

As the deployment of PV systems and other VREs (mainly wind power) increases in the Canary Islands, the challenges derived from the operation of the electric grid increase as well. With the new renewable and non-manageable generators added to the electric grid, the variability and uncertainty usually associated with the demand, are now extended to the generation. As the electric system works in a sense as a weighing scale, where generation and load should be balanced at all times to ensure the security and quality of the electric grid, sudden changes experienced by non-manageable renewable sources lead to imbalances in the equilibrium, resulting in hazardous operation conditions. The nature and severity of these hazards will depend on several factors, such as the voltage level of the grid, its robustness, the presence of other appliances, the geographical dispersion of the renewable energy plants.... The most common and worrisome impacts are probably those related with frequency deviations at a grid level and voltage violations in distribution voltage levels.

The effects of the variability will also be different depending on the relative size of the electric system. In continental well-interconnected grids, the effects of the disturbances are diluted on the larger size and increased robustness of the systems, which have

---

[4] For a more detailed description of the different forecasting methods and applications, the reader is referred to Section 2.2.

18

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

multiple large generators scattered throughout the geography, and the relative importance of disturbances is smaller, when compared to small isolated electric systems (such as the ones found in Canary Islands).

The uncertainty derived from the utilization of the non-manageable renewable sources should be dealt somehow, in order to warrant the quality and security of the electric supply. There is a vast plethora of strategies, including electricity storage in a large number of ways (different technologies with different uses and applications), advanced control mechanisms and strategies included with the new inverters and other power electronic equipment, power curtailment or limitation of the amount of energy dumped to the grid, forecasting…

The main reason behind the election of one alternative over the others will be, as it is usual in business, the cost of the option and the relative costs compared with the other alternatives. Solar forecasting has been identified as a key task in order to increase the host capacity of PV systems in electric grids, following the line of thought that, if we are able to confidently predict the sudden disturbances in the generation, we could adapt other parts of the electric system to respond efficiently to these changes and variations, and the security of the electric system would be assured.

As stated in previous paragraphs, the main barrier preventing the deployment of a technology is its costs in relation with the other alternatives. For that reason, one of the prime movers of this thesis is the development of a low-cost intra-hour solar forecasting system that is capable of doing confident predictions that could be used to ensure grid's security and reliability.

As mentioned before, the effects of the variability are different depending on the voltage level of the grid they are produced. Voltage violations are a common problem at the distribution level (medium and low voltage parts of the system). The increased resistivity of the cables at that voltage levels makes the disturbances in active power injections have a major influence on the voltage magnitudes. For that reason, if a distribution grid has a large PV penetration (for example, neighborhoods where most of the houses have a PV system connected to the grid), it is more likely that these situations arise more frequently, and problems such as under- or over-voltages in the lines occur. These situations need to be avoided since they affect the quality of the electricity delivered to the customers, and could lead to unsafe situations. Advanced control strategies for voltage regulation in PV inverters is an interesting approach to reduce the effects derived from a high PV penetration in distribution grids.

At the transmission level (high voltage systems), the variations of the active power do not lead to voltage deviations, since the resistivity of the lines is far smaller than in distribution grids. However, the most frequent problem, particularly in small or isolated grids, is the shift in frequency values from the nominal, as the imbalance between load and

generation occurs when sudden and rapid changes in generation take place. Large disturbances such as generator tripping or load outages can initiate cascading outages, system separation into islands and even the complete blackout [27]. Solar forecasting is particularly useful since PV solar plants are low-inertia systems (at least in the traditional ways of use, which is foreseen to change in the near future when PV plants will provide the so-called artificial or synthetic inertia), and the changes in output power are practically instantaneous [28]. Therefore, if accurate predictions are available, the system operator can adapt the rest of the elements of the grid to cope with the ramps that may occur, maintaining the voltage and frequency stability through the grid.

The combination of solar forecasting methods with advanced control strategies for solar inverters will also lead to more versatile electric systems with the ability of: (i) predict when solar variability and solar ramps will happen, trying to anticipate the possible outcomes and adapting the rest of the electric grid consequently and (ii) achieve an effective mitigation of voltage related impacts, especially in distribution grids. Small island systems will benefit the most from early adoption of these technologies, since their relative weakness and lack of robustness make them particularly vulnerable to the effects of the solar variability.

1.5 Hypothesis and objectives

Given the global energy outlook, regional background of Spain and the Canary Islands and prime movers and motivations for the research and the thesis, the initial hypothesis can be settled. In a general sense, the fundamental hypothesis can be summarized in the following question:

Can we integrate more PV installations in weak, islanded electric grids if we are able to predict the very near future outputs and control the behavior of the PV inverters?

If the hypothesis is broken down to a lower level, and deeper understanding of all the topics are considered, the following questions arise:

a. *How optimal and well managed is the current fleet of PV plants in the Canary Islands?*

b. *How future PV plants will perform in a determined location under specific conditions?*

c. *How large and sudden changes in PV output due to passing clouds affect the stability of electric grids?*

d. *How good an intra-hour PV forecasting can be in terms of error metrics? How state-of-art techniques for intra-hour solar forecasting can be improved?*

20

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 1: Introduction

    e. *What is the best technique to perform intra-hour PV forecasting, and specifically, for quick ramp detection?*

    f. *What is the true economic value of forecasting PV (and any renewable source)? How big is the added value of performing forecasts?*

    g. *How can we use the power electronics available on PV inverters to address the variability of solar energy and control the voltage and frequency of the grid to maintain stability?*

To answer these questions, a series of initial objectives laid out at the beginning of the thesis.

- The design and construction of a low-cost, image-based, intra-hour solar forecasting system capable of reducing, or at least matching, the error rates encountered in the literature.

- Test the forecasting system in a real work environment, with data provided by utilities in the framework of the projects compounding the thesis.

- Integration studies of PV systems in electric grids, adding the information obtained with the forecasting system, and proposing a series of measures and evaluating their suitability.

These initial objectives were clear and concise, but, as often happen, new insights obtained throughout the development of the thesis made the course deviate a little from the original path, and new objectives arose. Those objectives were updated yearly in the research plan.

- Make an evaluation of the current PV power plants in the Canary Islands and generate a model to estimate their specific yield (kWh/kW)

- Study the installation rates of PV systems in the Canary Islands and evaluate the impact of the different policies implemented in the last decades over them.

- Integration studies of PV systems with a particular focus on low- and mid-voltage distribution grids. Propose control methods for PV inverters that take into account information provided by the forecast system.

- Economic and environmental analysis of the impact of a good forecast over the electric system.

1.6 Relation of the papers with the content of the thesis

This thesis is presented as a compendium of three papers, which are listed and briefly described below:

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

1. The first paper "*A simple big data methodology and analysis of the specific yield of all PV power plants in a power system over a long time period*" aims to be an initial evaluation of the PV potential in the Canary Islands. A simple big data-based model to estimate the specific yield of PV plants is proposed, using only data easily available for the utilities, to compare the performance of the PV plants with a theoretical ideal situation.

2. The second paper "*Validation of All-Sky Imager Technology and Solar Irradiance Forecasting at Three Locations: NREL, San Antonio, Texas, and the Canary Islands, Spain*" is a paper in collaboration with researchers of the University of Texas, San Antonio. The paper describes the deployment of the early stages of the technology developed during the research process. Three different prototypes of the Sky Imagers were deployed in three different locations (two in the USA and one in Spain), and the key findings of each prototype and the valuable lessons extracted from each experience are summarized in the article.

3. The third paper "EDA-based optimized global control for PV inverters in distribution grids" deals with part of the thesis dealing with the integration of PV systems in electrical systems. In this paper, an optimization metaheuristic (Estimation of Distribution Algorithm, EDA) is used to control PV smart inverters in a distribution grid, in order to maintain voltage stability in a high PV penetration scenario.

The general reasoning for the research line is the following. First, an evaluation of the PV potential of the Canary Islands is conducted. In this part, the objective was to quantitatively evaluate the configuration and management of the PV plants currently installed in the Canary Islands, developing a simple procedure to estimate the specific yield of already deployed and future PV plants with data readily available to the utilities. With this study, a bulk estimation of how much energy may be produced by PV plants allocated in different parts of the Islands is achieved. Following the increasing global trend of PV deployment, it could be determined the energy produced in the future by the plants just with location data, and few other easily available data that could be modeled or retrieved without exposing the sensible information of the plants or the owners.

Second, once the preliminary analysis has been done, the problem of solar forecasting is tackled. Since our main concern is grid stability (rather than economic and market interactions or long-term planning aspects), intra-hour solar forecasting is the main focus of our research. To this end, a low-cost solar forecasting system based on sky-imagers is designed. Three early prototypes of the sky-imagers were installed in different locations around the world (two in the United States and one in La Graciosa Island), in the framework of a collaboration with the University of Texas, San Antonio (UTSA). The knowledge retrieved from these experiences is summarized in the presented article, and

22

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

the insights gained have helped in the decision-making process for the next steps of our research in solar forecasting (and forecasting in general), and in the last builds of our devices.

Third, as a final step to the integration of more PV capacity in weak electric grids, the use of advanced capabilities of PV inverters is studied. Power electronics in modern PV inverters can be used to adapt the output conditions of the PV power (both reactive and active power) in order to maintain voltage and frequency stability. To approach this problem, an optimized global control for all the PV inverters in a distribution grid is proposed. To obtain the optimal parameters for each PV inverter in the grid, the control is optimized using a metaheuristic called Estimation of Distribution Algorithm (EDA), which is an iterative population based metaheuristic that constructs a distribution function with the best solutions of each step, to sample the solutions for the next one.

More extensive summaries of each work with full explanations can be found in Chapter 3: Research, and the referenced articles can be found annexed at the end of this document.

1.7 Other contributions and Projects

1.7.1 Projects

A big part of the research in this thesis has been developed in the framework of various private research projects with utilities, and collaborations with different entities and universities. This fact has supposed a new approach to understanding the interactions between the universities and the utilities, since the objectives of the research were linked to more "earthly" goals, solving real world problems that the utilities are experiencing or anticipating that they will occur in the near future, rather than facing the problem from a completely academic perspective, sometimes disconnected from the real world. Some of the projects that have been part of the research of this thesis are described below.

– **GRACIOSA**. The GRACIOSA project is probably the most important (from a sentimental point of view) since it was the cornerstone of the forecasting research for our group. This project, led by the Spanish utility Endesa, set the foundations for our forecast technology at the very beginning of the thesis, and it helped us to test our technology in a smart-grid context in La Graciosa island, with several PV systems, new hybrid energy storage technologies and energy management systems. Although the project itself finished in 2018, the agreement was extended between the parts to continue with the tests of our forecasting system, extending and upscaling our technology to a grid with higher PV penetration, in the island of Tenerife, under the DEMFORE project.

23

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

- **DEMFORE** (ongoing). The DEMFORE project (also led by Endesa) was conceived as the extension of the GRACIOSA project to a larger electric system with higher shares of PV and larger PV plants. The objective of the project is to forecast the production of several multi-MW PV plants, and to simulate the interactions of the PV systems variability with the rest of generators in the largest power plant in the island (Granadilla). Two sky-imagers were installed near several multi-MW PV plants in the south of the Island, and advanced techniques of deep learning and data-based forecasting are used to make the predictions. The final aim of this project is to evaluate the quality of the system at predicting big PV ramps.

- **REDUCO2** (ongoing). This can be seen as a side project of the DEMFORE in which the main goal is to produce an optimization routine to reduce GHG emissions from the thermal generation units, taking into account the intra-hour solar forecasts made by the forecasting system.

- **SIMRIS** (ongoing). The SIMRIS project is collaboration between our research group and the German utility E.On, with the objective of integrating our forecast system in a smart-grid in the location of Simris, Sweden. It is still at an early stage of development, but three sky-imagers have been installed in the location, and the major novelty in comparison with the prior projects is that the system will be able to perform 3D reconstruction of clouds.

1.7.2 Contributions to conferences

Along with the papers presented in this document, several preliminary works of this research have been presented in international conferences, and some of the main contributions are listed below.

- **IEEE-PVSC-44** (June 25-30, 2017, Washington, USA): "First Results of a Low Cost All-Sky Imager for Cloud Tracking and Intra-Hour Irradiance Forecasting serving a PV-based Smart Grid in La Graciosa Island", David Cañadillas, Walter Richardson Jr., Sara González, Benjamín González-Díaz, Les E. Shephard and Ricardo Guerrero Lemus.

- **WCPEC-7** (June 10-15 2018, Hawaii, USA): "A low-cost two-camera sky- imager ground-based intra-hour solar forecasting system with cloud base height estimation capabilities working in a smart grid", David Cañadillas, Benjamín González-Díaz, Jacob Rodríguez, Jorge Rodríguez, and Ricardo Guerrero-Lemus.

- **ICREPQ'18** (March 21-23, 2018, Salamanca, Spain): "Reactive power management in photovoltaic installations connected to low voltage grids to avoid active power curtailment". D. Cañadillas-Ramallo, B. González-Díaz, J. F. Gómez-González, and R. Guerrero-Lemus.

24

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

– **ICREPQ'19** (April 10-12, 2019, La Laguna, Spain): "Reactive power management to enhance solar energy penetration in small grids: technical and framework approaches". B. González-Díaz, J.F. Gómez-González, D. Cañadillas-Ramallo, J.A. Méndez-Pérez, R. Guerrero-Lemus.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

26

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
            UNIVERSIDAD DE LA LAGUNA

            María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
            UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 2.   Theoretical Background

In this section, the background theory related to the different topics treated in the papers included in the compendium is reviewed in detail. It aims to be a compact summary of the most important topics, with a wide literature survey and simple explanations to contextualize the techniques and methods employed and described in the papers within the state-of-art of the different areas of research.

2.1   Solar resource estimation and impact of policies

Solar resource assessment and evaluation of PV potential has been a traditional line of research in the energy field for many years. With the irruption of solar energy technologies, more efforts were put to correctly estimate the potential generation capacity of determined regions, mainly for modeling purposes and analysis of economic revenue of PV plants. Nowadays, several approaches have been considered to make such estimations, generally made in a large scale using satellite data, and good PV potential estimations can be easily accessed for free in several sources and databases over the internet. However, these estimations lack in general the granularity required for some regions which may have special characteristics, such as small islands with pronounced orography, which can experience multiple microclimates in short distances. Therefore, more accurate methods in a local scale for estimating solar potential and performance of PV systems are required for those locations.

On a side note, emerging advanced data techniques may be used to tackle this problem from a different perspective. Big data (BD) and artificial intelligence (AI) techniques have gained a notorious relevance in solar energy research in the last decade. Advances in data acquisition devices, databases and related technologies, have granted the availability of large volumes of data to the scientific community, which is beginning to exploit it to its full potential. Among the multiple uses that big data has in the field, the estimation of PV potential or specific yield of PV plants are of great interest in terms of planning future facilities, analyzing performance of currently deployed plants, studying the economic value of installations in a particular region and even programming operation and maintenance requirements. Large amounts of data can also help to understand the real implications of regulation changes in the sector by analyzing the effect of new policies in several control variables, such as PV installation rates and annual generation.

27

In this section, some of the techniques found in the literature for PV potential and specific yield estimations are presented, as well as the use of BD and AI in similar research. An overview of the effects of policies and regulation changes on the PV sector is presented as well.

2.1.1 Estimation of PV specific yield and PV potential

Specific yield of PV plants is an interesting aspect to evaluate in terms of understanding the potential capabilities of producing PV power of a region. Specific yield is defined as the amount of energy that is produced per unit of capacity (in a determined time), and its unit is kWh/kW. Generally, the time considered is a year, so it represents the amount of energy per capacity that a PV plant will have in a year under determined circumstances. Specific yield can be used as an initial estimation of how PV plants will perform in a certain location under specific circumstances, but also to evaluate the performance of already deployed plants, compared to ideal conditions. Moreover, it could be seen as a kind of a bulk regional forecasting method, in the sense that obtains potential PV output estimations of specific regions in the long term.

Modelling PV specific yield is, in a sense, equivalent to model PV potential, or solar irradiation. Although the output of these models could be different (in terms of units, for example), clear relationships between them arise when results are compared. For that reason, in the literature one could find references to PV potential or solar irradiation estimations, and the methods and models are fairly similar to the ones describing specific yields. There are several methods to estimate solar resources, PV potential and specific yields of PV plants. Some of them use Geographic Information Systems (GIS) based approaches [29]–[31], physical methods [32], [33], or use of irradiance databases. The aims of each approach tend to be different as well, ranging from the assessment of the solar resource itself, to selecting the best locations for future PV plants. For an extensive review of PV potential assessment methods, the reader is referred to [34].

Although several studies have examined the PV yield, little to none of them have utilized data from individual plants coming from smart meter records along with advanced big data techniques. Actually, the availability of behind-the-meter data is scarce for researchers, although there are some examples of using this data to similar studies and estimations. For instance, in [35], data from smart meters is employed to feed a Deep Learning (DL) model to estimate the size of the PV systems, achieving errors of around 2%.

28

2.1.2 Solar databases and other available resources

A decade ago, it was difficult to find reliable irradiance and irradiation data with an acceptable resolution. Nowadays, there is a large number of solar databases that provide high quality data, covering the whole globe and even providing additional services such as PV potential estimation or specific yield. The irradiance and irradiation from these solar databases is estimated with satellite data, and although it generally will not be as accurate as a local or regional study, they provide a good initial estimation for many applications that may use them, such as the planning of PV systems. Some of the most utilized and complete databases provided by international and governmental agencies that one can find on the internet nowadays are listed below:

- Photovoltaic Geographical Information System (PVGIS) [36], from the European Commission, is one of the most utilized databases. Originally it contained information about Africa and Europe, but now it covers almost the whole Earth surface. It has five datasets that could be selected. The main differences are that they have better performance for determined areas or latitudes, and some of them have reanalysis data. The datasets are SARAH, NSRDB, CMSAF, ERA5 and COSMO. The details of the implementation can be consulted in [37].

- Renewable Energy Atlas (RE Atlas) [38], by the National Renewable Energy Laboratory (NREL), provides monthly averages of daily total solar resource, on a grid of 40 km by 40 km. It only covers the USA.

- Global Solar Atlas [39], from the World Bank Group, provides a summary of solar power potential and solar resources globally for almost 150 non-OECD countries and regions.

- NASA provides several web applications to consult irradiance data and even renewable energy potentials, for instance the Prediction Of Worldwide Energy Resources (POWER) project or the net Radiation Maps [40], [41].

- SOLEMI [42] from the German Aerospace Center.

Several private companies and universities also provide solar resource data from their own analysis of satellite data. Some of them are Meteonorm from Meteotest [43], SoDa (Solar radiation data) [44], SolarGIS [45], or SOLCAST [46].

2.1.3 Roof-mounted and ground-based PV plants

Once the PV potential or specific yield of PV plants for a particular region is determined, one could start modeling the PV plants in ideal conditions. However, there are other factors which prevent reaching the ideal conditions, in which PV plants are placed in the best locations with high irradiation values at the optimum tilt.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

2.1 Solar resource estimation and impact of policies

This is especially true in the case of the Canary Islands. It is known that they possess excellent solar resource and irradiation values, and they should provide good PV yields. But if we take into account other aspects, the analysis is not that straightforward and becomes a little more complex. Land use is a serious concern in the Canary Islands archipelago, and it is a common issue in similar locations and islands. In small size islands, land is a limited resource, and a high competition for it is encountered. Land may be used for other purposes (such as agriculture) or it may fall inside protected areas (natural or national parks). Moreover, large PV systems can extend in a large region depending on their capacity and configuration. This limitation in land use often leads to plants located in places with slightly less solar resource and could also lead to plant configurations with suboptimal parameters, such as the tilt of the PV modules.

An interesting alternative to limit high land use of large PV systems is to install smaller PV systems in the rooftops of the buildings or industrial areas. With decreasing overall costs for PV technology and adequate policies to stimulate their deployment, rooftops installations are expected to grow in the coming years, both in Spain and in the Canary Islands. From a technical point of view, the shift to a DG paradigm will translate into more complex distribution grids and a more challenging operation of electric grids[5], as briefly discussed in Section 1.3. But from an environmental and land use point of view, it definitely is an attractive alternative.

However, there are some drawbacks to the installation of PV on the roofs in terms of purely energetic performance (without regards to the impacts caused to the electric grid). Generally, the disposition of the roofs is not optimized to allocate PV modules. The number of modules that can be allocated, and therefore the capacity of the plant, are both limited by the dimensions of the roofs. In other cases, aesthetic aspects are considered when PV is integrated into the buildings. Also the influence of shadows is usually higher in roof mounted installation, especially if there are other buildings or trees nearby.

Probably the best case for roof-mounted installations would be in the industrial areas, where the aesthetic considerations are not that important, and the dimensions of the roofs are larger, enabling the allocation of larger plants. Moreover, labor hours tend to coincide with higher irradiance hours, and self-consumption schemes are more attractive.

On the other hand, ground-based PV plants have several advantages over roof mounted plants. Some of the most important are listed below:

– Higher capacities (utility scale), that can add up to several MW.

– They are optimized. Besides the land use constraints, plant design should ensure optimum tilt and orientation, minimum shadows and minimum losses.

[5] An extensive description of the technical problems associated with DG is given in Section 2.3.

30

- Decreased generation costs, benefiting from economies of scale (the larger the plant, the lower the cost of the unit of electricity).

- Better management, in terms of scheduled operation and maintenance.

Although each alternative has its advantages and drawbacks, it is expected that a combination of both approaches will be needed to reach the goal of decarbonizing the electric systems, particularly in areas where there are land use limitations.

2.1.4 Big data and Artificial Intelligence

During the last years, the term Big Data (BD) has gained more and more prominence in both the scientific community and the enterprise world. Yet there is no clear consensus about the real meaning of the term, despite being broadly used on a daily basis [47]. In the modern world, a ridiculously large amount of data is collected every second: sensors, cameras, power meters, electronic devices, social networks or web applications are continuously gathering huge amounts of information and, most of the time, this data is collected without any specific future purpose. The term big data came to explain this situation, in which the quantity of information is so abundant that its analysis is a task beyond human capabilities in most cases. It was not until recent developments in computer science that the algorithms and models needed to treat big data were available, and therefore, the recent popularity of the term is explained by that.

Some authors have attempted to define big data in terms of the dimensions a data set should have to be considered as such. Volume, Variety and Velocity (the so called "Three V's"), seems to be the defining characteristic of BD. Although the definition of the Three V's could be extended with extra dimensions: Veracity, Variability and Value [47]. In any case, it is a fact that there is no universal benchmark to each of the dimensions to conclude what is BD and what is not. Data structure is highly sector-dependent, and therefore, it could be the case that definitions for each sector would be more appropriate than a universal one. Nevertheless, in the context of this research we define BD as large volume datasets, which seems to be in line with most of the uses in the literature in our field [48].

The use of big data cannot be explained without mentioning its functional counterpart, that is, artificial intelligence (AI). AI techniques include a vast set of methods to exploit big data to its greatest extent, including machine learning (ML) and deep learning (DL) techniques among others (computer vision, data mining, etc.). ML and DL techniques are probably the most interesting from a functional point of view, since they are not confined to specific uses, and in general are problem agnostic, meaning that they try to model the data by its structure, leaving little to no room for researchers' biases.

BD and AI techniques are extensively used in recent research in many fields. Regarding energy and electric grids, a simple search including "energy" and "big data" or

31

"artificial intelligence" in any scientific database, will lead to thousands of results. This by itself could be the content of an entire thesis. Since the objective of this thesis is not to fully describe the use of these techniques on the solar energy field, the curious reader can found extensive reviews and state-of-art reports of BD techniques in smart- and micro-grids in [49]–[51]. Literature reviews on the use of BD and AI techniques in solar energy and forecasting applications are presented in [48], [52]. In [53], a wider review including more uses of big data in energy is presented.

2.1.5 The effect of policies and regulations in PV installation rates

Another interesting factor to consider is how the different regulations and cost trends have affected the installation rates during the last years. It is known that regulatory policies have a large impact on the deployment of a technology. Incentives, feed-in tariffs (FiT), and other economic mechanisms could be used to help the development of a new technology. During the last decade, we have assisted an unprecedented change in the way of thinking about electricity generation, as well as an impressive cost reduction in renewable energy technologies. It is possible that the correlation between the two facts is not negligible. The economic incentives made the investments in the sector attractive, and therefore, more economic resources were put into it. With more economic resources, there is an increase in the demand for new installations. And with this increase in the demand, more experience is accumulated by the manufacturers and therefore, the competitiveness of the industry also increases, and as a consequence, a reduction in the costs of the technologies [54].

The technologies with the higher cost reductions are undoubtedly wind and solar PV power. They are both now considered mature technologies and achieve better costs than conventional fuel-based energy alternatives in some energy systems, particularly when fuel costs are higher (for example, in islands where the logistics increments the price) and renewable resources are abundant. Moreover, renewable energies avoid the increased risk and dependence of the volatility of fuel prices. However, they have their own uncertainties, derived from their own nature (non-manageable energy sources) and also from a constantly changing regulatory framework (at least in some countries) that tries to balance the incentives given at a certain moment with the cost reductions experienced in the posterior years. The financial crisis of 2007-2008 also had an important impact on the costs of renewable energy, where manufacturers had to compete in a shrunk market [54].

Several studies can be found in the literature analyzing the impact of regulation and policies on the renewable energy sector. In [55], an analysis on how policies fostering innovation and research and development (R&D) of renewable technologies affect the number of patents is presented, highlighting the importance of policy instruments on the development of renewable technology, and concluding that long term policies significantly increase the patenting and R&D activity.

32

In [56], an extensive analysis of the different drivers and barriers in the deployment of renewable energy in the EU is made, where policies are found to be a key enabler of renewable energy deployment success. A similar analysis in China evidenced the positive impact of renewable energy laws over renewable installation rates [57]. Along the same lines, government subsidy policies are seemingly of critical importance in Japan deployment rates [58].

In [59], a similar study conducted in the USA examining the effects of different policies on the Balance Of System (BOS) costs concluded that cash and property incentives increase deployment and reduce BOS costs, while renewable portfolio standards potentially slow down reduction in costs. The presence of interconnection standards could potentially promote deployment and reduce BOS costs. A comparison between the impact of policies and oil prices was performed in [60], concluding that public policies play a role far more important than oil prices in renewable deployment rates. In [61], the effectiveness of different policy targets is examined, concluding that policies targeted toward market expansion will be more successful once market preparation policies, such as Power Purchase Agreements (PPA) standards, are in place.

In Spain, policies governing electricity generation have been very unstable in terms of incentives and feed-in tariffs, particularly for solar PV installations. First policies promoting renewable energy date from 1997, where FiTs were approved for the "special regime", which mainly referred to wind and solar power. A further increase in the retribution in 2004 and change in the conditions for the facilities approved in the following years lead to an unprecedented deployment of solar PV, placing Spain as the global leader in PV installation in the year 2008 (more capacity added than countries such as the USA or China). The installation boom in Spain is explained by the disproportionate warranted revenues for the investors. The situation was unsustainable, and solar PV FiT costs were around 50% of the total renewable support costs, while its participation was around 10% of total renewable electricity generation. This would lead later in the future to the so-called "tariff deficit", which was one of the reasons for the abrupt stop of solar PV development around 2010 [62]. By 2012, with the arrival of the conservative party to the government, financial incentives and FiT for renewable energy were suspended in a series of regulations, resulting in a complete slowdown in their expansion. Decreases in costs over the last years, and the arrival of the progressive party to the government with a new point of view and a set of policies for the promotion of renewable energy, have now put Spain back on track to acceptable renewable energy installation rates. It is worth mentioning that Spain is starting to adapt its national regulation framework to comply with the European Directives and Regulations approved in the last years. For instance the Directive (EU) 2018/2001 [63] that promotes the use of renewable energy sources across different sectors or the Directive (EU) 2019/944 [64] and the Regulation (2019/943) [65, p. 9] that establish common rules for the internal electricity markets. Some of the actions proposed by these Directives enable the interaction of new actors, such as independent

electricity aggregators or electricity storage owners, in the electricity markets, which ultimately will translate into higher competition in the electric sector, and therefore higher rates of renewable energy in the electric systems.

Below, the principal laws, royal decrees and other regulations concerning the retribution of renewable energy are listed and briefly described.

- Ley 54/1997 [66]. Electric sector law, which first introduced the Special Regime category to promote the installation of renewable energy plants.

- RD 436/2004 [67]. Approval of a FiT for special regime generators.

- RD 661/2007 [68]. Gave priority access to the grid to renewable generators, more favorable conditions for facilities larger than 100 kW, and a revision of the FiT rates only every four years is set.

- RD 1578/2008 [69]. Established an annual 400 MW cap on support for new PV installations, and a reduction in FiT levels for future PV installations. Also established a maximum period for which PV subsidies were available.

- RDL 1/2012 [70]. Suspension of financial incentives for all renewable energy generators.

- RDL 9/2013 [71]. Renewable energy installations no longer receive a FiT for the energy produced, but a special remuneration in reference to the installed capacity, guaranteeing a so-called "reasonable return" for the installation.

- RD 900/2015 [72]. Imposed several restrictions to the self-consumption and approved a new tax for the PV installations connected to the grid (the so-called "sun tax").

- RDL 15/2018 [73]. Introduced new modalities for self-consumption installations, suspending the "sun tax".

- RD 244/2019 [74]. Expanded the formulations in the RDL 15/2018, promoting self-consumption plants.

Figure 15 shows the solar PV installed capacity in Spain. It can be observed the huge impact that the different policies abovementioned have had during the years, being particularly noticeable the stagnation suffered after the approval of the RDL 1/2012, which caused several years of almost no capacity additions until 2019.

34
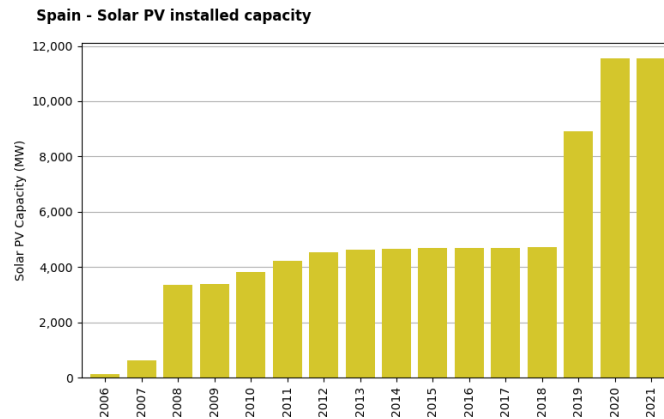
Figure 15. *Solar PV capacity installed in Spain over the last 15 years [75].*

2.1.6 Conclusions and future prospects

Compared to a decade ago, the availability of irradiance and solar databases, as well as the methods and models to estimate solar PV potential and specific yield, have hugely improved. Nowadays, very reliable estimations for solar PV potential can be found in several well-known solar databases. However, lack of resolution in some of the models may introduce errors for some particular regions, where climatic conditions are not as continuous as assumed by the models. Therefore, the proposal of new models with higher accuracy for these regions is required to have better estimations of the PV potential or specific yield. The introduction of new big data and artificial intelligence techniques in the models is an interesting approach that could yield better estimations, making use of the vast amount of data collected by utilities every day.

Given the history of renewable energy policies in Spain, the impacts that regulation have had over the installation rates and the overall energy transition are undeniable. Spain is a very particular case when it comes to renewable energy policies. A set of structural features of the electric system, design flaws and inflexibility on the policies push away Spain from world leader in PV installations in 2008, to installation rates close to zero during the next decade, evidencing that a stable regulatory framework is required to achieve higher PV shares in the electricity mix in the coming years. With the compliance of the new European Directives presented above, it is expected that such stable regulatory framework finally lands on the Spanish policy landscape, enabling higher and consistent renewable energy installation rates in the country.

2.2 Solar Forecasting

Forecasting solar irradiance and PV power output was identified by the International Energy Agency (IEA) at the beginning of the decade as a key element for enabling high PV power penetration rates in future electric grids [76]. Since then, a vast body of research has been conducted and a large amount of publications can be found in the literature, concerning all the different time horizons, techniques and applications that solar and PV forecasting have. In this section, an overview of solar and PV forecasting basics, including different classifications, methods and techniques, is presented.

2.2.1 Basic considerations

2.2.1.1 Solar forecasting and photovoltaic forecasting

First of all, it is important to differentiate between the terms solar forecasting and PV forecasting. Solar forecasting usually refers to the forecasting of solar irradiance or irradiation. That is, it does not necessarily take into account that these forecasts will be used for solar energy purposes (although most of the uses will be in the energy field). In other words, it does not take into account later interactions between irradiance and power generating systems. On the other hand, PV forecasting explicitly states that the forecasts will be produced for (and sometimes from) PV plants. This could be done by means of modelling the PV plants with other environmental variables, or by using PV data directly. While the first could be used to produce the second type of forecast, it could also be used for other applications, including forecasting of concentrated solar power CSP plants, or other applications beyond the energy field.

2.2.1.2 Solar irradiance components

The irradiance reaching the outer part of Earth's atmosphere is known as the extraterrestrial irradiance, which fluctuates depending on the day of the year, taking values of 1360 ± 50 W/m^2 [77]. As the radiation goes down through the atmosphere, several complex processes as reflections, absorptions, scattering and re-emissions take places, due to the interactions of irradiance with atmospheric molecules. As a consequence of these effects, the extraterrestrial irradiance reaching the surface is divided into two different components: the Direct Normal Irradiance (DNI) which is the irradiance reaching a surface perpendicularly placed, and the Diffuse Horizontal Irradiance (DHI), which accounts the radiation scattered coming from every other direction. These both components form the Global Horizontal Irradiance (GHI), which is defined by:

$$GHI = DHI + DNI \times \cos\theta_z$$

36

being θ_z the solar zenith angle. Frequently, the desired parameter is the irradiance reaching a tilted surface (since PV modules are usually tilted), so these components are translated to the plane of array incidence, calculating then the Global Tilted Irradiance (GTI), also known as Plane of Array (POA) irradiance. There are several models to estimate the GTI [78], which differs from each other in the complexity of the model itself, or the use of empirical values for its determination. Probably, the most known and utilized models are the Hay model [79] and the Perez-Ineichen [80] model.

DNI is usually used for CSP since this technology uses only direct irradiation, which, in general, is concentrated to heat up some intermediate medium to produce steam and run a steam (or other fluid) cycle, although this definition would depend on the kind of CSP technology considered. GHI is more broadly applied to solar PV models, since PV modules make use of both the direct and the diffuse irradiances.

2.2.1.3   Clear sky models

Undoubtedly, the most influencing factor affecting solar irradiance is the presence of clouds. Therefore, clouds are the main obstacle to make accurate irradiance predictions. However, it is possible to estimate the solar irradiance in clear sky conditions i.e. absence of clouds. Clear sky conditions are quite useful to model irradiance conditions of a particular location throughout the year, without taking into account the effect of clouds.

Clear sky models have a wide range of applicability in other forecasting techniques. For example, some persistence forecasts make use of clear sky models. Satellite based forecasts also could make use of the clear sky models to improve their performance. Clear sky models could also be used as a baseline for comparing other forecasting techniques. Finally, they could be used as inputs for the estimation of PV power generation under some particular conditions.

Generally, clear sky models are developed using Radiative Transfer Models (RTM) which require meteorological inputs such as ozone content, water vapor content or turbidity, as well as solar geometry data [77], [81]. There are numerous RTM used to construct clear sky models, such as the Solis model [82], European Solar Radiation Atlas (ESRA) model [83], the Ineichen model [84] or the Reference Evaluation on Solar Transmittance 2 (REST2) model [85], among others. The selection of models will mainly depend on the quality and availability of data [77].

2.2.1.4   Persistence forecasts

Persistence forecasts are the simplest forecasting techniques that could be implemented. As the name suggests, they are based in the persistence or continuity of the current conditions expanded to the designed forecast horizon. In the case of solar

irradiance, that means that the irradiance conditions at the time t would be equal to the irradiance at the forecasted time $t + \Delta t$. Generalizing:

$$fv(t + \Delta t) = fv(t)$$

where fv is the forecasted variable (PV output, irradiance, etc.) t is the time when the prediction is made and Δt is the forecast time horizon. Depending on the type of forecasting being made, persistence models can be tweaked to improve their prediction capabilities, producing the so called Smart Persistence (SP) models [77]. In the case of irradiance forecasts, SP models are usually made by adding a clear-sky index (CSI) in the formulation. CSI is the fraction of the measured irradiance over the clear-sky irradiance at time t. SP can be then defined by:

$$I(t + \Delta t) = k_t(t) \cdot I^{clr}(t + \Delta t)$$

where I is the irradiance, k_t is the clear sky index and I^{clr} is the clear sky irradiance for the forecasted time horizon. Since solar irradiance time-series are not stationary in the mid- and long-terms, other approaches have been proposed. For instance, in [86] a method introducing a stochastic component (representing the influence of clouds) to the persistence model is presented, which uses clear sky conditions as the stationary part of the model. Similar approaches can be used to improve persistence models for PV power forecasts.

Persistence models or forecasts are commonly used as a baseline to evaluate the performance of other forecasting models. Despite its naiveness, in some scenarios and situations, especially when time horizons are short, persistence and smart persistence models have proved to be very tough to beat. For instance, in intra-hour forecasts with clear sky conditions with very short time horizons, persistence models tend to perform better since predicted and measured values are close to each other. However, when clouds are present, their performance considerably decreases with cases where almost random predictions can outperform it.

2.2.1.5 *Performance metrics*

Due to the vast number of different approaches utilized to perform solar forecasts, it is very difficult to find a performance metric suitable to establish comparisons between all of them. Moreover, solar irradiance is inherently dependent on geographic location, time of the year and climate, which arises another problem of comparability. Despite that fact, there have been several attempts to estimate the performance of the models and devices

38

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 2: Theoretical Background

to make fair comparisons. Traditional statistical metrics may be sometimes used to characterize model quality. Some of them are listed and described below[6]:

– **Coefficient of determination (R^2).** The coefficient of determination shows how correlated the forecast values and the real values are.

$$R^2 = 1 - \frac{Var(\hat{I} - I)}{Var(I)}$$

– **Mean Absolute Error (MAE).** It shows the average distance between the measured values and the model predictions.

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|\hat{I}_t - I_t\right|$$

– **Mean Absolute Percentage Error ($MAPE$).** It is the MAE expressed as a percentage with respect to the real value and normalized.

$$MAPE = \frac{100\%}{N}\sum_{t=1}^{N}\left|\frac{\hat{I}_t - I_t}{I_t}\right|$$

Among the traditional statistical metrics, the International Energy Agency (IEA) in its Task 14 [76], has made recommendations for reporting the accuracy of the forecasting models, based on the following three performance metrics:

– **Root Mean Squared Error ($RMSE$).** It is a measure of the average spread of the errors (penalizes large errors). Model dispersion.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{I}_t - I_t\right)^2}$$

– **Mean Bias Error (MBE).** Evaluates if the model over or underestimates. Average Bias of the model

$$MBE = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{I}_t - I_t\right)$$

---

[6] In the following formulas, $\hat{I}_t$ (i hat) represents the forecasted variable (in this case irradiance, but they can apply to PV power as well) at time t and I_t represents the real measurement at time t.

39

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

2.2 Solar Forecasting

– **Kolmogorov-Smirnov Integral (KSI).** This aims to quantify the ability of the model to reproduce observed statistical distributions. Ability to reproduce frequency distributions.

$$KSI = \int_{x_{min}}^{x_{max}} D_n \, dx$$

where D_n is the difference between two cumulative distribution functions (the forecasted and the real).

– **Skill score or forecast skill (ss).** Sometimes it is desirable to present relative errors compared to a baseline rather than the absolute errors. For that reason, several authors have proposed alternatives metrics to show these magnitudes as a percentage. Recently, a novel performance metric for the evaluation of the quality of forecast models, the skill score or forecasting skill (ss), was presented in [87]. This metric is based on the comparison of two additional terms: the solar resource variability (V) and the forecast uncertainty (U).

Solar irradiance variability (V) is affected mainly by solar position and cloud cover. Fluctuations due to the solar position are deterministic and they are included in clear sky models. However, fluctuations due to cloud cover are completely stochastic processes and they are very difficult to model. For that reason, the accuracy of the solar irradiance forecasting models depends almost exclusively on how well they deal with these stochastic fluctuations. The authors in [87] addressed only the stochastic component, removing the fluctuations due to diurnal and yearly changes. With this idea, they defined the solar irradiance variability (V) as:

$$V = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{I_t}{I_t^{clr}} - \frac{I_{t-\Delta t}}{I_{t-\Delta t}^{clr}}\right)^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\Delta k_t)^2}$$

On the other side, forecast uncertainty (U) is a normalized version of the $RMSE$ with respect to the clear sky irradiance (I_t^{clr}).

$$U = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{\hat{I}_t - I_t}{I_t^{clr}}\right)^2}$$

With these two terms, the skill score (ss) can be obtained using the following equation.

$$ss = 1 - \frac{U}{V}$$

40

As *U* diminishes, the forecast skill tends to 1. That means that when *ss* is 1 (and therefore *U* is 0), the forecast method performs perfectly since no uncertainty exist. When *ss* tends to 0, forecasts are dominated by solar variability. This is the case of persistence forecast, which has a skill score of 0. As a result, this parameter can be seen as a measure of the quality performance of the model using the persistence model as a baseline, and can be approximated using the following equation:

$$ss \approx 1 - \frac{RMSE}{RMSE_p}$$

In this case, *ss* can take values from 1 (perfect forecasting) to arbitrary large negative values (persistence forecasts are better than the evaluated model).

2.2.2 Classification of forecasting techniques

Across the literature it can be found different classifications of solar forecasting techniques. Two of them can be found in the complete reviews in [77], [81]. However, a forecasting technique is defined by several aspects which are dependent one on another. In this section, the most important differentiations between techniques to the author's opinion are presented, taking into account that every forecasting model or technique will fall into one of the options presented for each category. It has to be noted that these aspects are often inter-related, for instance, methods with low spatial resolutions would probably require local measurements, while for models considering large areas, satellite images or numerical weather prediction models may be used. Solar and PV forecasting techniques can be classified according to several aspects, which will be discussed in the following subsections.

2.2.2.1 Approach: physical based methods vs data based methods

According to the approach used to produce the forecast, forecasting techniques can be classified into either physical based or data based methods. The prior family of methods aims to model the problem with analytic or semi-empirical equations, using different inputs such as weather variables, data from installations, satellite images, etc. In a sense, they try to simulate the different systems governing the irradiance values or even the PV output values. Data-based methods generally use past records of the forecasted variable as their only input.

Physical based methods usually rely on data coming from Numerical Weather Predictions (NWP) models. NWP models are by themselves forecasting models, which try to predict future weather variables, given the current status of the system. NWP models generally use mathematical functions to implement the physical equations governing atmospheric conditions such as wind, pressure, density and temperature: the Navier–

2.2 Solar Forecasting

Stokes, the mass continuity equations, the first law of thermodynamics and the ideal gas law [88]. Although their performance has improved dramatically over the last decades [88], NWP still are sources of uncertainty, which then propagates to solar forecasting. This is probably one of the main disadvantages of these methods. On the contrary, an advantage of physical based methods is that no historical data is needed to make the predictions. Therefore, they are suitable for obtaining power predictions before the plants have been constructed [77]. Different studies using these approaches can be found in [89]–[92].

Data based methods use, in general (but not exclusively), past data of the forecasted variable as their only input. Data-based methods can be further classified according to the nature of the operations used to treat the data in Statistical, Machine Learning or Deep Learning methods. These methods are then used to identify patterns and extract trends in the data, aiming to predict the forecasted variable. The principal constraints of these methods are that they need a huge amount of data, and the quality of the data used has a huge impact on the overall accuracy of the methods. Among the statistical methods, there are a series of autoregressive methods that are commonly used including Linear stationary models (Auto-Regressive (AR) [93], Moving Average (MA), Mixed Auto-Regressive Moving Average (ARMA) [94], [95], Mixed Auto-Regressive Moving Average models with exogenous variables (ARMAX)), Linear non-stationary models (Auto-Regressive Integrated Moving Average (ARIMA) [96], ARIMA with exogenous variables (ARIMAX), Seasonal Auto-Regressive Integrated Moving Average (SARIMA), SARIMA with exogenous variables (SARIMAX)) or Non-linear stationary models (Nonlinear Mixed Auto-Regressive Moving Average models with exogenous variables (NARMAX)

Additionally, we have the ML and DL methods (also called Artificial Intelligence methods). The variety of types of models is huge, and since they are both areas of active research, this number increases every year. Some of the most utilized algorithms are k-Nearest Neighbors (k-NN) [97], Support Vector Machines (SVM) [98], [99], Random Forests (RF) [100], [101], Artificial Neural Networks (ANN) [102], and all their derivations. For an in-depth discussion about the use of different machine learning techniques in the solar forecasting area, the reader is referred to [52].

The availability of data and its quality is particularly important for Machine and Deep Learning strategies, since a considerable part of the dataset would be used for training the models. Scarcity of data during the training process leads to poor quality models. Another disadvantage that ML and DL methods have is that they do not generalize as well as physical based approaches (or do not generalize at all), since they are trained with a specific dataset corresponding to a particular location, plant or problem. Thus if one wants to use the methods in different locations or problems, the models must be retrained. Despite that, the overall architecture of the models could still be reusable (for example, the structure of the model itself or parts of it, as the image encoders). A remarkable feature of

42

ML and DL approaches is that the models are agnostic and assumption-free (data should be self-explanatory)

Finally, **hybrid models** are models that combine both approaches (physical and data-based methods). In general, they are more complex, but they theoretically should yield better results. With the recent explosion of big data and AI, and the continuous improvement of computational capabilities of computers, current research trends seem to be more focused on data based and hybrid models. However, physical models will continue to be important in the long run, since they will benefit from some of the abovementioned improvements as well.

2.2.2.2   Time horizon

The main way in which solar forecasts are usually classified is considering their time horizon. Time horizon is probably the most relevant characteristic of a forecasting method, since it will determine the applications of the forecasts produced. Electric systems are very complex, and involve interactions of multiple actors at different levels (from the delicate equilibrium between generation and load, to the numerous bidding processes and market interactions occurring throughout the day), thus there are several aspects that should be considered in the decision making process, each of which requires specific information that can be obtained with forecasts for specific time horizons.

Regarding the terminology of the time horizons, there is no clear consensus about which terms should be used. Traditionally the terms "very short-term", "short-term", "mid-term" and "long-term" have been used, as one can check in the numerous studies and reviews found in the literature [52], [81], [103]–[106]. However, more precise descriptions of the time horizons are preferred, and are being adopted by the community in the last years [107]. So in order to follow the best practices, we are adopting the more descriptive definition of the terms as in [77], [106].

2.2.2.2.1   Intra-hour forecasting

Intra-hour forecasts span from 1 second up to 1 hour. Their main applications are the maintenance of grid stability (forecast ramps and high-frequency variability [108]), making a correct schedule of the optimal spinning reserves, power smoothing, real-time electricity dispatch and demand response [77], [109]. This type of forecasting is especially suited for islanded, isolated or weak electric grids, where high shares of solar PV energy is present in the mix [77]. Intra-hour forecasts are usually produced by means of local sensing networks (that could be sky-imagers, neighboring plants, irradiance sensors...) and real-time data of the variable to predict.

2.2.2.2.2 *Intra-day forecasts*

Intra-day forecasts typically cover from 1 to 6 hours and their main applications include load following or optimize bids in intra-day markets. Grid operators also use them for controlling different load zones or trading outside the boundaries of their area [77], [110]. The most common techniques for this time horizon are statistical methods and satellite imagery, although analysis of neighboring plants and NWP are also possible.

2.2.2.2.3 *Day-ahead forecasts*

Day-ahead forecasts cover the time span from 6 to 24 hours in the future. These forecasts are primarily used for planning, unit-commitment or placing biddings in the day ahead markets [77], [91]. The technique dominating this time frame is undoubtedly NWP, which gives an idea of the real value of NWP models, more suitable to study future trends and not sudden changes in the near future.

2.2.2.2.4 *Two days ahead or longer*

Forecasts with time horizons above the 24 hours include a variety of applications, from unit-commitment, transmission management, trading or optimize the maintenance planning and modeling of PV plants[77], [108]. For longer periods, there is a point where the boundaries between forecasting and modelling can start to blur. That is, predicting the output of a potential PV plant in its whole useful life can be seen as forecasting or modeling.

2.2.2.3 **Spatial resolution**

In terms of the spatial horizon or resolution, solar and PV forecasts can be point forecasts (made for a single plant), regional (for an ensemble of plants or a particular area) or even global (covering the whole globe). Commonly, grid operators would prefer regional forecasts since they are more useful to maintain the balance between the demand and the supply in the electric system. On the other hand, micro-grid operators may prefer individualized forecasts since their objective is to maintain quality and stability of the electricity in a smaller area.

In order to produce regional forecasting, upscaling is the preferred technique [76]. Upscaling is a well-known technique applied in wind power forecasts, and it basically tries to extrapolate to a large region the results from fewer points. It consists in selecting a set of representative data of the plants in a determined region, and upscale the results of a forecasting done to these data by means of normalization or other method utilizing nominal capacity of PV plants. Its popularity during the last decades has been due to the difficulties of obtaining detailed information of all the PV systems in a system. Although with new data acquisition devices and techniques, this could change in the near future.

44

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Generally, regional forecasts have significantly better accuracy as the geographic area of study increases. For instance, in [92] reductions in RMSE up to 64% for a forecast over an area comprising the whole country of Germany compared to a point forecast were reported. Similarly, in [89], reductions in RMSE up to 69% were found when forecasting 10 plants combined in comparison with a single forecast. This error reduction with aggregated information may seem obvious, since correlation between errors from two locations decreases as their distance increases, leading to a partial cancelation of such errors at individual sites [77], [92]. This is what is called the "smoothing effect". In a sense, is the same principle by which, the aggregated load forecast can be considered an easy problem, while individual load forecast is a very complex one.

2.2.2.4 Deterministic vs Probabilistic forecasts

According to the output produced by the forecasts, these can be classified into deterministic (often referred to as point forecasts, but this term should be avoided since it could also have a spatial meaning) or probabilistic. Deterministic forecasts output a single value and probabilistic forecasts output a range of weighted values, a probability density function of the possible outcomes that encodes the likelihood of each value to occur. Traditionally, deterministic approaches have been more widely adopted than probabilistic, probably because the complexity level of the models is lower, although some authors suggest that the lack of good performance metrics for probabilistic forecasting may have induced to think that their performance was worse than it really was [104].

Deterministic approaches do not consider valuable information such as upper and lower bounds of forecasted values or its degree of confidence. However, once the deterministic methods are evaluated, and their error rates are known, one could simply translate these errors into probabilities using any distribution (normal distribution generally is a good approximation to errors). Therefore, a sort of probabilistic forecasting can be obtained from deterministic forecasts.

2.2.2.5 Data sources

Another way to differentiate forecasting methods is regarding the data sources input on the models. A wide variety of data types can be used to feed the models: past records of PV plants, weather data, satellite images, sky images from ground based cameras… Besides the type of the data, the origin of the data can also be used as a way of classifying the models. There are models that use endogenous data, that is, time series of records of PV production or irradiance; and there are models that use exogenous data, which represent external data sources to the variable being forecasted, such as temperature, pressure, etc.

2.2.2.6 Application

Finally, a final characteristic would be the application of the forecasts. At this point most of the applications have been discussed in previous paragraphs, but they are listed below regardless for quick reference.

- Grid stability and power quality
- Markets interactions
- Optimization of conventional generators increasing and optimizing ramp response
- Load following
- Unit-commitment and scheduling
- Regional scale power prediction

Figure 16 aims to synthetize all the relevant information regarding applications and spatial and temporal horizons of solar forecasting methods in an image.
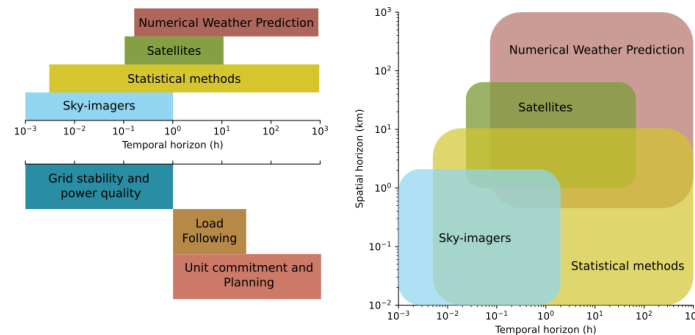
Figure 16. *Solar forecast methods and applications depending on the time and spatial horizons.*

2.2.3 Intra-hour solar forecasting with sky-imagers

Solar and PV intra-hour forecasting refers to forecasts made in the very short term, covering time horizons from a few seconds up to 1 hour. The main applications of this kind of forecast are to assure grid quality and stability (by predicting production ramps of PV plants), as well as making a correct schedule of the spinning reserves. The role becomes more important when talking about islanded or isolated power grids, or weak grids with high solar penetration. Also the role of solar intra-hour forecasting may change depending on the voltage level of the grid at which the PV plants are connected. For instance, if solar PV plants are connected at the low voltage level, the major objective may be the prevention of voltage limits violation. On the other hand, when forecasts are performed to

46

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
            UNIVERSIDAD DE LA LAGUNA

            María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
            UNIVERSIDAD DE LA LAGUNA

know the productions of large PV plants connected to higher voltage levels, the main purpose could be to regulate frequency or make better bids in intra-hourly markets.

In the intra-hour temporal horizon, the main factor affecting PV production is the clouds. Position and shape of clouds are one of the most difficult phenomena to predict in the atmosphere due to its stochastic and dynamic behavior, which makes it very tricky to find models to correctly define them. Since analytic modelling of individual or small groups of clouds through mathematical equations seems to be an almost impossible task, one of the main approaches has been to track the movement of the clouds trying to predict where these clouds will be in a short temporal interval. Cloud tracking can be made by means of different approaches and technologies (sky-imagers, irradiance sensor network, pyranometer arrays, satellite images, etc.). Each of them has its own advantages and disadvantages. For example, satellite images cover larger regions, but their sampling frequency is usually too low for the intra-hour forecast. Irradiance sensor networks provide a simplified dataset with potential low costs on its acquisition, but heavy computational post-processing is required to provide meaningful predictions. A sky-imager provide high quality images of the sky and the clouds, but its spatial range is limited to its surroundings and require complex post-processing and image-processing techniques. Probably, the best forecasting models to come will use a combination of some of the techniques. But since the main focus of this thesis regarding solar forecasting is the use of sky-imagers, the following paragraphs will describe the utilization of these devices and the research related to it.

The first reference to a sky-imager for cloud field or cloud cover estimation date from 1993 [111], although the device had been used for other purposes decades earlier. The first devices were far more complex than today's models in terms of the optics required: convex mirrors to emulate the wide effect of angular or fish-eye lenses and they used a device to block the sun from the image, with the associated mechanisms to move it. With the technological improvement in the manufacturing of lenses, and the decrease in costs of the electronic components, new sky-imager models are quite different to the earlier models.

The use of sky-imagers is restricted to local (up to a few kilometers depending on the camera properties) and very short-term (up to one hour depending on climatic conditions) forecasting. There is a vast body of literature detailing the uses of sky-imagers for solar forecasting. Multiple studies have used sky-imagers to forecast cloud positions and movement and derive GHI [112]–[122] or DNI [123]–[129]. Other studies have focused on the use of sky-images to classify sky conditions, cloud types, cloud movement estimations and cloud reconstruction [130]–[138]. There have been authors with more original approaches such as using sky-images to create cloud shadow maps on the ground to forecast power production ramps geospatially [139], or to predict irradiance ramps directly [140]. Besides the papers where solar irradiance was the variable forecasted, very

47

few studies have focused on the forecast of actual PV generation [141]–[146] (at least, using sky-imagers).

2.2.4 Combining sky-imagers with AI for intra-hour forecasting

The relatively small number of studies dedicated to forecast PV power generation in comparison with the forecast of GHI or DNI is partially explained by the lack or low availability of PV data existing a few years ago. Back then, the main objective of the research was to emulate the physics system governing irradiance conditions as well as possible, and then, make the predictions with the irradiance data for the PV plants. Nowadays, vast amounts of PV data are collected every second at every level of the electric system, and sometimes it is available for researchers. This explains a new trend to shift the research focus towards more data-based methods. ML and DL techniques have strongly emerged and have been applied to different data sets searching for a good solution that requires fewer infrastructures and could potentially achieve better results.

ML and DL methods do not generalize as well as physical based approaches, since they are trained with data or images corresponding to a particular location. Thus, if one wants to use the methods in different locations, models must be retrained. Despite that, parts or blocks of the models could still be reusable (for example, the structure of the model itself or parts of it, as the image encoders). A remarkable feature of ML and DL approaches is that the models are agnostic and assumption-free (data should be self-explanatory). This means that there should be little room for the conjectures of the researcher other than the suitability of the models' architecture, the fine tuning of the model or feature engineering. There is also a rich body of literature regarding the use of ML and DL models to short-term solar forecasting. Some of the most relevant for this study are discussed below.

Regarding the use of ML and DL for data-based solar forecasting (without sky-imagers), various authors have decided to forecast GHI and DNI [147]–[155] as it happened with the sky-imagers. But in this case, there are far more studies which use PV data to directly forecast PV power [156]–[164].

Combining the use of sky-imagers with ML or DL techniques is also possible. Several authors have used sky-images to feed DL models, particularly Convolutional Neural Networks (CNN). In [165], the authors use sky-images to now-cast the output of a 2.5 kW PV plant, using a set of different models. In [166], the authors used video inputs to forecast the output of a 30 kW PV plant. This study is extended to new model architectures and data inputs in [142], [167], [168].

48

2.2.5   Conclusions and future prospects

Intra-hour solar forecasting will be a key technology for ensuring quality of electricity and grid stability in the future, as larger shares of PV energy are deployed in the electric grids. It will be particularly useful in islands and weak electric grids, where the effects of PV variability are more severe than in continental and interconnected grids. The suitability of these types of forecasting systems and the added value they can provide to the system operation will depend on the environmental conditions and the particularities of the electric grid where they are installed.

Current trends in the field are heading towards the utilization of ML and DL techniques in conjunction with physical methods, and they are expected to dominate the future landscape as well, since the amount of generated data coming from PV systems or other sensing devices keeps growing. As the performance of forecasting methods continues improving, as well as the capabilities to integrate the information provided by forecasting systems with the operation of the electric systems, it is expected that the adoption of techniques to integrate forecasts in the decision making process (at several levels) will be much higher than it is today.

For intra-hour solar forecasting, sky-imagers are expected to continue decreasing in costs (in line with the fall in cost of electronic and optic parts), while the quality of their services will increase substantially with the introduction of new techniques (such as deep learning models). This could enable their massive deployment, especially (but not exclusively) in insular contexts and in micro- and smart-grids applications, where the added value of the technology is higher.

If the reader is interested in a more in-depth review of solar and PV forecasting techniques, there are several review papers that provide valuable structured information, and which some of them focusing on a particular topic (machine learning algorithms, probabilistic forecasting...) [77], [103], [104], [110], [169], [170].

2.3 Integration of PV systems in distribution grids

As pointed out in Section 1.3, where some of the problems regarding the integration of solar PV in the electric grids were briefly discussed, the integration issues will depend on a variety of factors: robustness of the grid, penetration levels of PV in the mix, voltage level of the connections, etc. In this section, an in-depth overview of the main characteristics of distribution grids and the impacts associated with a high penetration of PV systems is presented.

2.3.1 From centralized electric grids to distributed generation

The electric grid operates at different voltage levels, and thus is divided in different subgrids according to the voltage at which they operate. For example, transmission grid operates at high voltages to minimize the losses associated with the transportation of bulk electricity to larger distances through power lines. The voltage is then stepped-down and the power is injected in distribution grids. Distribution grids can either operate at mid- or low-voltage levels, and they are where all end-users are connected, and therefore, safety and stability issues are a major concern.

Due to the modular capabilities of the technology, PV systems can be designed in a wide range of capacities: from a few kW that households can install in their rooftops, to multi-MW and even macro projects in the GW scale. In general, large projects are connected to transmission grids to efficiently evacuate all the power generated, while small PV systems are connected to the distribution grid. The latter case is what is called Distributed Generation (DG), and it is completely changing the paradigm in the generation of electricity in the last years. This change poses operational and engineering challenges to the Distributions System Operators (DSO), which are responsible for the quality of the electric supply.

Traditionally, electric grids have been operated radially, with large generators evacuating the generated power in the transmission grids [171]. Power is then distributed to the loads, generally after reducing its voltage, in a unidirectional manner. With the introduction of DG in the distribution grid, now power injections occur at the distribution level, and power flows can be bidirectional, coming from the demand-side of the line (depending on the load conditions). This could represent a problem for a variety of reasons, mainly intrinsic to the characteristics of the distribution grid, which are discussed in the next sections.

50

2.3.2 Impacts of PV systems on electric grids

As stated in the section 1.3 there are a variety of impacts that the integration of PV systems has over the electric grid. Some of them affect the overall system, while others are localized in certain distribution grids or feeders. Factors conditioning the severity of the impacts include the voltage level at which the PV systems are connected, the PV penetration levels, and the robustness of the electric grid itself.

The robustness of an electric grid makes reference to all the characteristics of the grid that help support its stability, such as size, number of high inertia generators, interconnections, etc. For instance, electric systems on small islands are more likely to be less robust than continental systems. Continental systems have more interconnection potential, more rotating generators and their renewable energy generation is scattered in larger areas. On the contrary, small islands usually have most of the renewable sources concentrated in one limited region, high interconnection costs depending on the distances to other electric systems and the bathymetry of the surrounding sea floor, and in general, less number of large high inertia generators. Therefore, it is more likely that pernicious effects from a high PV penetration will manifest in small and isolated electric systems rather than in large continental ones. In the following subsections, some of the most common impacts of PV integration are discussed.

2.3.2.1 Frequency instability

Frequency is probably the most important factor governing power quality. It is controlled by maintaining the balance between generation and load at all times (if generation is higher than the demand, frequency increases, and if it is lower, frequency decreases). Frequency needs to be controlled within a tight range, which varies depending on the country and the electric system (in Spain the frequency is 50Hz and the range is ±0.15Hz for the mainland, and ±0.25Hz for the islands [172]). Disruptions in the balance between electric supply and demand lead to frequency fluctuations.

Frequency in power systems is controlled mainly by sources of inertia, such as large rotating generators and motors. Frequency level is maintained by control loops built into the generators. When deviations in frequency occur, generators and turbines adapt their fuel or steam flows to match the load with the demand, restoring the normal frequency value again [173].

The increasing power from VRE sources makes frequency control a harder task. One of the reasons that may provoke such frequency fluctuations is the passing clouds which cause rapid variations in the power output. Although nowadays the contribution of PV systems to frequency fluctuation is rather smaller compared with, for instance, wind power (due to uneven shares of both technologies), it is expected to become more noticeable as the deployment of such systems increases [174]. However, as presented in

51

the introduction, in weak electric grids with low number of high inertia sources, even small PV penetration levels could lead to frequency deviations. In any case, the increase in PV penetration will demand a higher frequency regulation capacity, with the subsequent derived costs [175].

2.3.2.2 Voltage fluctuations

As it happens with the frequency, the variability in solar power could lead to voltage fluctuations in the grid. If the voltage moves outside the safe level of operation, several performance problems may arise. Since voltage related issues are more important at the distribution level than in a general context of electric grids, and they are discussed in detail in the section 2.3.4.

2.3.2.3 Islanding operation

Unintentional islanding occurs when PV systems within a certain network continue to supply power to the loads even after the part of network is disconnected from the main grid for some reason [174]. Normally, when this situation occurs, power quality of the islanded network becomes degraded. In general, PV inverters are designed to detect abnormal power quality conditions and to disconnect from the network. However, although it is highly unlikely, if the PV generation and the load in the islanded line are identical, PV system protections might fail to detect the disconnection and will continue supplying power [174].

The unintentional islanding can cause a number of different problems: safety issues for network operators, maintenance of the fault conditions that tripped the circuit, damage to equipment due to the poor power quality, transient overvoltages, damage of inverters, switchgears and loads due to lack of synchronization during reconnection, etc. Islanding is a well-known problem and it is usual that inverters provide anti-islanding features. Although there are several ways of classifying them, islanding detection methods can be mainly implemented in two ways (assuming there is not communication link between equipment): passive and active methods, or also a combination of both (hybrid methods). However, the best options to improve islanding detection are based on the deployment of some communication infrastructure between the utility and the inverter [173].

Although traditionally islanding operation has been avoided, advances in PV inverters technology and other power electronic devices could lead to scenarios where islanding operation is possible, and even desirable to increase the resilience of the electric systems under certain circumstances. For example, in case of a major fault of the grid, some parts of the grid could disconnect from the main grid and, if grid-forming inverters are available, momentarily continue to supply electricity to the loads [176].

52

2.3.2.4 *Integration costs*

Another effect on the inclusion of VRE in the electric grid is the economic impact of the measures needed to cope with their variability, the so called integration costs (also known as hidden costs or system-level costs). Integrating VRE into power systems causes costs elsewhere in the system. Examples include distribution and transmission networks reinforcements, short-term balancing services, provision of firm reserve capacity, a different temporal structure of net electricity demand, and more cycling and ramping of conventional plants [25].

However, there is no consensus about the real definition of integration costs, and if they should be computed on the VRE. Other generation technologies impose integration costs which are not allocated to those technologies. For example, large generators impose contingency reserve requirements, block schedules increase regulation requirements, gas scheduling restrictions impose system costs, nuclear plants increase cycling of other baseload generation or hydro generators create minimum load reliability problems. And these costs are not computed to the generators causing them [177].

2.3.3 Basic characteristics of distributions grids

Distribution grids have their own special features in contrast with transport grids. Some of them are their unbalanced nature, the presence of dynamic loads and small distributed generators, higher R/X ratios... [178]. It is important to know the peculiarities of distribution grids in order to understand the effects that a high PV penetration has over its operation, and the problems arising from that condition. Some of the most important are discussed below.

2.3.3.1 *Decreasing section of lines with the distance*

As the ensemble of the electric grids, distribution lines were originally designed to operate radially, and it is common that the section of the cables decreases as the distance to the substation increases (the amount of current at the end of the lines is smaller than near the substations, so the shrink of the cables was justified from an economic and technical point of view). If power is injected directly in the distribution grid, the requirements to evacuate the generated power increase, and it is possible that some lines are not prepared for these distributed current injections, particularly if the generators are located at the end of the line.

2.3.3.2 *Unbalanced nature of distribution grids*

Voltage unbalance occurs when the magnitude or the phase angles of the voltages in a three-phase system are not equal (not exactly 120º). It is usually measured by the voltage

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

2.3 Integration of PV systems in distribution grids

unbalance factor (VUF) which is defined as the ratio of negative sequence to positive sequence voltage components [179].

Distribution grids are inherently unbalanced systems. Although the DSO tries to evenly distribute the capacity on each phase, shifts in the timing of loads by individual households can lead to imbalance especially as the size of the distribution grid diminishes. Another cause of voltage unbalance is the differences in line impedances, that could result from non-transposed lines [180]. Additionally, installation of single-phase distributed generators contributes to increase the voltage unbalance through the distribution grid. In unbalanced three-phase four-wire systems the power injections in one of the phases affects the other two, which is caused by neutral point shifting [181].

2.3.3.3 *Voltage regulation and low X/R ratios*

Another reason, which is also related to the original operational design, is the voltage regulation. Voltage regulation is usually done at the substation, where devices called On-load Tap Changers (OLTC) provide flexibility to the operational voltage of the line beyond the transformer, and they are usually set to values slightly above the nominal voltage. This is done to ensure that the farthest users in the line receive the electricity at an acceptable voltage, since there is a drop in the voltage magnitude as the current flows through the line. If voltage conditions change, OLTCs can regulate the voltage by stepping up or down. Other equipment, such as shunt capacitors, Static Synchronous Compensators (STATCOMs) and Static VAR Compensators (SVCs), that can contribute to the voltage regulation in every load condition are also possible [182].

With a large share of DG, active power is injected directly to the distribution grid. Lower X/R ratios[7] in distribution lines means that lines are more "resistive" and therefore, active power injections and absorptions have a larger effect on the voltage [183]. If the DG is mainly composed of PV systems, their high variability and low inertia make that a partly cloudy day could translate into a series of discontinuous injections of active power, with the subsequent effect on the voltage of the line.

2.3.4 Voltage-related impacts due to PV systems on distribution grids

Once reviewed the main characteristics of distribution grids, the impacts caused by a high PV penetration can be explained. The main impacts at the distribution level are related to voltage regulation and control, and to a lesser extent, harmonics or islanding operation. In this section, the potential voltage-related impacts that PV systems have over distribution systems are reviewed. Voltage is one of the most important parameters (if not

---

[7] Often in the literature, the inverse ratio is used (R/X). In that case, distribution grids have high R/X ratios.

54

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

the most) limiting the penetration of PV systems in distribution grids. Due to the strict quality requirements that distribution grids have, it is capital that voltage levels remain within the normative limits in order to ensure a secure and reliable service for the end-users.

If voltage levels fall outside the normative range, safety and performance problems arise. For example, undervoltage conditions can cause "brownouts", where some equipment would not be able to start up or the intensity of lighting would diminish. On the other hand, overvoltage conditions may decrease the lifetime of most equipment and damage sensible electronic devices [173]. Another consequence that voltage fluctuation has, is the continuous operation of voltage regulation devices, such as OLTC, with the subsequent decrease in lifetime [180]. The main voltage-related effects of a high PV penetration in distribution grids are described below.

2.3.4.1 *Voltage rise and reverse power flow.*

Voltage rise is inherent to any active power injection into the electric system, since, as injected current passes into the lines, it creates a voltage rise across the system impedance [184]. Due to lower X/R ratios in distribution grids, active power injections will have a major influence on this voltage rise. If a particular feeder has low PV penetration levels, overvoltages would not normally be an issue, since most of the produced power would be consumed by the users in the feeder. But if a significant PV capacity is installed, there could be moments when the produced PV power is higher than the load and a net export of power takes place. Therefore, load conditions of the feeder affect the severity of the voltage rise. PV power has its peak production around midday, sometimes coinciding with low load levels of households, so net exports are more likely to happen. Another factor affecting the magnitude of the voltage rise is the distance to the substation, where installations at the end of the line will require a higher voltage rise to evacuate the power. If these conditions are met, voltage rise and reverse power flows will happen and the stability of the line may be compromised.

2.3.4.2 *Power output fluctuation (variability)*

As discussed in previous sections, PV power systems have a high variability since they rely on solar irradiance, which at the same time experiences severe fluctuations as a consequence to the presence of passing clouds. Additionally, PV systems have very low inertia, and changes in their output power are almost instantaneous. These conditions may lead to power fluctuations in the feeder, which create high uncertainty on the operation of distribution systems. In general, PV power systems experience two types of fluctuations. On the one hand, there are short-term events with high frequency fluctuation due to passing clouds (common on partly cloudy days). These can cause power quality issues, such as light flickering or variable motor speeds, and increased number of operation of regulation devices [173], [184]. On the other hand, there are pronounced

ramps that severely reduce the power output (it can drop more than 60% in a matter of seconds [185]), which may happen in the transition from a clear sky to an overcast day, and which could require backup generation to maintain power supply. The effects of fluctuations are smoothed with the geographical dispersion of PV systems or the presence of storage devices in the feeder. Voltage regulation from PV inverters could be employed to mitigate them as well.

2.3.4.3 *Grid derived voltage fluctuations*

Another source of impacts is related to grid-derived voltage fluctuations. PV inverters are usually designed and set to operate in "grid voltage-following mode", disconnecting from the grid when voltage falls outside the predefined limits (programmed in the inverter). Some short-term events occurring in the grid such as voltage sags or peaks, could lead to the disconnection of PV inverters. If the PV penetration in a particular feeder is high, these events could result in the disconnection of a large number of DG, while loads do not, exacerbating the problem [184]. If this situation occurs, the network will still have to provide power to the loads without the support of PV systems, so additional power needs to be delivered by the lines of the network, leading to overload of the lines or long periods of undervoltage. Modern smart inverters now include the "ride-through" capabilities, to overcome this kind of event.

2.3.4.4 *Voltage unbalance*

PV systems in distribution grids can cause voltage unbalance on the line, particularly when single-phase inverters are deployed, and the capacity is unequally distributed on the different phases. The unbalance increase will also depend on the distance to the substation (the unbalance will be higher at larger distances). Voltage unbalance could have a negative effect on three-phase generators and loads, leading to rise in temperature, noise or vibration; and also in power electronic devices [173].

2.3.5 Technical solutions for voltage regulation in distribution grids

As discussed in the previous section, voltage related issues are the most critical among all the problems of integrating PV systems to distribution grids. Voltage regulation strategies can be implemented in order to prevent or mitigate the effect of these issues. There are a handful of strategies that can be implemented by either the DSO or the PV systems, or by both of them simultaneously. Technical solutions that are used to mitigate the effect of the impacts and achieve an effective voltage regulation throughout the network are described in this section.

56

2.3.5.1 Solutions provided by Distribution System Operators

DSOs are responsible for maintaining the power quality in distribution systems, and they can provide a series of solutions to mitigate the impacts of a massive deployment of PV systems in the distribution grid. Traditionally, overvoltage problems have been solved reinforcing the grid, but this is an expensive solution, and nowadays it is seen as a last resource measure. DSOs have other options available for voltage regulation, which can contribute to increase both the power quality and the hosting capacity of distribution grids. Some of these solutions are presented below.

2.3.5.1.1 On-Load Tap Changers (OLTC)

As discussed in previous sections, OLTCs are essential parts of the distribution grids, and they are the main voltage regulation device in the grids. They control the voltage changing the winding ratio of substation transformers, stepping up or down the taps without disconnecting the loads [180].

2.3.5.1.2 Line Voltage Regulators (LVR)

LVRs, also known as Step voltage regulators (SVR), or simply voltage regulators, are autotransformers with a regulation capacity of ±10% typically, used to stabilize voltage in long and heavy load feeders. They can be placed in the middle of the feeders to prevent voltage levels to fall outside the normative range.

2.3.5.1.3 Reactive power control or VAR-control form dedicated devices

The provision of reactive power through fast-acting power electronic devices, such as SVC, STATCOMs or capacitor banks can also help to deal with voltage fluctuations.

2.3.5.1.4 Grid reinforcement

Grid reinforcing is basically done by increasing the cross-sectional area of the lines to reduce the impedance and therefore the voltage rises due to active power injections. It is a rather expensive measure that in general is seen as a last resource.

2.3.5.1.5 Network reconfigurations and closed-loop operation

Grids can be reconfigured in terms of functional connections and paths for the power. An interesting option is to give the grid the ability to automatically reconfigure, thus behaving differently under different conditions. Closed-loop operation follows tries to emulate the operation of transmission grids in a sense. The idea is to increase the short-circuit power of interconnection points in the grid, strengthening the grid and reducing the effect of DG injections over voltage.

2.3.5.1.6 *Energy Storage Systems (ESS)*

Installation of centralized ESS can also contribute to the mitigation of impacts, with different technologies contributing to different control strategies.

2.3.5.2 ***Local control strategies provided by PV systems.***

DG, along with the improvement of power electronic devices, has brought with it the possibility of new operation strategies in distribution grids. Nowadays, PV inverters can be used to locally provide ancillary services to the grid, such as grid forming features, fault-ride-through capabilities or improved controls for voltage regulation. For voltage regulation, there are two main strategies, namely active power curtailment and reactive power compensation, which are described below.

2.3.5.2.1 *Active power curtailment.*

The simplest way to avoid overvoltage due to active power injections is to reduce the amount of power injected to the grid. Despite being the simplest strategy, it is not the most desirable, since it implies the spilling of renewable energy, although low X/R ratios in distribution lines makes this strategy quite effective. This strategy can be implemented in different ways, such as fixing a maximum percentage of active power over the rated power of the inverter, or make a gradual reduction of the power injected based on an external signal, such as line voltage. The latter strategy is called Volt-Watt (VW) control. The VW control is used to reduce the amount of active power delivered to the grid as a function of the voltage at the point of common coupling (PCC) of the solar inverter. An operation curve defining the voltage set point where the control should start to act is configured in the inverter [186]. An example of such curve is given in Figure 17. Active power curtailment can affect the economic viability of PV systems, and therefore, PV owners would be reluctant to apply it.
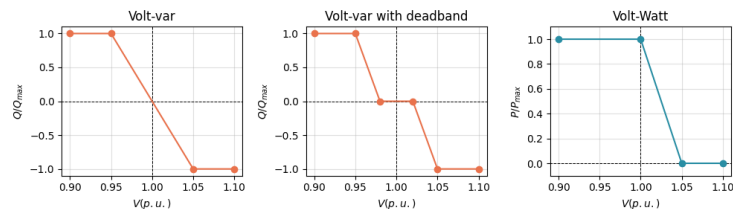
Figure 17. *Example curves for different voltage regulation strategies. Volt-var without deadband (left), Volt-var with deadband (center) and Volt-Watt (right).*

2.3.5.2.2 *Reactive power compensation.*

Another option is to control the voltage of the line through reactive power injections/absorptions. The main drawback of this method is that the effect of reactive power on the voltage depends on the reactance of the line, ignoring its resistance [187].

58

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Hence, the effectiveness of reactive power compensation is determined by the X/R ratio of the lines, which in distribution lines is low, meaning that reactive power interactions will have an overall weaker influence on the voltage than active power. Despite this, reactive power compensation has the potential to control the line voltage without the necessity of curtailing any active power, although some of the strategies can inherently limit the amount of active power injected. Reactive power compensation can be implemented in a variety of local control strategies.

- **Fixed power factor (PF).** Modern PV inverters can operate at PFs different from the unity. Fixed PF control can be used to limit the injections of active power while providing a reactive power absorption or injection throughout the day. Reactive power output increases proportionally to the PF, as the PV active power increases [186]. Fixed PF is suboptimal by all means, since it leads to a continuous curtailment of active power and inefficient use of reactive power irrespective of grid conditions.

- **Power factor as a function of active power.** This control is based on measuring the active power output of the PV inverter, to then control the PF (providing reactive power compensation) when power output reaches a certain value e.g. 50% of rated power.

- **Volt-var (VV) control.** The VV control is a local controller that reacts to voltage measurements at the PCC of the solar inverter with the feeder [188]. VV control can be configured with setting different functional points and slopes for the curves, the existence, or lack thereof, of a "deadband", the possibility of including hysteresis, etc. This control is also called Q(V).

2.3.5.2.3   Combined controls

There is also the possibility to make active power curtailment and reactive power compensation simultaneously, combining VW and VV. As in the only VW control, this strategy is only useful when experiencing overvoltage situations, otherwise, it would behave as an only VV function. A priority of which function dominates over the other has to be defined

2.3.5.3   Global control strategies for voltage regulation

There is another control strategy that involves both the DSO and PV systems. Due to the intrinsic characteristics of distribution grids (listed in section 2.3.3), local inverter control strategies will typically lead to solutions far from the optimum point of operation. Reactive power usage and active power curtailment may be managed suboptimally and control actions may be distributed unequally among all the PV inverters in the system. Global or centralized controls, which take into consideration the global state of the grid and acts over all the different devices, will improve the voltage regulation.

59

The evolution of communication infrastructures and capabilities of PV systems promotes the introduction of such global optimization methods. Global optimization methods retrieve information from every system on the grid and analyze them in real time to find optimal operating points for different devices involved to maintain power quality [189]. The increased communication of the devices in the power grid is the critical enabler of global controls [190]. Also, there should be a cooperation between the DSOs and the PV owners in order to achieve an effective global control strategy.

There are multiple ways to achieve an effective global control, but in the literature the most common methods act over the PV inverters either by changing the setup of a local controller, or by commanding them to inject/absorb reactive power or curtail active power. Several examples can be found in [183], [191]–[195].

2.3.6 Metaheuristics for global control optimization problems

The possibility of including a global control for PV inverters in the distribution grid opens the door to find the best set of parameters for each PV inverter, that is, to optimize the control. As mentioned in the previous section, global controls receive the signals from all the controlled devices in the grid (in this case PV inverters) and send commands to each of them to regulate their parameters. The decision making process of which commands should be given to each controller, can be done by means of different approaches. Probably the most interesting is to optimize the response of all the devices in order to achieve global optimal operating conditions. In other words, global control can be converted into an optimization problem with a determined formulation dependent on the pursued objective. In the case of PV inverters in distribution grids, common objectives include minimizing the amount of active power curtailment, minimizing the use of reactive power compensation, preventing the voltage from falling outside the normative limits, etc. Different approaches to the problem will lead to different problem formulations.

2.3.6.1 Optimization problems and metaheuristics

Optimization problems are a family of mathematical problems that aims to find the best possible solution from all the feasible solutions for a particular function or problem. In general, optimization problems can be classified according to the nature of the variables being studied (discrete or continuous), the existence or not of constraints (constrained or unconstrained), mono or multi-objective, the nature of the problem itself (if it is linear, non-linear, quadratic, etc.) or if they are static or dynamic [196]. There are also several techniques to solve optimization problems: classic optimization algorithms, iterative methods, global methods or metaheuristics. Roughly speaking, they can be classified into exact methods and heuristics [197]. Heuristics are particularly interesting for real world optimization problems (also called hard optimization problems), since most real world problems are complex enough to be almost computationally intractable to approach with

60

exact or deterministic methods, and they usually cannot be solved to optimality, or to any guaranteed bound in reasonable time [196].

Metaheuristics are high-level frameworks of algorithms designed to solve a wide range of hard optimization problems with little or no adaptation needed to each problem [196]. They are usually applied to hard optimization problems for which there are no clear exact methods applicable or they are computationally intractable with them. They are generally nature-inspired, meaning that they try to mimic some natural or physical process, and they commonly use metaphors to name the mathematical objects used in the algorithmic routines.

For a metaheuristic to be successful for a given problem, it should have a balance between exploration (search in larger regions of the solution space) and exploitation (intense search along the most promising regions). Metaheuristics can be classified according to the number of simultaneous solutions they work with in single-solution metaheuristics, and population-based metaheuristics. As one may expect, the prior group favors exploitation while the latter favors exploration.

Single-solution methods, sometimes called trajectory methods, provide an initial guess for the solution, and then describe a trajectory along the solution space, often using the gradients of the function, to find the best direction to move, and the iterative process continues until some conditions are met. They are more vulnerable to getting stuck in local minima/maxima if there is no mechanism to prevent it, and if functions are non-convex. They also favor exploitation, since they will most likely find the best solution on a local scale, but they will be unaware of the existence of better solutions outside. Some examples of single-solution methods include simulated annealing (SA), tabu search and greedy randomized adaptive search procedure (GRASP) method

Population-based methods work with multiple solutions (a population) at the same time. This family of methods are highly inspired by Darwin's evolutionary theory, in which populations of individuals (solutions) must evolve in order to adapt to environmental conditions (optimization problem), and only the most valid individuals survive (best solutions). Population-based randomly initialize the population of solutions trying to cover the major part of the solution space. After evaluation, the population is modified through different operators, which will depend on the selected strategy. Often these modifications will have some random components, and some objective-guided components. Depending on the kind of strategy followed by the algorithm, population-based metaheuristics can be classified into:

– **Evolutionary Computation (EC)**. As the name suggests, they keep a major resemblance with the evolutionary theories, and they use modern terminology from biology (genes, mutations, chromosomes…). This category includes Genetic Algorithm (GA), Evolutionary programming and Genetic programming.

61

– **Swarm Intelligence (SI)**. SI techniques try to emulate the collective social behavior of several animal species, for example ant or bee colonies. One of the main features is that there is global information about the problem that is shared with all the colony members (population of solutions), guiding them to more favorable regions of the solution space. Some examples of SI algorithms are Ant colony optimization (ACO), Particle Swarm Optimization (PSO) or Bacterial foraging optimization, artificial immune system.

– **Other Evolutionary algorithms** that do not exactly lie in the previous groups can be found in the literature. Probably the most interesting are the ones using probability theory and statistics to deal with the reproduction process of solutions. Estimation of Distribution Algorithms (EDA), Differential Evolution or Co-evolutionary algorithms can be found under this description.

For extended information of metaheuristics, their origins and main applications, the reader is referred to [196]–[199].

2.3.6.2 *Metaheuristics applied to PV inverter global controls*

The problem of global optimized control for PV inverters is a rather complex method, particularly as the number of PV systems in the grid increases. With a high number of PV inverters and other regulation devices, the number of variables to be tracked and optimized increases, along with the complexity of the task of monitoring and control. Therefore, fast optimization algorithms are needed to find near optimal solutions in a reduced timeframe, and here metaheuristics play an important role.

Metaheuristics have been used in multiple renewable energy and power systems-related problems. For example, in [200], GA is used to determine the best location of hybrid PV-wind power plants, and in [201], another GA is used to optimize the operation of hydro-PV power systems. A combination of SA and PSO is featured in [202] to estimate the optimal capacity of different renewable energy power systems. In [203] four different EC algorithms were used to optimize the design of grid-connected PV systems.

There is a vast quantity of studies that have also applied metaheuristics to solve some form of optimization problems related to PV inverters or other devices in distribution and smart grids. For instance, in [204] Non-dominated Sorting Genetic Algorithm (NSGA) is used to solve voltage control in a grid with high DG. In [205], a modified ACO is proposed for solving nonlinear voltage and reactive power control problems. A PSO for reactive power and voltage control is presented in [206]. In [207], a modified version of PSO, named chaotic PSO, was used to adjust generators, transformer OLTCs and compensators to minimize power losses on a grid under high renewable penetration. In [208], yet another modification to the PSO algorithm, mixing it with fuzzy logic, is used to relieve the overvoltage caused by high PV penetration minimizing total line loss. In [209], a "big bang-

62

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

big crunch" optimization method (a modified GA) is used to minimize voltage deviations by controlling the regulator taps and reactive power contributions of capacitors and DG. Differential evolution to solve the optimal reactive power dispatch is presented in [210], [211]. In [212] an EDA is used to optimize the charge patterns of a large fleet of electric vehicles.

The list of studies utilizing metaheuristic to solve grid-related problems is huge, and only its analysis probably would require the dedication of an entire thesis to do it. For extensive reviews of metaheuristic applied to power system problems, the reader is referred to [213], [214].

It seems that it is easy to get lost in the whole vocabulary of metaheuristic techniques, where combination of several approaches lead to a list of acronyms impossible to quantify or classify, or where dressing classic metaheuristic with other metaphors or words lead to "novel" metaheuristics, which are basically the same as other well-established heuristics from a functional point of view. As warned in [197], where a profound reflection and critique to the proliferation of new metaheuristic is exposed, this plethora of different techniques searching for the novelty may represent a distraction from actual research, and it also generates suspicion among researchers. Nevertheless, if all the competition for novelty is put aside, metaheuristics prove to be a very useful tool to solve complex optimization problems, such as power systems related ones, in acceptable times. These kind of approaches have more value with the ever increasing addition of devices to the electric grid, conforming what are called smart-grids.

2.3.7 Conclusions and future prospects

In this section, the main impacts concerning the deployment of PV systems in the electric grids have been reviewed. As it has been shown, the integration of PV systems on electric grids and more specifically, in distribution grids, is not exempt of complexity. Solar variability poses a high stress over the grid, affecting the functional and quality parameters of electric grids (frequency, voltage), thus increasing the operational complexity of the system. As the number of PV systems continues to increase, the probability of these events to happen will increase as well.

However, simultaneous advances in other areas such as power electronic devices and PV inverters, improved communication infrastructures, integration of energy storage systems and forecasting systems, etc., will help to cope with the issues arising from a massive deployment of PV systems. Taking advantage of the new capabilities of modern devices at a large scale will mean the switch from traditional electric grids with a passive mode of operation, to more interconnected, intelligent and optimized grids (the so called smart-grids).

2.3 Integration of PV systems in distribution grids

An increase in the number of intelligent devices in the grid means that the number of variables to be tracked and controlled will shoot up as well. For that reason, fast responding optimization algorithms will be required to handle complex problems with large number of variables in almost real-time, in order to ensure optimal operation of the system. Metaheuristics are a very promising (although not exempt of controversy) set of optimization algorithms that are able to handle hard optimization problems (real world problems) of high complexity in acceptable times.

64

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 3.   Research

In this chapter, the papers composing the compendium for the thesis are summarized.

3.1   A simple big data methodology to PV specific yield estimations

3.1.1   Reference

3.1.2   Abstract

In this work, we have reviewed recent literature concerning the performance of PV plants in different power systems and detected that, usually, the raw data used is not complete in terms of the number of PV plants and also main parameters of poor quality need to be removed. Then, for the first time, a study of the specific yield of all PV plants in a power system with single-PV-plant resolution is presented. Thus, we have analyzed the official 237,588 monthly energy values obtained from 2005 to 2017 for the 1523 PV plants in the Canary Islands. This dataset is obtained from PV plants ranging from 0.53 kWp up to 9 MWp, and it has been supplied by the distribution system operator and main utility (ENDESA). Then, this dataset is compared to 153,120 irradiance and temperature data from a PVGIS database. Results show that the Spanish regulation has a direct effect not only on the development of the PV capacity in the Canary Islands, but also on the specific yields. Moreover, only combining meteorological data (irradiance, temperature and wind

65

speed) from satellites, starting year of operation, and nameplate capacity we have developed a very simple theoretical model to predict the specific yield of a PV plant at any location in the Canary Islands, avoiding the requirement of any data from the owners of the PV plants. The simulation values obtained have been validated with the real specific yields for PV plants assumed to be well managed (multi-MW power plants placed in best locations) showing errors below a 3%. This theoretical model has also been used for detecting suboptimal PV plant designs and anomalous specific yield of PV plants above the clear sky limit. Recommendations to avoid anomalous specific yields in future are included.

3.1.3 Research objectives

The main objective of the paper was the development of a simple model to estimate specific yield of PV power plants in the Canary Islands, using data readily available for the utilities and other entities interested in the resulting specific yields.

Along the main objective, a series of side goals were pursued during this research:

– Study the influence of the legislation and regulations on the installation rates and management of PV plants

– The difference in performance between roof-mounted PV installations and ground-based installations.

– Compare the quality of the different available solar irradiance databases for complex microclimatic and orographic conditions in Canary Islands

– Study the performance's dependence on the capacity of the installations in terms of specific yield

– Study the dependence of modules from different manufacturers on the specific yields of plants within the same municipalities (similar irradiance values)

– Estimate the slope of the PV modules using the simple model

– Study failing injection or disconnection rates compared to the capacity of the plants

Whereas there exists a considerable number of methods to obtain the specific yield of PV plants, the objective of the paper was to obtain a very simple model to quickly estimate the theoretical specific yield of a PV plant given only readily available data. In the article, a big data procedure is described to develop the model, which could be easily used with data held by the electric utilities. The model could be used either by the utilities or the governing organs in terms of preventing frauds or irregular situations with some of the installations.

66

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 3: Research

3.1.4 Methodology

Data provided by the utility Endesa, containing 237,588 monthly records of smart meters for all the renewable energy plants in the Canary Islands, was filtered to get only PV data and to guarantee the consistency of the values, excluding outliers (which were mainly zeros due to technical problems or administrative issues) and removing duplicates. The same dataset contained the geographic location of the plants, the nameplate capacity and the voltage level of the connection to the electric grid. With the geographical information, solar irradiance, temperature and wind speed were obtained for solar and weather databases. It was important to maintain a high spatial resolution on the weather and solar databases since Canary Islands, due to their unique location and particular orography, have an important number of microclimatic conditions that can change substantially between close locations. In order to select the best solar database, a simple procedure of comparison between the different solar datasets was performed, discarding the ones with the higher errors in comparison with real data.

The developed model makes use of the abovementioned sources of data, that is, nameplate capacity, smart meter readings, geographical location and weather variables (mean temperature, average wind speed, irradiance values...) which are then fed to our simple model, where a set of analytical and empirical formulas are used to estimate the specific yield of the PV plants. The data requirements for the model were minimized to match data that could be obtained easily from the utilities. An extensive description of each calculation done by the model is given in the article, but a summary of the calculations is given below.

- **Temperature dependence on PV efficiency.** First, the temperature dependence on PV efficiency is obtained by:

$$\eta_{mpp}(G_{mod}, T_{mod}) = \eta_{mpp,25}(G_{mod}) \cdot [1 + b(T_{mod} - 25\,°C)]$$

where b is the temperature coefficient of silicon PV modules (approximately -0.0035 ºC), G_{mod} is the irradiance on the plane of the module, and T_{mod} is the module temperature. Module temperature can be approximated by:

$$T_{mod} = T_{amb} + c \cdot G_{mod} - a \cdot v^{-1}$$

where T_{amb} is the ambient temperature, c is the Ross coefficient, a is an empirical parameter that accounts for the influence of wind on the module temperature, and v is the wind speed at the height of the PV modules (z), that can be calculated using the following equation:

67

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

3.1 A simple big data methodology to PV specific yield estimations

$$v = v_{ref} \frac{\ln\left(\frac{z}{z_0}\right)}{\ln\left(\frac{z_{ref}}{z_0}\right)}$$

where v_{ref} is the wind speed at the reference height z_{ref}, and z_0 is the roughness length in the wind direction.

– **Degradation rates and other losses.** To estimate the degradation rate of PV modules, an average 0.5% loss of rated power per year of operation is considered. Other losses related to PV inverters efficiency and other miscellaneous losses are considered as 9.5% of the model output. Overall losses can be introduced in the calculations by a correction factor (L), defined by:

$$L = 0.905 \cdot [1 - 0.005 \cdot (Y_i - Y)]$$

where Y_i is the year of study for which the PV yield is being determined, and Y is the installation year of the PV plant.

For introducing the influence of improving technology over the period of study on the module efficiency, a linear approximation among the average efficiency of commercial silicon modules over the period of study (from 12% in 2005 to 17% in 2017) is made:

$$\eta'_{mod} = (Y - 2005) \cdot 0.417 + 12$$

$$\eta_{mod} = \frac{\eta'_{mod}}{100}$$

– **Estimation of the specific surface in terms of module efficiency.** For this calculation we have obtained a relation between efficiency (in %) and surface (in m2) considering the record efficiency for large silicon modules measured under the global AM1.5. A polynomial regression is made, obtaining the following expression for the estimation of the surface:

$$S = 131.2 \cdot \eta'_{mod}{}^2 - 78.379 \cdot \eta'_{mod} + 15.423$$

– **Specific yield of PV plant (kWh/kWp).** Finally, the PV specific yield can be determined using the following expression:

$$E_i = G_x \cdot \eta_{mpp} \cdot \eta_{mod} \cdot S \cdot L$$

68

3.1.5 Results

The main objective of the paper was to define a simple procedure for the estimation of specific yield of PV plants. The goal was achieved, producing a very simple model that gets deviations from the real values of around 2% on average. All the errors are below 3%.

Data showed that deviations from real values were larger starting in 2015. These deviations experienced in the last years of the study, where the real outputs are quite below the predicted values by the model, coincide with changes in the retribution system implemented in 2014 (IET 1459/2014) [215]. With the new regulation, a larger share of the retribution is based on a minimum amount of hours producing energy, beyond which, the retribution does not apply, thus discouraging the cleaning of the PV modules. At least in one of the plants exhibiting this behavior, the cease in cleaning labors starting in 2015 has been confirmed.

Regarding regulation and policies, it is observed an evident impact of the different regulations approved on the installation rates of PV plants in the Canary Islands, where beneficial politics enabling feed-in-tariffs lead to higher installation rates, while restrictive policies that suspend the retributions decreased installation rates and the attention of the investors in the sector.

Additionally, some other interesting results were derived from the analysis. First, multi-MW plants, which are supposedly to be well managed in terms of operation and maintenance, have specific yield values very close to the clear sky limit defined for their municipalities. This lines up with what is expected from the initial hypothesis. This also provides the model with the additional functionality of evaluating the performance of any PV system given its location and age. The specific yields of these large PV plants could also be used as references for calibrating the overall procedure.

Second, the model can be used to determine the slope of the PV modules of the PV plants. This can be done by minimizing the difference between the real values and the simulated values, solving for the slope angle. As the irradiance at a defined slope is obtained directly from the solar databases and is not coded in the model, an iterative process could be used for this purpose. Although this could highlight inefficiencies in the design of the plants, one of the expected impacts of a large share of PV in the electricity generation is the decrease of the prices in months with higher irradiance, and therefore, the plants could be designed to maximize the output in months with lower irradiance, since the retributions will be more attractive.

Regarding the failure rates in terms of capacity, it is observed that smaller PV plants with lower capacities tend to fail more injecting electricity to the grid, than the larger ones. This could be explained by the fact that larger PV plants will be better managed (in terms of operation and maintenance), and also, malfunctioning of a single component of the

plant will not result in a complete zero of the output, but rather a minimal reduction. In small plants, that is not the case, since malfunctioning of any equipment could imply a total disconnection of the installation.

3.1.6 Conclusions

The simple model for the estimation of specific yield of PV plants in the Canary Islands, using only location, weather and nameplate capacity has been validated, obtaining good approximations to the real performance of well managed plants. The model could be used to prevent unrealistic specific yields (and possible frauds derived from it), to determine the lack of cleaning or defective maintenance of PV plants, and also could be used for the estimation of the slope of the PV modules.

It has been demonstrated that policies and regulation has a huge impact in terms of both new installation rates and maintenance operations. Changes in the fed-in-tariffs for the PV systems have had a huge impact in both the installation rates, by increasing the uncertainty of the investments (and pulling away the risk-averse investors), and also the performance of the already deployed plants, by discouraging optimal management of the plants (lack of cleaning). A set of stable policies and regulations is required in order to ensure that the future deployment of PV installations is not slowed down, if the Canary Islands want to achieve their objectives for the energy transition.

A very simple alert system implementation, based on the clear sky values presented in this study as theoretical maximum specific yields for plants depending on their location and configuration, could be adopted by the utilities or by the Spanish Markets and Competency Commission (CNMC) to avoid fraud and other irregular situations among the electricity producers.

70

3.2 Validation of sky-imagers and intra-hour solar forecasting

3.2.1 Reference

3.2.2 Abstract

Increasing photovoltaic (PV) generation in the world's power grid necessitates accurate solar irradiance forecasts to ensure grid stability and reliability. The University of Texas at San Antonio (UTSA) sky-imager was designed as a low cost, edge computing, all-sky imager that provides intra-hour irradiance forecasts. The sky-imager utilizes a single board computer and high-resolution camera with a fisheye lens housed in an all-weather enclosure. General Purpose IO pins allow external sensors to be connected, a unique aspect is the use of only open source software. Code for the sky-imager is written in Python and calls libraries such as OpenCV, Scikit-Learn, SQLite, and Mosquito. The sky-imager was first deployed in 2015 at the National Renewable Energy Laboratory (NREL) as part of the DOE INTEGRATE project. This effort aggregated renewable resources and loads into micro-grids which were then controlled by an Energy Management System using the OpenFMB Reference Architecture. In 2016 a second sky-imager was installed at the CPS Energy micro-grid at Joint Base San Antonio. As part of a collaborative effort between CPS Energy, UT San Antonio, ENDESA, and Universidad de La Laguna, two sky-imagers were also deployed in the Canary Islands that utilized stereoscopic images to determine cloud heights. Deployments at three geographically diverse locations not only provided large amounts of image data, but also operational experience under very different climatic conditions. This resulted in improvements/additions to the original design: weatherproofing techniques, environmental sensors, maintenance schedules, optimal deployment locations, OpenFMB protocols, and offloading data to the cloud. Originally, optical flow followed by ray-tracing was used to predict cumulus cloud shadows. The latter problem is ill-posed and was replaced by a machine learning strategy

with impressive results. R2 values for the multi-layer perceptron of 0.95 for 5 moderately cloudy days and 1.00 for 5 clear sky days validate using images to determine irradiance. The sky-imager in a distributed environment with cloud-computing will be an integral part of the command and control for today's smart-grid and Internet of Things

3.2.3 Research objectives

The main objective of the research was the development of a low-cost intra-hour forecasting system that could compete with commercial sky imaging-based forecasting technology, which often are too costly and have proprietary software. In the literature, several attempts to achieve so are found. The intention was to further decrease construction costs, taking advantage of the improvements on microprocessor capabilities and decreases in costs of electronic components, while improving performance of the devices. With this objective in mind, a series of sky-imager prototypes were designed, built and deployed in conjunction with the University of San Antonio, Texas (UTSA).

Therefore, the goal for this paper is to validate the performance of the early prototypes of the sky-imagers developed on the framework of the thesis, in three different locations (two in the USA and one in Spain), comprising three different micro-grid architectures. In the paper, details about each location are provided, highlighting the main differences between them, as well as the particularities, difficulties and barriers encountered during installation and operation of the sky-imagers.

A series of side objectives were pursued as well:

− Design and construction of the prototypes. The idea was to build inexpensive devices to enable the massive deployment of sky-imagers as key components of future micro-grids.

− Evaluation of the different approaches taken at each location, and the different functionalities of each prototype

− Evaluation of extra functionalities of the devices: weather data acquisition, cloud base height estimations, irradiance data directly from the images...

− Test the capabilities of different Machine Learning and Deep Learning algorithms for the intra-hour solar forecasting task.

− Perform suitability and durability tests for the prototypes and components, and identify the aspects to improve in future versions of the device, looking towards a commercial product.

− Analyze the interactions between de sky-imagers and the different micro-grids architectures present in each location.

72

3.2.4 Methodology

During the first months after the deployment of the sky-imagers, a series of tests were made to identify the critical aspects to take into account in later stages. In one of those tests, it was found that the correlation between ramps in GHI measurements and the presence of clouds occluding the sun in the images could be learned by AI models. In another test, the relationship between PV output and GHI turned out to be almost linear, as expected from theory, which leads to think that forecasting GHI indirectly leads to good predictions in PV power.

The early prototypes of the sky-imager were composed of a series of inexpensive components, but each of the prototypes deployed had slightly different characteristics and components. While the core of the devices was basically the same, different components were added to each prototype to expand their functionalities (weather data acquisition, cloud base height estimations...), and several strategies were tested in each location (ray tracing, machine learning, elliptical trajectories...). A detailed description of the three prototypes highlighting their differences is provided in the paper, but a brief summary is presented below.

- **National Renewable Energy Laboratory (NREL).** The NREL prototype had two microprocessors: an Odroid C1 for most of the computation, and a Raspberry Pi 2 for the image acquisition.

- **Joint Base San Antonio (JBSA).** At Fort Sam, the sky-imager internals were the same as in NREL, with the particularity that it was deployed in conjunction with a meteorological tower which had a Kipp&Zonen CMP11 pyranometer and a Vaisala WXT520 weather transmitter.

- **La Graciosa.** In La Graciosa, two prototypes were deployed in order to estimate the cloud base height by means of stereographic approaches. The build was based on a single Raspberry Pi 3 model B which handled the image acquisition, preprocessing and transmission to the server, where the algorithms were run with the inputs from both images.

The methods employed at each location were site-specific, since different micro-grids architectures and climatic conditions were present. Despite that, there are some parts of the overall process that are shared by all the prototypes. In general, the forecasting process can be described with the following steps:

- Image acquisition. All the sky-imager used a similar algorithm to capture the images. As one may expect, the luminic sky conditions are constantly changing during the day, therefore different settings should be used in different moments of the day. Given the inability to control the aperture of cheap cameras, the only way to control the exposure of the images is by setting the shutter speed. To

73

generate a balanced image, a High Dynamic Range (HDR) procedure was used to capture the images, which takes images at different exposures (different shutter speeds) and merges them together, reducing overexposure in the circumsolar region of the image.

– Image processing. The image processing pipeline is almost identical in all the devices.

  a. First the images are undistorted to remove the effect of the fisheye lenses.

  b. Second, the images are cropped and masked to remove any static obstacle that could be in the image (such as buildings, trees, power towers, etc.).

  c. Cloud decision algorithm. In order to segment the clouds in the image, a red to blue ratio operation is applied. This step could be slightly different in all three prototypes, but the overall definition is the same. Basically it consists in dividing the red channel of the images over the blue channel of the images. The idea is that, since clear sky is normally blue in the pictures (high values in the blue channel), and clouds are normally white (high values in all three channels), the ratio of the red to blue channels would lead to a different representation where clouds have high values and clear sky areas have low values, thus enabling an easy segmentation of the clouds.

  d. Optical flow to analyze the movement of clouds between consecutive frames. There are two main techniques applied in this procedure, one is the Lucas-Kanade method, that aims to identify the most valuable features to track over the images, and the other is the Dense Optical Flow. Both methods were applied in different experiments.

– GHI calculations. GHI calculations can be done either in the sky-imager or in a dedicated server. Both strategies were tested, since in the case of La Graciosa, double input from two images made it more convenient to use a server, while in the locations in the USA, especially in JBSA, edge-computing architecture seemed more suitable due to cybersecurity constraints of the micro-grid. Several approaches were taken to make the GHI forecasts. A set of 4 machine learning algorithms, including Multi-layer perceptrons (MLP), Random Forest (RF), Gradient Boosted Trees (GBT) and a Deep Learning (DL) architecture, were tested to forecast GHI values in the locations of the USA. The work was later expanded with the inclusion of four different DL architectures. In all of the tests, input images were heavily downsampled to enable the computation of the ML models on the microprocessor, and only the circumsolar region of the images was used (instead of the whole image), reducing even further the computational requirements for the microprocessors.

74

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 3: Research

– Cloud Base Height (CBH) estimations. In La Graciosa Island, a stereographic method was implemented to estimate CBH from paired sky images. The proposed method was purely geometric-based, using the relative position of the sky-imagers and the position of the clouds in the images, along with the projection matrix of the lenses utilized. First, an algorithm called Scale-Invariant Feature Transform (SIFT) is used to identify the useful features in simultaneous images from both sky-imagers. This algorithm is capable of identifying the same feature on two different images, even if transformations, rotations or changes in luminosity are present. After applying a filter operation with a determined threshold, the best features are selected to continue with the CBH estimation process. Valid features are then transposed from the image space (pixels) to the real space coordinates (azimuth and zenith). Geometric computation is then applied to obtain the length of the distance vectors (from the camera to the feature), and a distribution of CBHs is produced as a result.

For each of the previous steps, there are different approaches that could be taken, which are detailed in the paper.

3.2.5 Results

The first analysis comparing the four different machine learning models selected did not achieve great results, yet they provided valuable insights about the direction to take for the next tests. The models were fed with heavily downsampled images (8x8 pixels) of the circumsolar region of the original image. All of the models had over 20% of normalized Root Mean Squared Error (nRMSE), being the DL architecture and the GBT, the best models with nRMSE of 21.6 and 21.1% respectively. MLP and RF exhibit nRMSE values of 33.05% and 29% respectively.

After analyzing the preliminary results, it was decided to continue the tests with DL architectures, since it was concluded that they had more room for improvement (in terms of different architectures that could be used and hyperparameters that could be tuned to achieve better results). Four different fully connected models with different hyperparameters and increasing complexity were tested. Error metrics for all the models were diminishing with the increasing complexity, until a point of diminishing returns was reached. However, nRMSE were not far from those achieved with other ML models (all above 21.6%), and improvements were only observed in the Mean Absolute Error (MAE). These results lead to the conclusion that further improvements in forecasts would require using larger input images, changing in architecture and layers of the models, or fine-tuning the hyperparameters.

Another conclusion drawn was that, since the outputs for clear sky and cloudy days are quite different, maybe two different models for each situation could be trained

75

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

independently to achieve better results. The next step was to put that hypothesis to test. The four ML models previously used were fed with larger images (32x32), with different datasets for partly cloudy and clear sky days. MAE for all the models decreased considerably, primarily due to the increase of the number of input features (from 64 to 1024). Also, a clear reduction in the errors of clear sky days compared to cloudy days (5 times on average) was observed, as expected.

Results from CBH estimations from La Graciosa prototypes were compared with LIDAR measurements from a weather station 30 km south of the location of the cameras. This obviously has a significant effect on the comparison of results. Moreover, the different nature of the measurements (where LIDAR is punctual and SI covers a larger area) affects the comparability of results as well. The main conclusion that can be drawn from this experience is that the estimations are coherent with previous knowledge of the atmospheric conditions of the region.

Other valuable lessons were learned from the different locations. From the experience in JBSA, it was clear that the sampling frequency of the data provided by the micro-grid management system (15 minutes) is not enough to capture the effect of ramps in the irradiance. Data with a sampling frequency of 1 minute was obtained with the SI using inexpensive components attached to the microprocessor, enabling the collection of high resolution data with low cost equipment. In La Graciosa, high resolution data was also obtained using a cheap off-the-shelf mini PV module and a current sensor, from which irradiance data was derived.

3.2.6 Conclusions

The use of low cost sky imagers has been validated to quality produce solar irradiance forecasts in a variety of settings and applications. Improvement in the quality of the forecasts could be achieved considering several aspects of both the devices and the processing pipelines. On the one hand, weather and environmental factors have an effect on the durability of the devices (high temperatures in Texas damaged some components and proximity to the sea in La Graciosa led to water infiltrations in the enclosure, despite its IP67 degree of protection). Atmospheric dust depositions in La Graciosa, which are very common in the area due to its proximity to the Sahara desert, severely affected the quality of the images, evidencing a need for cleaning the devices after strong events of atmospheric dust.

On the other hand, several improvements can be made in the overall processing pipeline, from image preprocessing to architecture of the DL models. Different architectures with different layers should be tested to identify the steps with more improving potential. For instance, image processing could be done with DL algorithms,

76

using convolutional layers, or time-series predictions could be implemented with the use of recurrent layers.

In general, although physical methods will remain to be important in the coming years, data-based methods seem to open a new perspective on the use of sky-imagers for intra-hour solar forecasting. The utilization of hybrid models is expected to dominate the scene in the coming years, by generating models that brings the best from both perspectives: the theoretical knowledge and formulations from physical based models, and the power and versatility of data-driven techniques.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

3.3 Optimized global control for PV inverters in distribution grids

3.3.1 Reference

- **Authors**: David Cañadillas, Hamed Valizadeh, Jan Kleissl, Benjamín González-Díaz, Ricardo Guerrero-Lemus.

- **Full title**: "EDA-based optimized global control for PV inverters in distribution grids".

- **Journal**: *IET Renewable Power Generation.* 2020; 1– 15.

- **DOI**: https://doi.org/10.1049/rpg2.12031

3.3.2 Abstract

Operating distribution grids is increasingly challenging due to the increasing penetration of photovoltaic systems. To address these challenges, modern photovoltaic inverters include features for local control, which sometimes lead to suboptimal results. Improved communication infrastructure and photovoltaic inverters favor global control strategies, which receive information from all the systems in the grid. An estimation of distribution algorithm is used to optimize a global control strategy that minimizes active power curtailment and use of reactive power of the photovoltaic inverters, while maintaining voltage stability. Optimized global control outperforms every other local control evaluated in terms of apparent energy used for control (9.9% less usage compared to the second best alternative in all scenarios studied) and ranks second in terms of voltage stability (with a 0.14% of total time outside the voltage limits). Two new indicators to compare control strategies are proposed, and optimized global control strategy ranks best for both efficiency index (0.98) and average apparent power use (0.48 kVA).

3.3.3 Research objectives

The major goal of this research is to design and implement an optimization algorithm for global control of distributed PV inverters in distribution grids. Reviewing the literature for the use of different metaheuristics and its applications in the power field, it was seen that Estimation of Distribution Algorithms (EDA) were not widely applied for any power system related problems at all (only one reference dealing with the optimization of Electric vehicle charging were found), and it was not found any research using this metaheuristic to control the behavior of PV inverters in distribution grids. The following additional objectives were identified:

78

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 3: Research

- Study the viability of using PV inverters as the only source of voltage regulation in the distribution grid (beyond the transformer at the substation)

- Comparison and evaluation of the different strategies currently available in commercial PV smart inverters for voltage regulation (reactive power compensation)

- Development of two metrics to assess the overall performance of control methods for distributed generation in distribution grids, namely Effective index and Average apparent power use.

- Benchmarking of the EDA-based optimized control against other metaheuristics: Genetic Algorithms (GA) and Particle Swarm Optimization (PSO).

3.3.4 Methodology

Before the evaluation of the optimization algorithm, an optimization problem is formulated. The objective function encoded the maintenance of the voltage levels within regulatory limits, while minimizing power curtailment and reactive power usage. The optimization algorithm is computed every 15-minutes. Its inputs are the active power of every PV system in the distribution grid. The outputs are a series of splines (two per PV system, corresponding to active and reactive power) composed of 11 points, one for each decimal percentage from 0 to 100%, of possible active power over the rated power of the PV inverter. The idea of finding splines for each PV inverter (instead of a set of fixed values) is to account for the output variations in PV power and the possible combinations for the next 15-minutes until the next optimization cycle starts.

The optimization function is subject to several constraints. Constraint functions account for physical limits of the devices (maximum rated power of PV inverters), logical constraints (curtailed power cannot be higher than actual active power) and voltage limits. Constraints were modelled with penalty functions, meaning that large values will be added to the cost functions when a constraint is violated. With this approach, the algorithm has more degrees of freedom to search for feasible solutions, instead of getting stuck in a suboptimal region of the solution space finding little improvement at each iteration. With hard constraints, computational time increases considerably. If hard constraints are imposed, all the populations must be filled with feasible solutions, and in the first iterations, the distributions to sample from are too wide. This means that the chance of sampling thousands of individuals that satisfy all the constraints at once is quite low, so the sampling process should be repeated until all the population is filled with feasible solutions, leading to a bottleneck in the computational pipeline.

To formulate the optimization problem, first the constraints need to be defined. The first constraint accounts for the use of total active and reactive power that cannot exceed the rated apparent power of the PV inverter.

$$(P_{p,k}^{PV} - P_{p,k}^{curt})^2 + \left(Q_{p,k}^{PV}\right)^2 \leq \left(|S_p^R|\right)^2$$

where $P_{p,k}^{PV}$ is the actual PV active power produced by the PV system; $P_{p,k}^{curt}$ is the curtailed active power; $Q_{p,k}^{PV}$ is the reactive power absorption/injection.

The second constraint takes into consideration that the curtailed power cannot exceed the actual active power being produced:

$$P_{p,k}^{curt} \leq P_{p,k}^{PV}$$

The most important constraint (and the most difficult to satisfy) deals with voltage limits, and it is defined by:

$$|V^{min}| \leq |V_{n,k}| \leq |V^{max}|$$

where $|V_{n,k}|$ is the voltage magnitude of each node n of the circuit at every power output level k and $|V^{min}|$ and $|V^{max}|$ are the normative limits defined in the ANSI C84.1 norm ($|V^{min}| = 0.95\ p.u.$ and $|V^{max}| = 1.05\ p.u.$)

Running a complete power flow for each candidate solution and for each knot in the splines would be computationally infeasible. For that reason, voltage constraints were evaluated with a linearized power flow. The linearized version of the power flow is calculated by:

$$\left|V_{i,k}(\bar{x})\right| = \left|V_i^{base}\right| + \sum_{n=1}^{n_{nodes}} \left(s_{i,n}^P \cdot \left(P_{p,k}^{PV} - P_{p,k}^{curt}\right) + s_{i,n}^Q \cdot \left(Q_{p,k}^{PV}\right)\right)$$

where $\left|V_{i,k}(\bar{x})\right|$ is the voltage magnitude in node i at knot k for the solution $\bar{x}$. $\left|V_i^{base}\right|$ is the voltage magnitude of node i excluding the effect of PV systems, and $s_{i,n}^P$ and $s_{i,n}^Q$ are respectively the sensitivity of voltage magnitude at node i to active and reactive power injections/absorptions at node n. Sensitivity matrices are obtained using the perturbation-observation method, by making a small change in network state and measuring the effect of that change. The linearized version was evaluated against the Newton-Raphson method, giving errors of less than 0.2%, which indicates that it is a good model to estimate the voltage magnitude.

80

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 3: Research

Finally, the final optimization function is defined by:

$$\min_{Q_{p,i}^{PV}, P_{p,i}^{curt}} \sum_{k=1}^{n_{knots}} \sum_{p=1}^{n_{PVs}} \left(w\left(Q_{p,k}^{PV}\right)^2 + (1-w)P_{p,k}^{curt} \right),$$

$$\text{subject to} \qquad (P_{p,k}^{PV} - P_{p,k}^{curt})^2 + \left(Q_{p,k}^{PV}\right)^2 \leq \left(\left|S_p^{Rated}\right|\right)^2,$$

$$P_{p,k}^{curt} \leq P_{p,k}^{PV},$$

$$\left|V^{min}\right| \leq \left|V_{n,k}\right| \leq \left|V^{max}\right|,$$

where w is a weighting factor which has been set to 0.01 to favor the use of reactive power over active power curtailment. For a more detailed explanation of the distance function applied to each constraint, the reader is referred to the original Paper.

The simulations were run on a slightly modified version of the IEEE-123 node test feeder, where 9 PV systems of 1,000 kVA each were added, and voltage regulators and shunt capacitors were disabled. The optimization problem is solved for 6 scenarios, resulting from the combination of 3 irradiance curves (corresponding to clear day, partly cloudy day and overcast day) and 2 load curves (minimum and maximum demand). Power flows are performed every minute, and the optimization algorithm calculates the splines every 15 minutes (reducing computational costs). The splines obtained at each optimization, are then used for the 15 next minutes. The PV inverters adapt to the current irradiance and load conditions at each minute following the splines output by the optimization routine.

The optimization problem is solved using a modified version of the metaheuristic Estimation of Distribution Algorithm (EDA). The basics on the metaheuristic can be summarized as follows. First, the initial population is sampled randomly, trying to cover large parts of the solution space. Then, the initial population is evaluated and sorted. The individuals with the best score (or fitness) value, are then selected to build the new distribution from which the next generation will be sampled. At this point, some tweaks could be added to slightly randomize the process again, although in our case it was not done. Finally, the generation, evaluation, selection and distribution build processes are repeated until either the maximum number of generations set or the minimum fitness value is reached. The pseudo-code for the implementation of the algorithm is described below.

81

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

3.3 Optimized global control for PV inverters in distribution grids

---

Algorithm 1. UMDA$_c$ for the optimization problem

1	$\boldsymbol{P_0} \leftarrow$ Generate a random initial population P, of individuals $\bar{x} = (x_1, x_2, x_3, \dots, x_n)$, where n is the number of variables ($n_{knots} \times 2n_{PVs}$), according to a multivariate normal distribution with mean μ_0 and covariance matrix Σ_0 such as:

$$X \sim \mathcal{N}(\mu_0, \Sigma_0), \qquad \mu_0 = \boldsymbol{0}$$

$$\Sigma_0 = \begin{bmatrix} \sigma_{0,i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{0,j}^2 \end{bmatrix}$$

$$\sigma_0^2 = large\ constant\ value$$

2	For $\boldsymbol{g = 1, 2, 3} \dots$ in number of generations:
2	$\boldsymbol{P_{g-1}^{Se}} \leftarrow$ Evaluate population, sort them according to their fitness value and select Se best individuals
3	$\boldsymbol{p_g(x) = p(x|P_{g-1}^{Se}) = \mathcal{N}(\mu_g, \Sigma_g)} \leftarrow$ Build the probabilistic model: estimate the probability distribution of an individual being among the selected individuals, where the mean and the covariance matrix are computed as follows:

$$\mu_g = \bar{X} = \frac{1}{N} \sum_{r=1}^{N} x_{i,r}. \quad i = 1, \dots, n$$

$$\Sigma_g = \begin{bmatrix} \sigma_{i,i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{j,j}^2 \end{bmatrix}$$

$$\sigma_{i,i}^2 = \frac{1}{N} \sum_{r=1}^{N} (x_{i,r} - \bar{X}_i)^2 \quad i = 1, \dots, n$$

4	$\boldsymbol{P_g} \leftarrow$ Sample a new population from $p_g(x)$

82

Chapter 3: Research

Figure 18. Description of the selection process for the operational points for each PV inverter at each time step the simulation is run.

Further explanations about the functionality of both the metaheuristic and the overall simulation process, can be found in the paper. In figure XX, a diagram of the workflow of the overall process is presented in Figure 18:

The EDA-Optimized Global Control is compared with the methods available in current smart inverters (fixed power factor, Volt-var, Volt-Watt, both Volt-var and Volt-Watt simultaneously...). To benchmark the performance of EDA against similar strategies, a series of metaheuristics, namely Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) were applied to the same problem. All the metaheuristics were coded using Python programming language, and the grid related calculations were made using OpenDSS, a dedicated software for the resolution of power flows in electric grids.

To comprehensively evaluate the efficiency and efficacy of the proposed global control, two indicators were proposed to find out how good the regulation was for each method, by linking the total energy used for control (both curtailed active power and reactive power) with the resulting voltages over all scenarios. The first indicator, average apparent power usage ($\dot{S}_{avg}$) estimates the average apparent power needed for successful voltage regulation, and it is defined by:

$$\dot{S}_{avg} = \frac{S_{control}}{t_{V good}}$$

where $S_{control}$ is the total amount of apparent power used by each control in kVAh and $t_{V good}$ is the number of hours with voltages within the accepted range achieved with that same control. The second indicator, the effectiveness index (EI), quantifies both how effective the control is relative to the usage of apparent power, and its success in voltage regulation. EI is defined as:

$$EI = \left(1 - \frac{S_{control}}{S_{PV}}\right)\left(\frac{t_{V good}}{t_{total}}\right)$$

where S_{PV} is the total available apparent power (i.e. the sum of the inverter rated powers), and t_{total} is the total number of hours evaluated (6 scenarios x 24 hours). The effectiveness index ranges from 0 to 1 and a value of 1 indicates maximum effectiveness.

3.3.5 Results

Results for the simulations of all 6 scenarios are presented in the paper. In general, EDA-optimized global control outperformed every naive method for voltage control already available in the PV inverters. Only a Volt-var without deadband function achieved

83

better results in terms of voltage violations, but at the cost of using about 7 times more reactive power in maximum load conditions, and 17 to 40 times more reactive power in minimum load scenarios. In general, OptGC makes a better and more efficient use of the apparent power. *EI* for OptGC is 0.98 while $\dot{S}_{avg}$ is 0.48, being the first among all methods in both rankings.

When comparing with the other metaheuristics, EDA converges rapidly over the first generations, reaching a plateau after generation 100, and it achieves a smaller fitness value than PSO and GA. Regarding the percentage of voltage violations for each method, it can be seen that, although GA and PSO obtain better results, this comes at the cost of a 10 times more curtailed active power and 37 times more reactive power usage. The small improvement for GA and PSO in the voltages violations (less than a 0.1%) does not justify the excessive increase of curtailed power and reactive power usage.

3.3.6 Conclusions

The proposed optimized global control for PV inverters outperformed every other alternative in terms of apparent energy dedicated to control, averaging 0.48 kVAh per hour of good voltage values. In terms of voltage quality, only the VV local control (VV) performs slightly better (with voltage violations 0.07% of the time versus 0.14% for the OptGC), but with a less efficient use of the energy based on the indicators presented in the paper. In terms of effectiveness index, which measures the effective use of apparent power for voltage regulation, the optimized global strategy outperforms every local control and also the other metaheuristics analyzed (GA and PSO).

Additionally, the viability of using PV inverters as the only source of voltage regulation has been demonstrated, at least for the IEEE123 test feeder. The main limitations encountered by the strategy are related to the formulation of the optimization problem, which does not take into account variations in load conditions, which lead to some under voltages. A further analysis considering load variations and introducing forecasts will lead to better results. Another interesting

The suitability of using EDA as approximators in complex power system related optimization problems with a high number of variables was validated. The metaheuristic is able to provide quality solutions in a reasonable time, with convergence rates to feasible solutions and improvement rates once in the feasible region far superior to the other metaheuristics. The strategy of sampling the populations from distributions constructed with the information of the best individuals, proved to be superior to the relatively higher randomness of GA or the cooperative approach of PSO, at least for this specific problem.

84

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Chapter 4.   Conclusions and further research

In this section, the overall conclusions of the research, as well as the future lines of research that may be explored, are presented.

4.1   Conclusions

In this research, several interrelated topics, revolving around photovoltaic energy and photovoltaic systems, have been studied. The importance of modelling solar resources and plant performance, as well as forecasting solar energy, has been highlighted, especially thinking ahead in future scenarios where the share of renewable energy sources will be majority in the electric mixes. The adoption of new big data and artificial intelligence techniques in the modeling and forecasting pipelines of the processes has been proposed. The use of advanced control strategies, making use of the improvement of communications infrastructure and modern power electronics, is proposed, by laying out the use of fast optimization metaheuristics that can solve the optimization problems in acceptable times.

The deployment of renewable energies will continue to increase in the upcoming years, lining up with the global objectives set to combat climate change and the overall decrease in cost experienced (and expected to continue decreasing) for all renewable technologies. Modelling renewable energy plants and evaluating their performance is an important aspect in order to draw the picture of future electric and energy systems. Renewable resource assessments and potential estimations are a powerful tool to determine the real capacity that regions have to allocate alternative energy generation, and can be used to predict future energy scenarios. In the case of solar PV, the technology experiencing the highest falls in costs, the installation rates are expected to skyrocket over the next few years. Therefore, the modeling and simulation of their energy outputs and performances will prove to be valuable when selecting locations, configurations or ways of operation. These models and assessments can also be used to evaluate the performance of already deployed generators, pointing at possible inefficiencies in operation and maintenance of some PV plants.

The importance of energy policies over the deployment rates and overall performance of renewable energies, and specifically of PV solar, has been demonstrated.

85

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
        UNIVERSIDAD DE LA LAGUNA

        María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
        UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

4.1 Conclusions

In order to foster the installation of more PV (and other renewables) plants in the near future, a stable regulatory framework that clears up the investors uncertainties is required. With the grid parity in costs achieved by some renewable technologies nowadays, feed-in-tariffs may be not necessary. However other support strategies are recommended to promote the deployment of more clean energy in the electric grids.

Intra-hour solar forecasting will be a key technology in order to reach the objective of higher penetrations of solar PV energy in the electric mix, particularly in islands and weak electric grids. From a technical perspective, the variability of solar power plants leads to unwanted situations in terms of stability and quality of electric supply, and preventing these operational conditions is paramount for ensuring the reliability of electric systems. Island and weak electric grids are more vulnerable to the effects of the variability, and therefore, the relative importance of this technology in these regions is higher. Non-interconnected electric grids (including small micro-grids) will benefit the most from these technologies, by improving their operational conditions, leading to lower electricity costs, reductions in GHG emissions and extended lifetime of different devices (such as batteries).

Production of quality intra-hour forecasting seems to be tied to the use of sky-imagers, seems the stochastic nature of clouds is very difficult to capture with time-series analysis only. The improvement in computational capabilities of microprocessors, along with the overall decrease in costs of electronic components enable the production of low-cost and high-quality sky-imagers. The success of the adoption of this technology will depend on the overall costs of the forecasting systems, and the quality of the predictions made by it. If low-cost intra-hour forecasting systems are able to provide high-quality predictions, the added value to the electric system will be much higher, and this could lead to a massive implementation of the technology in those regions and electric systems where the value provided exceeds that from other alternatives (including operating the grid suboptimally).

The use of power electronics included in PV systems is another interesting alternative to cope with some of the impacts associated with the massive deployment of this technology in the electric grid. This is especially true (but not exclusive) in the case of PV systems connected to the distribution grid. Massive deployment in distribution grids poses a challenge over the traditional operation of electric grids, and the use of PV inverters to provide ancillary services is a very interesting strategy with low requirements. The main concern at this voltage level is the overvoltages due to active power injections on high resistive lines. Modern PV inverters can provide local voltage regulation by means of different strategies (active power curtailment and reactive power compensation). The improvement in communications infrastructure and computational capabilities is opening the possibility of using global control strategies, in which a higher-level service is fed with data from the grid and PV inverters. This high-level service can then analyze the data

86

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

received and optimize the operation of the grid, by sending commands to each device in the grid. This optimization should be done almost in real time, and for that purpose, fast algorithms or metaheuristics are required. From all the metaheuristics available in the literature, the use of distributions-based metaheuristics such as EDA, seems to be a good approach to these kinds of problems due to its fast convergence, decreased computational time and quality of solutions. By using optimized global control strategies, the hosting capacity of the distribution grids could be increased, and the supply of electricity to end-users will be more stable, safe and reliable.

In conclusion, PV energy will be a main actor in electric systems in the future. However, the current design of electric grids is not suited for a massive deployment of PV systems, and therefore, several changes in the way the grid is operated have to be made. Technologies and services such as solar forecasting or PV inverter-based regulation (among others) will play an important role in the transformation of electricity grids, helping to overcome the technical limitations encountered and enabling the integration of higher shares of solar energy into the grid.

4.2 Further research

Several future lines of research and extensions to the presented works could be identified for all the topics discussed in this thesis. Regarding the estimation of solar resource, PV specific yield, big data-based models and policy analysis, the research could be extended inspired by the following ideas:

- In the estimation of PV potential and PV specific yield, the use of advanced IA techniques, such as DL models, may be studied if larger datasets are available. DL models will probably outperform the simple model proposed in the first paper, and the inclusion of new features as inputs could improve even further the performance of the models.

- Extend the utility of the model to assess the performance of deployed PV plants by linking their performance to other environmental or regulatory variables, such as Saharan dust advections or imposed power curtailments by the system operator.

- Additionally, apply the same techniques to make similar studies for different renewable energies, such as wind power. In the dataset used for the first paper and provided by the utility, information about all the renewable generators in all the Canary Islands is presented. Thus characterization and modeling of wind power plants, and assessment of wind resource may be possible.

- An in-depth analysis of the effect of policies over the different installation rates on a regional scale could be interesting for the Canary Islands, in order to highlight

87

the importance of a robust and stable regulatory framework, and help in the decision making process for future policy makers.

In the research line of intra-hour solar forecasting using sky-imagers, there are multiple areas that could be extended in future works.

- Improvements in the sky-imagers can be proposed, by adding new sensors and sources of information that could be included in the posterior models to estimate the PV power or irradiance. There is also room for improvement in the information flow from the sky-imagers to the dedicated servers running the models.

- Hybrid models with DL approaches. Current research in our group is utilizing sequence of images from the sky-imagers as input in deep learning models, and by using specialized, architectures, operators and layers, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), the overall performance of the forecasts can be highly improved, at the cost of losing generality. However, if good models are found, these could be retrained with different dataset, conserving their structure and even their internal weights.

- Improvement of image and data preprocessing to extract more relevant features to input in the posterior models.

- Deployment of more sky-imagers in the surrounding areas of large PV plants to create a global and integrated forecasting system for all the island of Tenerife (extensible to other islands).

For the optimization of global control strategies for PV inverters and voltage regulation, the use of different metaheuristics besides EDA, as well as the implementation of different sampling strategies could be studied.

- Using different strategies in the EDA algorithm (e.g. efficiently building a model with multivariate dependencies). It is expected that if multivariate dependencies are considered when building the distribution functions to sample the populations, the overall performance of the optimization will increase, but it is likely that the computational cost will increase as well.

- Adding more PV systems on the studied grid. Although the EDA-based strategy still needs to be evaluated with a larger number of PV systems in the grid, based on our preliminary analysis we expect a similar performance if the same strategy without variable interdependencies (UMNA) is used.

- Extending the optimization problem to consider variations in load. The solutions input to the global control would then have an extra dimension (that is, they would be *splanes* instead of *splines*) to represent load variation. This is a very

88

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
            UNIVERSIDAD DE LA LAGUNA

            María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
            UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

interesting approach that will prepare the system from voltage variations coming from both sides (distributed generation and loads).

– Targeting runs for specific load and solar PV value based on probabilistic forecasts; and reducing the computational cost, e.g. by decreasing the number of generations being created at each optimization run. With the current strategy, there are multiple scenarios being simulated that are unlikely to occur. If reliable forecasts are available, then optimization runs could be done only for the most likely situations.

– Reformulate the optimization problem to find the best curves for the Volt-var control. Since in our research, VV control perform considerably good, maybe the global optimization of the VV curves (depending on the location of the PV system, distance to the substation…) could be an interesting approach.

– Test the optimized global control for PV inverters in other feeders, or even in real distribution networks, if the conditions for it are met.

– Analyze the interactions and behavior of different PV inverters when there is no active communications between them.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

4.2 Further research

90

References

[1] "Climate Change: Evidence & Causes," The Royal Society and The US National Academy of Sciences, 2020.

[2] T. A. Carleton and S. M. Hsiang, "Social and economic impacts of climate," *Science*, vol. 353, no. 6304, pp. aad9837–aad9837, Sep. 2016, doi: 10.1126/science.aad9837.

[3] W. J. Ripple, C. Wolf, T. M. Newsome, P. Barnard, and W. R. Moomaw, "World Scientists' Warning of a Climate Emergency," *BioScience*, p. biz088, Nov. 2019, doi: 10.1093/biosci/biz088.

[4] REN21, "Renewables 2020 Global Status Report," REN21, 2020. Accessed: Dec. 28, 2020. [Online]. Available: https://www.ren21.net/wp-content/uploads/2019/05/gsr_2020_full_report_en.pdf.

[5] "Climate change: US formally withdraws from Paris agreement," *BBC News*, Nov. 04, 2020.

[6] P. Friedlingstein *et al.*, "Global Carbon Budget 2019," *Earth Syst. Sci. Data*, vol. 11, no. 4, pp. 1783–1838, Dec. 2019, doi: https://doi.org/10.5194/essd-11-1783-2019.

[7] "Paris Agreement to the United Nations Framework Convention on Climate Change, T.I.A.S.," United Nations Framework Convention on Climate Change, Paris, Dec. 2015. Accessed: Jan. 12, 2021. [Online]. Available: https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

[8] IRENA, "Global Renewables Outlook: Energy Transformation 2050," IRENA.

[9] "The European Green Deal Investment Plan and JTM explained," *European Commission - European Commission*. https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_24 (accessed Jan. 22, 2021).

[10] IRENA, "Renewable power generation costs in 2019," p. 144, 2019.

[11] D. Ray, "Lazard's Levelized Cost of Energy Analysis—Version 14.0," p. 21, 2020.

[12] Carbon Tracker, "Powering down coal: Navigating the economic and financial risks in the last years of coal power," Carbon Tracker, 2018. Accessed: Feb. 26, 2021. [Online]. Available: https://carbontransfer.wpengine.com/wp-content/uploads/2018/12/CTI_Powering_Down_Coal_Report_Nov_2018_4-4.pdf.

[13] International Energy Agency, *Sustainable Recovery: World Energy Outlook Special Report*. OECD, 2020.

[14] BP, "Statistical Review of World Energy 2020 | 69th edition," BP, 2020. Accessed: Dec. 21, 2020. [Online]. Available: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2020-full-report.pdf.

[15]   "IEA Data & Statistics," *IEA*. https://www.iea.org/data-and-statistics (accessed Feb. 26, 2021).

[16]   P. Friedlingstein *et al.*, "Global Carbon Budget 2020," *Earth Syst. Sci. Data*, vol. 12, no. 4, pp. 3269–3340, Dec. 2020, doi: https://doi.org/10.5194/essd-12-3269-2020.

[17]   "Plan Nacional Integrado de Energía y Clima 2021-2030," Jan. 2020. Accessed: Jan. 20, 2021. [Online]. Available: https://www.miteco.gob.es/images/es/pnieccompleto_tcm30-508410.pdf.

[18]   "Plan Nacional de Adaptación al Cambio Climático 2021-2030." Accessed: Jan. 20, 2021. [Online]. Available: https://www.miteco.gob.es/images/es/pnacc-2021-2030_tcm30-512156.pdf.

[19]   "Canary Islands protest Spanish government's oil drilling approval," *the Guardian*, Mar. 27, 2012. http://www.theguardian.com/environment/2012/mar/27/spain-canaries-deepwater-drilling-protest (accessed Jan. 22, 2021).

[20]   "Anuario Energético de Canarias 2018," Gobierno de Canarias, Dec. 2019.

[21]   "Anuario del sector eléctrico de Canarias 2019," Gobierno de Canarias, Oct. 2020. Accessed: Jan. 25, 2021. [Online]. Available: https://www.gobiernodecanarias.org/cmsweb/export/sites/energia/doc/Publicaciones/AnuarioEnergeticoCanarias/AnuarioElectricoCanarias2019_pub.pdf.

[22]   T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, Jul. 2016, doi: 10.1016/j.ijforecast.2015.11.011.

[23]   A. Mills *et al.*, "Understanding Variability and Uncertainty of Photovoltaics for Integration with the Electric Power System," p. 18.

[24]   E. Ela, V. Diakov, E. Ibanez, and M. Heaney, "Impacts of Variability and Uncertainty in Solar Photovoltaic Generation at Multiple Timescales," *Renew. Energy*, p. 41, 2013.

[25]   M. Milligan *et al.*, "Integration of Variable Generation, Cost-Causation, and Integration Costs," *Electr. J.*, vol. 24, no. 9, pp. 51–63, Nov. 2011, doi: 10.1016/j.tej.2011.10.011.

[26]   T. Stetz, M. Rekinger, and I. Theologitis, "Transition from uni-directional to bi-directional distribution grids," IEA PVPS T14. Accessed: Jan. 29, 2021. [Online]. Available: https://iea-pvps.org/wp-content/uploads/2014/11/Transition_from_uni_directional_to_bi_directional_distribution_grids_REPORT_PVPS_T14_03_2014-1.pdf.

[27]   N. Amjady and F. Fallahi, "Determination of frequency stability border of power system to set the thresholds of under frequency load shedding relays," *Energy Convers. Manag.*, vol. 51, no. 10, pp. 1864–1872, Oct. 2010, doi: 10.1016/j.enconman.2010.02.016.

92

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

References

[28]   O. Gandhi, D. S. Kumar, C. D. Rodríguez-Gallegos, and D. Srinivasan, "Review of power system impacts at high PV penetration Part I: Factors limiting PV penetration," *Sol. Energy*, vol. 210, pp. 181–201, Nov. 2020, doi: 10.1016/j.solener.2020.06.097.

[29]   L. Sun *et al.*, "A GIS-based multi-criteria decision making method for the potential assessment and suitable sites selection of PV and CSP plants," *Resour. Conserv. Recycl.*, p. 105306, Nov. 2020, doi: 10.1016/j.resconrec.2020.105306.

[30]   Y. Sun, A. Hof, R. Wang, J. Liu, Y. Lin, and D. Yang, "GIS-based approach for potential analysis of solar PV generation at the regional scale: A case study of Fujian Province," *Energy Policy*, vol. 58, pp. 248–259, Jul. 2013, doi: 10.1016/j.enpol.2013.03.002.

[31]   Q. Yang *et al.*, "A GIS-based high spatial resolution assessment of large-scale PV generation potential in China," *Appl. Energy*, vol. 247, pp. 254–269, Aug. 2019, doi: 10.1016/j.apenergy.2019.04.005.

[32]   T. AlSkaif, S. Dev, L. Visser, M. Hossari, and W. van Sark, "A systematic analysis of meteorological variables for PV output power estimation," *Renew. Energy*, vol. 153, pp. 12–22, Jun. 2020, doi: 10.1016/j.renene.2020.01.150.

[33]   R. Singh, "Approximate rooftop solar PV potential of Indian cities for high-level renewable power scenario planning," *Sustain. Energy Technol. Assess.*, vol. 42, p. 100850, Dec. 2020, doi: 10.1016/j.seta.2020.100850.

[34]   R. O. Bawazir and N. S. Cetin, "Comprehensive overview of optimizing PV-DG allocation in power system and solar energy resource potential assessments," *Energy Rep.*, vol. 6, pp. 173–208, Nov. 2020, doi: 10.1016/j.egyr.2019.12.010.

[35]   K. Mason, M. J. Reno, L. Blakely, S. Vejdan, and S. Grijalva, "A deep neural network approach for behind-the-meter residential PV size, tilt and azimuth estimation," *Sol. Energy*, vol. 196, pp. 260–269, Jan. 2020, doi: 10.1016/j.solener.2019.11.100.

[36]   "JRC Photovoltaic Geographical Information System (PVGIS) - European Commission." https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html (accessed Feb. 01, 2021).

[37]   T. Huld, R. Müller, and A. Gambardella, "A new solar radiation database for estimating PV performance in Europe and Africa," *Sol. Energy*, vol. 86, no. 6, pp. 1803–1815, Jun. 2012, doi: 10.1016/j.solener.2012.03.006.

[38]   "NREL RE Atlas." https://maps.nrel.gov/re-atlas/ (accessed Feb. 01, 2021).

[39]   "Global Solar Atlas." https://globalsolaratlas.info/map?c=11.609193,8.261719,3 (accessed Feb. 01, 2021).

[40]   "NASA - POWER Data Access Viewer." https://power.larc.nasa.gov/data-access-viewer/ (accessed Feb. 01, 2021).

[41]   "NASA - Net Radiation," Jun. 30, 2020. https://earthobservatory.nasa.gov/global-maps/CERES_NETFLUX_M (accessed Feb. 01, 2021).

93

[42]   "Institute of Engineering Thermodynamics - SOLEMI - Solar Energy Mining."
https://www.dlr.de/tt/en/desktopdefault.aspx/tabid-2885/4422_read-6581/
(accessed Feb. 01, 2021).

[43]   Meteotest,   "Meteotest,"   *Meteotest*,   Jan.   31,   2021.
https://meteotest.ch/en/product/meteonorm (accessed Feb. 01, 2021).

[44]   "SODA-PRO." http://www.soda-pro.com/ (accessed Feb. 01, 2021).

[45]   "SOLARGIS - Solar Irradiance data." https://solargis.com/ (accessed Feb. 01, 2021).

[46]   "Solcast - Solar Forecasting & Solar Irradiance Data," *Solcast*. https://solcast.com/
(accessed Feb. 01, 2021).

[47]   A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and
analytics," *Int. J. Inf. Manag.*, vol. 35, no. 2, pp. 137–144, Apr. 2015, doi:
10.1016/j.ijinfomgt.2014.10.007.

[48]   G. de Freitas Viscondi and S. N. Alves-Souza, "A Systematic Literature Review on big
data for solar photovoltaic electricity generation forecasting," *Sustain. Energy
Technol. Assess.*, vol. 31, pp. 54–63, Feb. 2019, doi: 10.1016/j.seta.2018.11.008.

[49]   K. Moharm, "State of the art in big data applications in microgrid: A review," *Adv.
Eng. Inform.*, vol. 42, p. 100945, Oct. 2019, doi: 10.1016/j.aei.2019.100945.

[50]   M. Kezunovic, P. Pinson, Z. Obradovic, S. Grijalva, T. Hong, and R. Bessa, "Big data
analytics for future electricity grids," *Electr. Power Syst. Res.*, vol. 189, p. 106788,
Dec. 2020, doi: 10.1016/j.epsr.2020.106788.

[51]   Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities,"
Cities, vol. 89, pp. 80–91, Jun. 2019, doi: 10.1016/j.cities.2019.01.032.

[52]   C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A
review," *Renew. Energy*, vol. 105, pp. 569–582, May 2017, doi:
10.1016/j.renene.2016.12.095.

[53]   H. Jiang, K. Wang, Y. Wang, M. Gao, and Y. Zhang, "Energy big data: A survey," *IEEE
Access*, vol. 4, pp. 3844–3861, 2016, doi: 10.1109/ACCESS.2016.2580581.

[54]   A. M. Elshurafa, S. R. Albardi, S. Bigerna, and C. A. Bollino, "Estimating the learning
curve of solar PV balance–of–system for over 20 countries: Implications and policy
recommendations," *J. Clean. Prod.*, vol. 196, pp. 122–134, Sep. 2018, doi:
10.1016/j.jclepro.2018.06.016.

[55]   E. Hille, W. Althammer, and H. Diederich, "Environmental regulation and innovation
in renewable energy technologies: Does the policy instrument matter?," *Technol.
Forecast. Soc. Change*, vol. 153, p. 119921, Apr. 2020, doi:
10.1016/j.techfore.2020.119921.

[56]   A. Shivakumar, A. Dobbins, U. Fahl, and A. Singh, "Drivers of renewable energy
deployment in the EU: An analysis of past trends and projections," *Energy Strategy
Rev.*, vol. 26, p. 100402, Nov. 2019, doi: 10.1016/j.esr.2019.100402.

References

[57]   X. Zhao and D. Luo, "Driving force of rising renewable energy in China: Environment, regulation and employment," *Renew. Sustain. Energy Rev.*, vol. 68, pp. 48–56, Feb. 2017, doi: 10.1016/j.rser.2016.09.126.

[58]   Y. Zhang, J. Song, and S. Hamori, "Impact of subsidy policies on diffusion of photovoltaic power generation," *Energy Policy*, vol. 39, no. 4, pp. 1958–1964, Apr. 2011, doi: 10.1016/j.enpol.2011.01.021.

[59]   G. Shrimali and S. Jenner, "The impact of state policy on deployment and cost of solar photovoltaic technology in the U.S.: A sector-specific empirical analysis," *Renew. Energy*, vol. 60, pp. 679–690, Dec. 2013, doi: 10.1016/j.renene.2013.06.023.

[60]   M. Escoffier, E. Hache, V. Mignon, and A. Paris, "Determinants of solar photovoltaic deployment in the electricity mix: Do oil prices really matter?," *Energy Econ.*, p. 105024, Nov. 2020, doi: 10.1016/j.eneco.2020.105024.

[61]   A. J. Ryan, F. Donou-Adonsou, and L. N. Calkins, "Subsidizing the sun: The impact of state policies on electricity generated from solar photovoltaic," *Econ. Anal. Policy*, vol. 63, pp. 1–10, Sep. 2019, doi: 10.1016/j.eap.2019.04.012.

[62]   K. Gürtler, R. Postpischil, and R. Quitzow, "The dismantling of renewable energy policies: The cases of Spain and the Czech Republic," *Energy Policy*, vol. 133, p. 110881, Oct. 2019, doi: 10.1016/j.enpol.2019.110881.

[63]   *Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources (Text with EEA relevance.)*, vol. OJ L. 2018.

[64]   *Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU (Text with EEA relevance.)*, vol. OJ L. 2019.

[65]   *Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity (Text with EEA relevance.)*, vol. 158. 2019.

[66]   *Ley 54/1997, de 27 de noviembre, del Sector Eléctrico. Agencia Estatal Boletín Oficial del Estado (BOE).* .

[67]   *Real Decreto 436/2004, de 12 de marzo, por el que se establece la metodología para la actualización y sistematización del régimen jurídico y económico de la actividad de producción de energía eléctrica en régimen especial.* .

[68]   *Real Decreto 661/2007, de 25 de mayo, por el que se regula la actividad de producción de energía eléctrica en régimen especial.* .

[69]   *Real Decreto 1578/2008, de 26 de septiembre, de retribución de la actividad de producción de energía eléctrica mediante tecnología solar fotovoltaica para instalaciones posteriores a la fecha límite de mantenimiento de la retribución del Real Decreto 661/2007, de 25 de mayo, para dicha tecnología.* .

[70]   *Real Decreto-ley 1/2012, de 27 de enero, por el que se procede a la suspensión de los procedimientos de preasignación de retribución y a la supresión de los incentivos*

económicos para nuevas instalaciones de producción de energía eléctrica a partir de cogeneración, fuentes de energía renovables y residuos. .

[71] *Real Decreto-ley 9/2013, de 12 de julio, por el que se adoptan medidas urgentes para garantizar la estabilidad financiera del sistema eléctrico.* .

[72] *Real Decreto 900/2015, de 9 de octubre, por el que se regulan las condiciones administrativas, técnicas y económicas de las modalidades de suministro de energía eléctrica con autoconsumo y de producción con autoconsumo.* .

[73] *Real Decreto-ley 15/2018, de 5 de octubre, de medidas urgentes para la transición energética y la protección de los consumidores.* .

[74] *Real Decreto 244/2019, de 5 de abril, por el que se regulan las condiciones administrativas, técnicas y económicas del autoconsumo de energía eléctrica.* .

[75] "Series estadísticas nacionales | Red Eléctrica de España." https://www.ree.es/es/datos/publicaciones/series-estadisticas-nacionales (accessed Feb. 26, 2021).

[76] S. Pelland, J. Remund, J. Kleissl, T. Oozeki, and K. De Brabandere, "Photovoltaic and Solar Forecasting: State of the Art," IEA PVPS T14, 2013. Accessed: Jan. 19, 2021. [Online]. Available: https://iea-pvps.org/wp-content/uploads/2013/10/Photovoltaic_and_Solar_Forecasting_State_of_the_Art_R EPORT_PVPS_T14_01_2013.pdf.

[77] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016, doi: 10.1016/j.solener.2016.06.069.

[78] E. G. Evseev and A. I. Kudish, "The assessment of different models to predict the global solar radiation on a surface tilted to the south," *Sol. Energy*, vol. 83, no. 3, pp. 377–388, Mar. 2009, doi: 10.1016/j.solener.2008.08.010.

[79] J. E. Hay, Canada, and Atmospheric Environment Service, *Study of shortwave radiation on non-horizontal surfaces.* Downsview: Atmospheric Environment Service, 1979.

[80] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Sol. Energy*, vol. 44, no. 5, pp. 271–289, Jan. 1990, doi: 10.1016/0038-092X(90)90055-H.

[81] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," *Prog. Energy Combust. Sci.*, vol. 39, no. 6, pp. 535–576, Dec. 2013, doi: 10.1016/j.pecs.2013.06.002.

[82] R. W. Mueller *et al.*, "Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module," *Remote Sens. Environ.*, vol. 91, no. 2, pp. 160–174, May 2004, doi: 10.1016/j.rse.2004.02.009.

96

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

References

[83]   C. Rigollier, O. Bauer, and L. Wald, "On the clear sky model of the ESRA — European Solar Radiation Atlas — with respect to the heliosat method," *Sol. Energy*, vol. 68, no. 1, pp. 33–48, Jan. 2000, doi: 10.1016/S0038-092X(99)00055-9.

[84]   P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," *Sol. Energy*, vol. 73, no. 3, pp. 151–157, Sep. 2002, doi: 10.1016/S0038-092X(02)00045-2.

[85]   C. A. Gueymard, "REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset," *Sol. Energy*, vol. 82, no. 3, pp. 272–285, Mar. 2008, doi: 10.1016/j.solener.2007.04.008.

[86]   C. F. M. Coimbra and H. T. C. Pedro, "Chapter 15 - Stochastic-Learning Methods," in *Solar Energy Forecasting and Resource Assessment*, J. Kleissl, Ed. Boston: Academic Press, 2013, pp. 383–406.

[87]   R. Marquez and C. F. M. Coimbra, "Proposed Metric for Evaluation of Solar Forecasting Models," *J. Sol. Energy Eng.*, vol. 135, no. 1, pp. 011016-011016–9, Oct. 2012, doi: 10.1115/1.4007496.

[88]   P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction," *Nature*, vol. 525, no. 7567, Art. no. 7567, Sep. 2015, doi: 10.1038/nature14956.

[89]   S. Pelland, G. Galanis, and G. Kallos, "Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model," *Prog. Photovolt. Res. Appl.*, vol. 21, no. 3, pp. 284–296, 2013, doi: https://doi.org/10.1002/pip.1180.

[90]   M. Kudo, A. Takeuchi, Y. Nozaki, H. Endo, and J. Sumita, "Forecasting electric power generation in a photovoltaic power system for an energy network," *Electr. Eng. Jpn.*, vol. 167, no. 4, pp. 16–23, 2009, doi: https://doi.org/10.1002/eej.20755.

[91]   D. P. Larson, L. Nonnenmacher, and C. F. M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renew. Energy*, vol. 91, pp. 11–20, Jun. 2016, doi: 10.1016/j.renene.2016.01.039.

[92]   E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, and C. Kurz, "Regional PV power prediction for improved grid integration," *Prog. Photovolt. Res. Appl.*, vol. 19, no. 7, pp. 757–771, 2011, doi: https://doi.org/10.1002/pip.1033.

[93]   J. Huang, M. Korolkiewicz, M. Agrawal, and J. Boland, "Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model," *Sol. Energy*, vol. 87, pp. 136–149, Jan. 2013, doi: 10.1016/j.solener.2012.10.012.

[94]   Rui Huang, T. Huang, R. Gadh, and Na Li, "Solar generation prediction using the ARMA model in a laboratory-level micro-grid," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, Nov. 2012, pp. 528–533, doi: 10.1109/SmartGridComm.2012.6486039.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

References

[95]   B. Singh and D. Pozo, "A Guide to Solar Power Forecasting using ARMA Models," in *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, Bucharest, Romania, Sep. 2019, pp. 1–4, doi: 10.1109/ISGTEurope.2019.8905430.

[96]   J. Wu and C. K. Chan, "The Prediction of Monthly Average Solar Radiation with TDNN and ARIMA," in *2012 11th International Conference on Machine Learning and Applications*, Dec. 2012, vol. 2, pp. 469–474, doi: 10.1109/ICMLA.2012.225.

[97]   Z. Liu and Z. Zhang, "Solar forecasting by K-Nearest Neighbors method with weather classification and physical model," in *2016 North American Power Symposium (NAPS)*, Sep. 2016, pp. 1–6, doi: 10.1109/NAPS.2016.7747859.

[98]   J. Zeng and W. Qiao, "Short-term solar power prediction using a support vector machine," *Renew. Energy*, vol. 52, pp. 118–127, Apr. 2013, doi: 10.1016/j.renene.2012.10.009.

[99]   Jie Shi, Wei-Jen Lee, Yongqian Liu, Yongping Yang, and Peng Wang, "Forecasting power output of photovoltaic system based on weather classification and support vector machine," in *2011 IEEE Industry Applications Society Annual Meeting*, Oct. 2011, pp. 1–6, doi: 10.1109/IAS.2011.6074294.

[100] A. Khalyasmaa *et al.*, "Prediction of Solar Power Generation Based on Random Forest Regressor Model," in *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, Oct. 2019, pp. 0780–0785, doi: 10.1109/SIBIRCON48586.2019.8958063.

[101] L. Benali, G. Notton, A. Fouilloy, C. Voyant, and R. Dizene, "Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components," *Renew. Energy*, vol. 132, pp. 871–884, Mar. 2019, doi: 10.1016/j.renene.2018.08.044.

[102] A. K. Yadav and S. S. Chandel, "Solar radiation prediction using Artificial Neural Network techniques: A review," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 772–781, May 2014, doi: 10.1016/j.rser.2013.08.055.

[103] F. Barbieri, S. Rajakaruna, and A. Ghosh, "Very short-term photovoltaic power forecasting with cloud modeling: A review," *Renew. Sustain. Energy Rev.*, vol. 75, pp. 242–263, Aug. 2017, doi: 10.1016/j.rser.2016.10.068.

[104] D. W. van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1484–1512, Jan. 2018, doi: 10.1016/j.rser.2017.05.212.

[105] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renew. Sustain. Energy Rev.*, vol. 27, pp. 65–76, Nov. 2013, doi: 10.1016/j.rser.2013.06.042.

[106] J. Kleissl, *Solar Energy Forecasting and Resource Assessment*. Elsevier, 2013.

References

[107] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy Forecasting: A Review and Outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020, doi: 10.1109/OAJPE.2020.3029979.

[108] M. Zamo, O. Mestre, P. Arbogast, and O. Pannekoucke, "A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production," *Sol. Energy*, vol. 105, pp. 792–803, Jul. 2014, doi: 10.1016/j.solener.2013.12.006.

[109] U. K. Das *et al.*, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 912–928, Jan. 2018, doi: 10.1016/j.rser.2017.08.017.

[110] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization," *Renew. Sustain. Energy Rev.*, vol. 124, p. 109792, May 2020, doi: 10.1016/j.rser.2020.109792.

[111] J. E. E. Shields, R. W. Johnson, and T. L. Koehler, "Automated whole sky imaging systems for cloud field assessment," presented at the Fourth Symposium on Global Change Studies, Anaheim CA, Jan. 1993, [Online]. Available: http://www-mpl.ucsd.edu/people/jshields/publications/pdfs/13 Shields AMS 93.pdf.

[112] C. W. Chow *et al.*, "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed," *Sol. Energy*, vol. 85, no. 11, pp. 2881–2893, Nov. 2011, doi: 10.1016/j.solener.2011.08.025.

[113] R. Marquez, V. G. Gueorguiev, and C. F. M. Coimbra, "Forecasting of Global Horizontal Irradiance Using Sky Cover Indices," *J. Sol. Energy Eng.*, vol. 135, no. 1, pp. 011017-011017–5, Oct. 2012, doi: 10.1115/1.4007497.

[114] C.-L. L. Fu and H.-Y. Y. Cheng, "Predicting solar irradiance with all-sky image features via regression," *Sol. Energy*, vol. 97, pp. 537–550, Nov. 2013, doi: 10.1016/j.solener.2013.09.016.

[115] C. W. Chow, S. Belongie, and J. Kleissl, "Cloud motion and stability estimation for intra-hour solar forecasting," *Sol. Energy*, vol. 115, pp. 645–655, 2015, doi: 10.1016/j.solener.2015.03.030.

[116] H. Huang *et al.*, "Cloud motion estimation for short term solar irradiation prediction," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct. 2013, pp. 696–701, doi: 10.1109/SmartGridComm.2013.6688040.

[117] D. Bernecker, C. Riess, E. Angelopoulou, and J. Hornegger, "Continuous short-term irradiance forecasts using sky images," *Sol. Energy*, vol. 110, pp. 303–315, Dec. 2014, doi: 10.1016/j.solener.2014.09.005.

[118] Y. Chu, H. T. C. Pedro, L. Nonnenmacher, R. H. R. H. Inman, Z. Liao, and C. F. M. Coimbra, "A Smart image-based cloud detection system for intrahour solar irradiance forecasts," *J. Atmospheric Ocean. Technol.*, vol. 31, no. 9, pp. 1995–2007, Sep. 2014, doi: 10.1175/JTECH-D-13-00209.1.

[119] S. R. West, D. Rowe, S. Sayeef, and A. Berry, "Short-term irradiance forecasting using skycams: Motivation and development," *Sol. Energy*, vol. 110, pp. 188–207, Dec. 2014, doi: 10.1016/j.solener.2014.08.038.

[120] H. Yang *et al.*, "Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego," *Sol. Energy*, vol. 103, pp. 502–524, May 2014, doi: 10.1016/j.solener.2014.02.044.

[121] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "Estimation of solar irradiance using ground-based whole sky imagers," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2016, pp. 7236–7239, doi: 10.1109/IGARSS.2016.7730887.

[122] Y. Ai, Y. Peng, and W. Wei, "A model of very short-term solar irradiance forecasting based on low-cost sky images," Hangzhou, China, 2017, p. 020022, doi: 10.1063/1.4982387.

[123] R. Marquez and C. F. M. C. F. M. C. F. M. Coimbra, "Intra-hour DNI forecasting based on cloud tracking image analysis," *Sol. Energy*, vol. 91, pp. 327–336, 2013, doi: 10.1016/j.solener.2012.09.018.

[124] P. Blanc *et al.*, "Short-term forecasting of high resolution local DNI maps with multiple fish-eye cameras in stereoscopic mode," Abu Dhabi, United Arab Emirates, 2017, p. 140004, doi: 10.1063/1.4984512.

[125] Y. Chu, M. Li, and C. F. M. C. F. M. Coimbra, "Sun-tracking imaging system for intra-hour DNI forecasts," *Renew. Energy*, vol. 96, pp. 792–799, Oct. 2016, doi: 10.1016/j.renene.2016.05.041.

[126] Y. Chu, H. T. C. Pedro, and C. F. M. Coimbra, "Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning," *Sol. Energy*, vol. 98, no. PC, pp. 592–603, Dec. 2013, doi: 10.1016/j.solener.2013.10.020.

[127] Y. Chu, M. Li, H. T. C. Pedro, and C. F. M. Coimbra, "Real-time prediction intervals for intra-hour DNI forecasts," *Renew. Energy*, vol. 83, pp. 234–244, 2015, doi: 10.1016/j.renene.2015.04.022.

[128] R. Chauvin, J. Nou, S. Thil, and S. Grieu, "Intra-day DNI forecasting under clear sky conditions using ANFIS," in *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 2014, vol. 19, pp. 10361–10366, [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84929773933&partnerID=40&md5=3b67cb832b32ed5ea41629327caff899.

[129] R. Chauvin, J. Nou, S. Thil, and S. Grieu, "A new approach for assessing the clear-sky direct normal irradiance in real time," 2016, doi: 10.1109/EEEIC.2016.7555435.

[130] L. Liu, X. Sun, F. Chen, S. Zhao, and T. Gao, "Cloud Classification Based on Structure Features of Infrared Images," *J. Atmospheric Ocean. Technol.*, vol. 28, no. 3, pp. 410–417, Apr. 2010, doi: 10.1175/2010JTECHA1385.1.

[131] M. S. Ghonima, B. Urquhart, C. W. Chow, J. E. Shields, A. Cazorla, and J. Kleissl, "A method for cloud detection and opacity classification based on ground based sky

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

References

imagery," *Atmospheric Meas. Tech.*, vol. 5, no. 11, pp. 2881–2892, 2012, doi: 10.5194/amt-5-2881-2012.

[132] A. Kazantzidis, P. Tzoumanikas, A. F. Bais, S. Fotopoulos, and G. Economou, "Cloud detection and classification with the use of whole-sky ground-based images," *Atmospheric Res.*, vol. 113, pp. 80–88, Sep. 2012, doi: 10.1016/j.atmosres.2012.05.005.

[133] L. Liu, X. Sun, T. Gao, and S. Zhao, "Comparison of Cloud Properties from Ground-Based Infrared Cloud Measurement and Visual Observations," *J. Atmospheric Ocean. Technol.*, vol. 30, no. 6, pp. 1171–1179, Jun. 2013, doi: 10.1175/JTECH-D-12-00157.1.

[134] R. Tapakis and A. G. Charalambides, "Equipment and methodologies for cloud detection and classification: A review," *Sol. Energy*, vol. 95, pp. 392–430, Sep. 2013, doi: 10.1016/j.solener.2012.11.015.

[135] J. Alonso, F. J. J. Batlles, G. López, and A. Ternero, "Sky camera imagery processing based on a sky classification using radiometric data," *Energy*, vol. 68, pp. 599–608, Apr. 2014, doi: 10.1016/j.energy.2014.02.035.

[136] J. Alonso and F. J. J. Batlles, "Short and medium-term cloudiness forecasting using remote sensing techniques and sky camera imagery," *Energy*, vol. 73, pp. 890–897, Aug. 2014, doi: 10.1016/j.energy.2014.06.101.

[137] R. Chauvin, J. Nou, S. Thil, and S. Grieu, "Cloud motion estimation using a sky imager," in *AIP Conference Proceedings*, 2016, vol. 1734, doi: 10.1063/1.4949235.

[138] A. Taravat, F. Del Frate, C. Cornaro, and S. Vergari, "Neural Networks and Support Vector Machine Algorithms for Automatic Cloud Classification of Whole-Sky Ground-Based Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 666–670, Mar. 2015, doi: 10.1109/LGRS.2014.2356616.

[139] P. Kuhn *et al.*, "Shadow camera system for the generation of solar irradiance maps," *Sol. Energy*, vol. 157, pp. 157–170, Nov. 2017, doi: 10.1016/j.solener.2017.05.074.

[140] Y. Chu, H. T. C. Pedro, M. Li, and C. F. M. Coimbra, "Real-time forecasting of solar irradiance ramps with smart image processing," *Sol. Energy*, vol. 114, pp. 91–104, Apr. 2015, doi: 10.1016/j.solener.2015.01.024.

[141] P. Kuhn *et al.*, "Validation of an all-sky imager-based nowcasting system for industrial PV plants," *Prog. Photovolt. Res. Appl.*, vol. 26, no. June, pp. 1–14, Nov. 2017, doi: 10.1002/pip.2968.

[142] V. Venugopal, Y. Sun, and A. R. Brandt, "Short-term solar PV forecasting using computer vision: The search for optimal CNN architectures for incorporating sky images and PV generation history," *J. Renew. Sustain. Energy*, vol. 11, no. 6, p. 066102, Nov. 2019, doi: 10.1063/1.5122796.

[143] Z. Zhen *et al.*, "Pattern Classification and PSO Optimal Weights Based Sky Images Cloud Motion Speed Calculation Method for Solar PV Power Forecasting," *IEEE*

Trans. Ind. Appl., vol. 55, no. 4, pp. 3331–3342, Jul. 2019, doi: 10.1109/TIA.2019.2904927.

[144] T. Schmidt *et al.*, "Short-term solar forecasting based on sky images to enable higher PV generation in remote electricity networks," *Renew. Energy Environ. Sustain.*, vol. 2, p. 23, 2017, doi: 10.1051/rees/2017028.

[145] Z. Zhen *et al.*, "SVM based cloud classification model using total sky images for PV power forecasting," *2015 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. ISGT*, no. 51277075, pp. 1–5, 2015, doi: 10.1109/ISGT.2015.7131784.

[146] Z. Zhen *et al.*, "Deep Learning Based Surface Irradiance Mapping Model for Solar PV Power Forecasting Using Sky Image," *IEEE Trans. Ind. Appl.*, pp. 1–1, 2020, doi: 10.1109/TIA.2020.2984617.

[147] Y. Yu, J. Cao, and J. Zhu, "An LSTM Short-Term Solar Irradiance Forecasting Under Complicated Weather Conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019, doi: 10.1109/ACCESS.2019.2946057.

[148] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms," *Appl. Energy*, vol. 253, p. 113541, Nov. 2019, doi: 10.1016/j.apenergy.2019.113541.

[149] M. Husein and I.-Y. Chung, "Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach," *Energies*, vol. 12, no. 10, p. 1856, May 2019, doi: 10.3390/en12101856.

[150] R. Marquez and C. F. M. Coimbra, "Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database," *Sol. Energy*, vol. 85, no. 5, pp. 746–756, 2011, doi: 10.1016/j.solener.2011.01.007.

[151] F. V. Gutierrez-Corea, M. A. Manso-Callejo, M. P. Moreno-Regidor, and M. T. Manrique-Sancho, "Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations," *Sol. Energy*, 2016, doi: 10.1016/j.solener.2016.04.020.

[152] N. Dong, J.-F. Chang, A.-G. Wu, and Z.-K. Gao, "A novel convolutional neural network framework based solar irradiance prediction method," *Int. J. Electr. Power Energy Syst.*, vol. 114, p. 105411, Jan. 2020, doi: 10.1016/j.ijepes.2019.105411.

[153] J. Cao and X. Lin, "Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks," *Energy Convers. Manag.*, vol. 49, no. 6, pp. 1396–1406, 2008, doi: 10.1016/j.enconman.2007.12.030.

[154] C. Chen, S. Duan, T. Cai, and B. Liu, "Online 24-h solar power forecasting based on weather type classification using artificial neural network," *Sol. Energy*, vol. 85, no. 11, pp. 2856–2870, 2011, doi: 10.1016/j.solener.2011.08.027.

[155] H. Lee and B.-T. Lee, "Confidence-aware deep learning forecasting system for daily solar irradiance," *IET Renew. Power Gener.*, vol. 13, no. 10, pp. 1681–1689, 2019, doi: 10.1049/iet-rpg.2018.5354.

102

References

[156] L. Liu *et al.*, "Prediction of short-term PV power output and uncertainty analysis," *Appl. Energy*, vol. 228, pp. 700–711, Oct. 2018, doi: 10.1016/j.apenergy.2018.06.112.

[157] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Appl. Energy*, vol. 251, p. 113315, Oct. 2019, doi: 10.1016/j.apenergy.2019.113315.

[158] F. Wang *et al.*, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," *Energy Convers. Manag.*, vol. 181, pp. 443–462, Feb. 2019, doi: 10.1016/j.enconman.2018.11.074.

[159] H. Zang, L. Cheng, T. Ding, K. W. Cheung, Z. Wei, and G. Sun, "Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning," *Int. J. Electr. Power Energy Syst.*, vol. 118, p. 105790, Jun. 2020, doi: 10.1016/j.ijepes.2019.105790.

[160] C.-J. Huang and P.-H. Kuo, "Multiple-Input Deep Convolutional Neural Network Model for Short-Term Photovoltaic Power Forecasting," *IEEE Access*, vol. 7, pp. 74822–74834, 2019, doi: 10.1109/ACCESS.2019.2921238.

[161] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Jul. 2019, doi: 10.1007/s00521-017-3225-z.

[162] K. Wang, X. Qi, and H. Liu, "Photovoltaic power forecasting based LSTM-Convolutional Network," *Energy*, vol. 189, p. 116225, Dec. 2019, doi: 10.1016/j.energy.2019.116225.

[163] H. Wang *et al.*, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manag.*, vol. 153, pp. 409–422, Dec. 2017, doi: 10.1016/j.enconman.2017.10.008.

[164] Z. Li, S. Rahman, R. Vega, and B. Dong, "A Hierarchical Approach Using Machine Learning Methods in Solar Photovoltaic Energy Production Forecasting," *Energies*, vol. 9, no. 1, p. 55, 2016, doi: 10.3390/en9010055.

[165] J. Zhang, R. Verschae, S. Nobuhara, and J.-F. Lalonde, "Deep photovoltaic nowcasting," *Sol. Energy*, vol. 176, pp. 267–276, Dec. 2018, doi: 10.1016/j.solener.2018.10.024.

[166] Y. Sun, G. Szűcs, and A. R. Brandt, "Solar PV output prediction from video streams using convolutional neural networks," *Energy Environ. Sci.*, vol. 11, no. 7, pp. 1811–1818, 2018, doi: 10.1039/C7EE03420B.

[167] Y. Sun, V. Venugopal, and A. R. Brandt, "Short-term solar power forecast with deep learning: Exploring optimal input and output configuration," *Sol. Energy*, vol. 188, pp. 730–741, Aug. 2019, doi: 10.1016/j.solener.2019.06.041.

[168] Y. Sun, V. Venugopal, and A. R. Brandt, "Convolutional Neural Network for Short-term Solar Panel Output Prediction," p. 5.

103

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
             UNIVERSIDAD DE LA LAGUNA

             María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
             UNIVERSIDAD DE LA LAGUNA

[169] M. N. Akhter, S. Mekhilef, H. Mokhlis, and N. Mohamed Shah, "Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques," *IET Renew. Power Gener.*, vol. 13, no. 7, pp. 1009–1023, 2019, doi: 10.1049/iet-rpg.2018.5649.

[170] S. Sobri, S. Koohi-Kamali, and N. Abd. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Convers. Manag.*, vol. 156, pp. 459–497, Jan. 2018, doi: 10.1016/j.enconman.2017.11.019.

[171] A. Molina-Garcia, R. A. Mastromauro, T. Garcia-Sanchez, S. Pugliese, M. Liserre, and S. Stasi, "Reactive Power Flow Control for PV Inverters Voltage Support in LV Distribution Networks," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 447–456, Jan. 2017, doi: 10.1109/TSG.2016.2625314.

[172] Red Eléctrica de España, "P.O. SEIE – 1 Funcionamiento de los sistemas eléctricos insulares y extrapeninsulares." .

[173] R. Passey, T. Spooner, I. MacGill, M. Watt, and K. Syngellakis, "The potential impacts of grid-connected distributed generation and how to address them: A review of technical and non-technical factors," *Energy Policy*, vol. 39, no. 10, pp. 6280–6290, Oct. 2011, doi: 10.1016/j.enpol.2011.07.027.

[174] "Overcoming PV grid issues in the urban areas," IEA PVPS-T10, Oct. 2009. Accessed: Feb. 03, 2021. [Online].

[175] M. A. Eltawil and Z. Zhao, "Grid-connected photovoltaic power systems: Technical and potential problems—A review," *Renew. Sustain. Energy Rev.*, vol. 14, no. 1, pp. 112–129, Jan. 2010, doi: 10.1016/j.rser.2009.07.015.

[176] P. Unruh, M. Nuschke, P. Strauß, and F. Welck, "Overview on Grid-Forming Inverter Control Methods," *Energies*, vol. 13, no. 10, Art. no. 10, Jan. 2020, doi: 10.3390/en13102589.

[177] L. Hirth, F. Ueckerdt, and O. Edenhofer, "Integration costs revisited – An economic framework for wind and solar variability," *Renew. Energy*, vol. 74, pp. 925–939, Feb. 2015, doi: 10.1016/j.renene.2014.08.065.

[178] R. Yan and T. K. Saha, "Investigation of Voltage Stability for Residential Customers Due to High Photovoltaic Penetrations," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 651–662, May 2012, doi: 10.1109/TPWRS.2011.2180741.

[179] M. H. J. Bollen, "Definitions of Voltage Unbalance," *IEEE Power Eng. Rev.*, vol. 22, no. 11, pp. 49–50, Nov. 2002, doi: 10.1109/MPER.2002.4311797.

[180] M. M. Haque and P. Wolfs, "A review of high PV penetrations in LV distribution networks: Present status, impacts and mitigation measures," *Renew. Sustain. Energy Rev.*, vol. 62, pp. 1195–1208, Sep. 2016, doi: 10.1016/j.rser.2016.04.025.

[181] L. Degroote, B. Renders, B. Meersman, and L. Vandevelde, "Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks," in *2009 IEEE Bucharest PowerTech*, Jun. 2009, pp. 1–8, doi: 10.1109/PTC.2009.5281998.

104

References

[182] M. Juamperez, G. Yang, and S. B. Kjær, "Voltage regulation in LV grids by coordinated volt-var control strategies," *J. Mod. Power Syst. Clean Energy*, vol. 2, no. 4, pp. 319–328, Dec. 2014, doi: 10.1007/s40565-014-0072-0.

[183] S. Weckx, C. Gonzalez, and J. Driesen, "Combined Central and Local Active and Reactive Power Control of PV Inverters," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 776–784, Jul. 2014, doi: 10.1109/TSTE.2014.2300934.

[184] M. McGranaghan, T. Ortmeyer, D. Crudele, T. Key, J. Smith, and P. Barker, "Advanced Grid Planning and Operations," p. 123.

[185] A. Mills *et al.*, "Dark Shadows," *IEEE Power Energy Mag.*, vol. 9, no. 3, pp. 33–41, May 2011, doi: 10.1109/MPE.2011.940575.

[186] J. Seuss, M. J. Reno, R. J. Broderick, and S. Grijalva, "Analysis of PV Advanced Inverter Functions and Setpoints under Time Series Simulation.," SAND2016-4856, 1259558, May 2016. doi: 10.2172/1259558.

[187] S. M. Alizadeh, C. Ozansoy, and T. Alpcan, "The impact of X/R ratio on voltage stability in a distribution network penetrated by wind farms," in *2016 Australasian Universities Power Engineering Conference (AUPEC)*, Brisbane, Australia, Sep. 2016, pp. 1–6, doi: 10.1109/AUPEC.2016.7749289.

[188] A. O'Connell and A. Keane, "Volt–var curves for photovoltaic inverters in distribution systems," *IET Gener. Transm. Distrib.*, vol. 11, no. 3, pp. 730–739, Feb. 2017, doi: 10.1049/iet-gtd.2016.0409.

[189] P. Martí, M. Velasco, J. M. Fuertes, A. Camacho, J. Miret, and M. Castilla, "Distributed reactive power control methods to avoid voltage rise in grid-connected photovoltaic power generation systems," in *2013 IEEE International Symposium on Industrial Electronics*, May 2013, pp. 1–6, doi: 10.1109/ISIE.2013.6563803.

[190] R. Ma, H. Chen, Y. Huang, and W. Meng, "Smart Grid Communication: Its Challenges and Opportunities," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 36–46, Mar. 2013, doi: 10.1109/TSG.2012.2225851.

[191] Y. P. Agalgaonkar, B. C. Pal, and R. A. Jabr, "Distribution Voltage Control Considering the Impact of PV Generation on Tap Changers and Autonomous Regulators," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 182–192, Jan. 2014, doi: 10.1109/TPWRS.2013.2279721.

[192] A. Bonfiglio, M. Brignone, F. Delfino, and R. Procopio, "Optimal Control and Operation of Grid-Connected Photovoltaic Production Units for Voltage Support in Medium-Voltage Networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 1, pp. 254–263, Jan. 2014, doi: 10.1109/TSTE.2013.2280811.

[193] X. Su, M. A. S. Masoum, and P. J. Wolfs, "Optimal PV Inverter Reactive Power Control and Real Power Curtailment to Improve Performance of Unbalanced Four-Wire LV Distribution Networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 967–977, Jul. 2014, doi: 10.1109/TSTE.2014.2313862.

[194] X. Su, M. A. S. Masoum, and P. Wolfs, "PSO based multi-objective optimization of unbalanced lv distribution network by PV inverter control," in *2014 China International Conference on Electricity Distribution (CICED)*, Sep. 2014, pp. 1744–1748, doi: 10.1109/CICED.2014.6992001.

[195] E. Demirok, P. C. González, K. H. B. Frederiksen, D. Sera, P. Rodriguez, and R. Teodorescu, "Local Reactive Power Control Methods for Overvoltage Prevention of Distributed Solar Inverters in Low-Voltage Grids," *IEEE J. Photovolt.*, vol. 1, no. 2, pp. 174–182, Oct. 2011, doi: 10.1109/JPHOTOV.2011.2174821.

[196] I. Boussaïd, J. Lepagnot, and P. Siarry, "A survey on optimization metaheuristics," *Inf. Sci.*, vol. 237, pp. 82–117, Jul. 2013, doi: 10.1016/j.ins.2013.02.041.

[197] K. Sörensen, "Metaheuristics—the metaphor exposed," *Int. Trans. Oper. Res.*, vol. 22, no. 1, pp. 3–18, 2015, doi: https://doi.org/10.1111/itor.12001.

[198] K. Sörensen and F. W. Glover, "Metaheuristics," in *Encyclopedia of Operations Research and Management Science*, S. I. Gass and M. C. Fu, Eds. Boston, MA: Springer US, 2013, pp. 960–970.

[199] J. A. Parejo, A. Ruiz-Cortés, S. Lozano, and P. Fernandez, "Metaheuristic optimization frameworks: a survey and benchmarking," *Soft Comput.*, vol. 16, no. 3, pp. 527–561, Mar. 2012, doi: 10.1007/s00500-011-0754-8.

[200] M. A. S. Masoum, S. M. M. Badejani, and M. Kalantar, "Optimal placement of hybrid PV-wind systems using genetic algorithm," in *2010 Innovative Smart Grid Technologies (ISGT)*, Jan. 2010, pp. 1–5, doi: 10.1109/ISGT.2010.5434746.

[201] L. Liu, Q. Sun, Y. Wang, Y. Liu, and R. Wennersten, "Research on Short-term Optimization for Integrated Hydro-PV Power System Based on Genetic Algorithm," *Energy Procedia*, vol. 152, pp. 1097–1102, Oct. 2018, doi: 10.1016/j.egypro.2018.09.132.

[202] L. jing Hu, K. Liu, Y. Fu, and P. Li, "Capacity Optimization of Wind /PV/Storage Power System Based on Simulated Annealing-Particle Swarm Optimization," in *2018 37th Chinese Control Conference (CCC)*, Jul. 2018, pp. 2222–2227, doi: 10.23919/ChiCC.2018.8482706.

[203] D. Gómez-Lorente, I. Triguero, C. Gil, and A. Espín Estrella, "Evolutionary algorithms for the design of grid-connected PV-systems," *Expert Syst. Appl.*, vol. 39, no. 9, pp. 8086–8094, Jul. 2012, doi: 10.1016/j.eswa.2012.01.159.

[204] Dr. P. N. Hrisheekesha and Dr. J. Sharma, "Evolutionary Algorithm Based Optimal Control in Distribution System with Dispersed Generation," *Int. J. Comput. Appl.*, vol. 1, no. 14, pp. 35–41, Feb. 2010, doi: 10.5120/305-471.

[205] T. Niknam, "A new approach based on ant colony optimization for daily Volt/Var control in distribution networks considering distributed generators," *Energy Convers. Manag.*, vol. 49, no. 12, pp. 3417–3424, Dec. 2008, doi: 10.1016/j.enconman.2008.08.015.

106

References

[206] H. Yoshida, K. Kawata, Y. Fukuyama, S. Takayama, and Y. Nakanishi, "A particle swarm optimization for reactive power and voltage control considering voltage security assessment," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1232–1239, Nov. 2000, doi: 10.1109/59.898095.

[207] Y. Hong, F. Lin, Y. Lin, and F. Hsu, "Chaotic PSO-Based VAR Control Considering Renewables Using Fast Probabilistic Power Flow," *IEEE Trans. Power Deliv.*, vol. 29, no. 4, pp. 1666–1674, Aug. 2014, doi: 10.1109/TPWRD.2013.2285923.

[208] H. Yang and J. Liao, "MF-APSO-Based Multiobjective Optimization for PV System Reactive Power Regulation," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1346–1355, Oct. 2015, doi: 10.1109/TSTE.2015.2433957.

[209] M. M. Othman, M. H. Ahmed, and M. M. A. Salama, "A Coordinated Real-Time Voltage Control Approach for Increasing the Penetration of Distributed Generation," *IEEE Syst. J.*, vol. 14, no. 1, pp. 699–707, Mar. 2020, doi: 10.1109/JSYST.2019.2904532.

[210] A. A. A. E. Ela, M. A. Abido, and S. R. Spea, "Differential evolution algorithm for optimal reactive power dispatch," *Electr. Power Syst. Res.*, vol. 81, no. 2, pp. 458–464, Feb. 2011, doi: 10.1016/j.epsr.2010.10.005.

[211] M. Varadarajan and K. S. Swarup, "Differential evolution approach for optimal reactive power dispatch," *Appl. Soft Comput.*, vol. 8, no. 4, pp. 1549–1561, Sep. 2008, doi: 10.1016/j.asoc.2007.12.002.

[212] W. Su and M.-Y. Chow, "Performance Evaluation of an EDA-Based Large-Scale Plug-In Hybrid Electric Vehicle Charging Algorithm," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 308–315, Mar. 2012, doi: 10.1109/TSG.2011.2151888.

[213] D. G. Rojas, J. L. Lezama, and W. Villa, "Metaheuristic Techniques Applied to the Optimal Reactive Power Dispatch: a Review," *IEEE Lat. Am. Trans.*, vol. 14, no. 5, pp. 2253–2263, May 2016, doi: 10.1109/TLA.2016.7530421.

[214] Y. A. Rahman, S. Manjang, Yusran, and A. A. Ilham, "An Empirical Metaheuristic Assessment for Solving of Multi-type Distributed Generation Allocation Problem," in *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Nov. 2018, pp. 699–702, doi: 10.1109/ISRITI.2018.8864438.

[215] *Orden IET/1459/2014, de 1 de agosto, por la que se aprueban los parámetros retributivos y se establece el mecanismo de asignación del régimen retributivo específico para nuevas instalaciones eólicas y fotovoltaicas en los sistemas eléctricos de los territorios no peninsulares.* .

107

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

References

108

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annexes

Annexes

109

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annexes

110

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annexes

Annex I: A simple big data methodology and analysis of the specific yield of all PV power plants in a power system over a long time period

111

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annexes

112

A simple big data methodology and analysis of the specific yield of all PV power plants in a power system over a long time period

R. Guerrero-Lemus[a,*], D. Cañadillas-Ramallo[a], T. Reindl[b], J.M. Valle-Feijóo[c]

[a] *Departamento de Física, Universidad de La Laguna, Avenida Astrofísico Francisco Sánchez S/N 38206 S/C de, Tenerife, Spain*
[b] *Solar Energy Research Institute of Singapore (SERIS), National University of Singapore (NUS), Singapore 117574, Singapore*
[c] *ENDESA, Carlos Jr. Hamilton s/n, 38001 S/C de, Tenerife, Spain*

A B S T R A C T

In this work, we have reviewed recent literature concerning the performance of PV plants in different power systems and detected that, usually, the raw data used is not complete in terms of the number of PV plants and also main parameters of poor quality need to be removed. Then, for the first time, a study of the specific yield of all PV plants in a power system with single-PV-plant resolution is presented. Thus, we have analyzed the official 237,588 monthly energy values obtained from 2005 to 2017 for the 1523 PV plants in the Canary Islands. This dataset is obtained from PV plants ranging from $0.53\,kW_p$ up to $9\,MW_p$, and it has been supplied by the distribution system operator and main utility (ENDESA). Then, this dataset is compared to 153,120 irradiance and temperature data from a PVGIS database. Results show that the Spanish regulation has a direct effect not only on the development of the PV capacity in the Canary Islands, but also on the specific yields. Moreover, only combining meteorological data (irradiance, temperature and wind speed) from satellites, starting year of operation, and nameplate capacity we have developed a very simple theoretical model to predict the specific yield of a PV plant at any location in the Canary Islands, avoiding the requirement of any data from the owners of the PV plants. The simulation values obtained have been validated with the real specific yields for PV plants assumed to be well managed (multi-MW power plants placed in best locations) showing errors below a 3%. This theoretical model has also been used for detecting suboptimal PV plant designs and anomalous specific yield of PV plants above the clear sky limit. Recommendations to avoid anomalous specific yields in future are included.

1. Introduction

Photovoltaics (PV) is growing sharply around the world [1], and it is expected to play a dominant role in regions where the cost of producing electricity from this renewable source is far below the cost from conventional power sources [2]. However, there is a lack of research articles studying the performance of all the PV plants in a power system with a single-PV-plant resolution. This is attributed to legal concerns for providing such information from utilities, system operators and energy commissions to researchers [3].

Works detected in the literature only describe and analyze representative subsets of the PV plants operating inside the power system [4–6], or subsets that are obtained from web pages where owners of small capacity PV systems installed on roofs altruistically upload energy production data that is only partially valuable [7–9]. Experience has highlighted the often poor quality of such self-reporting for rooftop PV systems [3]. Monitoring data from PV systems installed in meteorological weather stations is another alternative, but they are also subsets and include a disproportionate share of large PV systems, because the power produced in small systems is usually not monitored [4]. Other works use datasets obtained from industry associations and energy agencies, but only for nameplate capacities and not energy production [10]. Also, there are works in the literature where proprietary dataset are obtained from public commissions regulating utilities, but only for studying contracts under specific incentive programs [11]. Recognized world electric power plant databases offer data characterizing power plants, but only about plant operators, geographic location, capacity, age, technology, fuels manufacturers, etc.[12], but not about energy production.

To obtain weather conditions for estimating the potential energy of a PV power plant, geographic information services (GIS) from satellite-based datasets are usually employed. However, these datasets can also provide accuracy uncertainties, e.g., NASA-SSE reported global mean square errors of 10.25% for the global horizontal radiation at a defined

---

113

time resolution [13]. The alternative option is the use of ground measurement data stations, as they can achieve a higher accuracy if the sensors are well calibrated. For islands as the Canaries, a rich variation of microclimatic conditions exists due to an uneven relief of most islands. Then, as there are only a few weather stations offering the required dataset, the use of ground measurement weather stations is not an option for the study of most PV systems connected to the power grid. Also, networks of weather stations have a high cost not only for installation, but for operation and maintenance.

We believe that it is necessary for researchers to have access to the performance of all PV plants in a power system if advances in the integration of this renewable technology are required from governments and utilities. However, the specific yield of each PV plant must be protected. In this work we expose and analyze the main results obtained from the performance of all the PV systems injecting electricity to the 7 power grids that compose the Canary Islands power system (the Fuerteventura and Lanzarote power grids are partially connected). Detailed results have been discarded from publishing for assuring that the identification of individual PV plants is not possible. To our knowledge, this is the first time a research article is showing and analyzing such information. We expect that it shows a convenient methodology and precedent for the scientific community in an effort to have access to similar information from other power systems.

The paper is organized as follows: In Section 2, the methodology used in this work is described, including the characteristics of the energy production data, data filtering process, geographic location of each PV plant and solar resource. In Section 3, the input data for our model, and our clear sky model is described. In Section 4, the results are exposed about the evolution of the PV capacity in relation to the regulation approved by the Spanish government in the period under analysis, the geographical distribution of the PV plants influenced by different parameters, upper sky limits, and specific yields. In section 5, the main conclusions are included.

2. Materials and methods

The 237,588 monthly energy production data have been obtained from the AC billing meters provided by the distribution system operator and main utility (ENDESA). This dataset has a monthly frequency because the PV plants are remunerated on a monthly basis. The uncertainty of AC billing meters is rather small (0.2–0.5%) and well documented [14]. The data combines production from large PV plants placed on ground at the best locations and orientation in the Canary Islands, and from small PV plants placed on roofs and often poorly placed, oriented and/or managed [3]. Then, the results from the large and optimized PV plants can be considered as references for calibrating the simulation tool we develop in this article. Also, it is important to mention the need to work with data from PV plants injecting electricity to the grid instead of installed capacity, as it is many times reported, because there are cases where PV plants are out of service or not connected to the grid but accounted for official statistics about PV capacity.

The datasets have been filtered to guarantee the consistency of the values, the periodicity required for this study and for avoiding any error in terms of duplicated information. No data filling has been necessary because of any lack of missing data at any period. Also, for calculating the specific yield (kWh per kW$_p$ in a time period) of each PV plant per year we have avoided the months without energy production in the first year of operation. Moreover, as the aim of this paper is mostly devoted to the understanding of the PV potential in the Canary Islands and not purely the economic issues related, months with zero specific yield have not been considered for calculating average outputs of each PV plant, as these zeros can be related to a breakdown in the PV system, but also to technical, administrative or legal conflicts. Moreover, data about the geographic location of each PV plant, nameplate capacity, and voltage for grid integration were considered in the analysis.

The geographic location of each PV plants has been provided by

ENDESA and it has been used for obtaining the 153,120 solar irradiance, temperature and wind speed data used in this work, adding also irradiance, temperature and wind speed data from the most important urban areas. This amount of irradiation data is important to be considered for defining the solar resource because most Canary Islands have an abrupt relief. Therefore, different microclimatic conditions are affecting the performance of close PV plants as the solar resource varies substantially between close locations. It is quite usual to filter irradiance data based on their low accuracy, introducing different conditions [15]. In this case, a simple filtering data process for the different solar resource datasets used has been necessary, as it will be described below. No missing irradiance data at certain time periods was detected in the solar resource databases used. Then, no approaches were used to fill missing data, in contrast with other works [16].

The global horizontal irradiation (GHI) has been downloaded from the PVGIS-SARAH, PVGIS-ERA5 and PVGIS-CMSAF databases [17,18] on a monthly base from 2005 to 2016. Recent reports from the International Energy Agency (IEA) show that typical normalized root mean square errors for satellite monthly irradiation reported in the literature are between 4% and 8% [14,19]. Of course, irradiance measurements with calibrated and well maintained on-site radiometers are much more precise than satellite data, first and second class pyranometers (ISO 9060 classification), and silicon irradiance sensors, but such stations are not common. Then, the precision of the satellite services is generally comparable or even better than that of the on-site measurements for monthly performance reporting for small to medium size PV plants [19]. The optimized slope angle used for the PV plants in the Canary Islands should be about 30°. However, there are many ground mounted PV plants tilted at 10° and other angles. A 30° slope should be used as the Canary Islands are situated at a range of latitudes from 27°35′N to 29°23′N, but the 10° slope is used as an alternative to avoid additional consumption of land and to produce a low visual impact in islands with a high environment protection.

For a standard model about the performance of a PV plant, input about PV module orientation is essential [20]. However, orientation data is not typically collected by utilities [3]. Also, as we plan a simple methodology only based on data provided by utilities and satellites, input about orientation is avoided in the model.

For defining the upper limit of the monthly specific yield in a PV plant located in the Canary Islands, a simple clear sky model [21] has been used. No consideration to PV bifacial technology [22] is introduced as most PV plants have been built before 2012 and we have no information of any PV plant based on bifacial technology. Then, we avoid the introduction of any rear surface irradiance term for calculating the clear sky values. Thus, for obtaining the potential specific yield at any location of the Canary Islands, an equation has been defined considering average temperature and wind speed in the area and the year the PV plant started supplying electricity to the grid.

3. Theory

There are different approaches currently used for introducing the temperature correction in the efficiency of the PV cell [23–28]. In our case, the temperature correction is introduced following Evans [29], and recently applied by Nobre et al.[30], where the efficiency at the maximum power point at a defined temperature, $\eta_{mpp}(G_{mod}, T_{mod})$, compared to the efficiency at the maximum power point at 25 °C, $\eta_{mpp,25}(G_{mod})$, depends on the irradiance on the plane of the module, G_{mod}, and the module temperature, T_{mod}:

$$\eta_{mpp}(G_{mod}, T_{mod}) = \eta_{mpp,25}(G_{mod}) \cdot [1 + b(T_{mod} - 25^oC)] \qquad (1)$$

where b is the temperature coefficient. We consider –0.0035 °C^{-1} as the value for b because almost all PV modules placed in the Canaries can be considered as based on silicon [4].

To estimate the module temperature, we consider also the average

wind speed because many locations, particularly in coastal areas, are very windy. Then, we have used a parameter described for the wind influence in the determination of the module temperature [31], the wind data [32], and the average wind speed values obtained for an specific height. The wind speed, v [m/s], at a height, z [m], above the ground level is obtained applying the equation [33]:

$$v = v_{ref} \cdot \frac{\ln\left(\frac{z}{z_0}\right)}{\ln\left(\frac{z_{ref}}{z_0}\right)}$$

(2)

where v_{ref} is the reference speed; z, is the height above the ground level for the desired velocity, v; and z_0 is the roughness length in the wind direction (we consider $z_0 = 0.03$ m and $z = 2$ m for PV plants mounted on ground, and $z_0 = 0.4$ and $z = 1$ for PV plants mounted on roof).

Then, the temperature of the module can be described as follows:

$$T_{mod} = T_{amb} + c \cdot G_{mod} - a(v-1)$$

(3)

where T_{amb} is the ambient temperature, c the Ross coefficient [34], and a is an empirical parameter (2.42 °C s/m for mc-Si [31]). It is expected that the a coefficient should be somewhat higher, as the daily wind speed profile in coastal areas of the Canary Islands is quite similar to the clear sky irradiance profile [2]. Moreover, in this paper we consider that all PV systems are free standing installations with improved ventilation compared to roof integrated systems, as almost all PV modules installed in the Canaries are placed on aluminum structures to offer a defined tilt angle. Then, the Ross coefficient can be approximated to 0.02 °C m²/W [4].

For introducing the temperature losses in our estimations of PV module output we introduce the average ambient temperature and wind into the efficiency at the maximum power point at a defined temperature and irradiance at $x°$ tilt:

$$\eta_{mpp}(G_{x°}, T_{amb}) = \eta_{mpp.25}(G_{x°}) \cdot [1 + b(T_{amb} + c \cdot G_{x°} - a(v-1) - 25°C)]$$

(4)

where $G_{x°}$ [W/m²] is the irradiance at $x°$ slope.

The PV system losses, as the losses related to the efficiency of the inverters and other miscellaneous losses (e.g., mismatch between modules and ohmic cable losses) are considered as a 9.5% of the module output [4]. This approximation is fair for well managed PV plants, but the value is larger when small and not so well managed PV plants are under consideration [7]. In our case, most multi-MW PV plants on ground can be considered as well managed. Degradation rates for real PV systems are mostly considered as variable, larger in the first year of operation, and calculated over monthly values of performance [14,19]. However, an average 0.5% loss of power per year of operation of the PV system can be considered [17], and it is introduced in the final equation. There were only PV plants with a combined power capacity of 425.09 kW in 2005 compared to 166.75 MW in 2017. Then, we consider that the 0.5% power loss per year started in 2005 for the PV plants injecting electricity to the grid in 2005. There are other losses effects not considered in these equations as snow, partial shadowing, dust and dirt.

Then, the overall losses can be described by the L correction factor, defined as:

$$L = 0.905[1 - 0.005(Y_i - Y)]$$

(5)

where Y is the year when the PV plant started injecting electricity to the grid and Y_i is the year when the specific yield is determined.

The efficiency of average commercial wafer-based silicon modules is considered to increase from about 12–17% in the period 2005–2017 [1]. This increase in efficiency is considered quasilinear in our model and applied to all the modules in the PV plant under consideration. Thus, a linear dependence between the year, Y, when the PV plant starts injecting electricity to the grid, and the efficiency of the PV modules, η_{mod}, is defined:
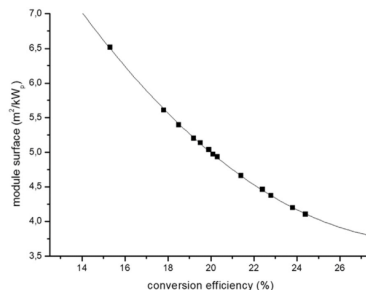

Fig. 1. Polynomial adjustment between conversion efficiency and surface obtained from the world record efficiencies reported in the literature for PV modules [35,36].

$$\eta_{mod}(\%) = 12 + (Y-2005) \cdot 0.417$$

(6)

Then, the estimation of the potential specific yield is calculated for each year i using the following equation:

$$E_i\left(\frac{kWh}{kW_p}\right) = G_{x°} \cdot \eta_{mpp} \cdot \eta_{mod} \cdot S \cdot L$$

(7)

being S [m²/kW$_p$] the surface per unit of capacity for the modules that constitute the PV system.

For this calculation we have obtained a relation between efficiency (in %) and surface (in m²) considering the record confirmed terrestrial large c-Si and mc-Si module efficiencies measured under the global AM1.5 reported from June 2007 [35] to June 2018 [36] (Fig. 1). The following equation is obtained:

$$S(m²) = 15.42346 - 78.379\eta_{mod} + 131.2\eta_{mod}^2$$

(8)

We do not consider changes in the distribution of the solar spectral irradiance during the daytime and how they influence the module performance because this influence is mostly for high band gap semiconductor materials [37,38].

For example, from the equations described above, the following equation for the potential specific yield, E_i, is found depending on the irradiance at 30° tilt, $G_{30°}$, ambient temperature, T_{amb}, the year when the PV module starts injecting electricity to the grid, Y, and the year when the output is obtained, Y_i:

$$E_i\left[\frac{kWh}{kW_p}\right] = G_{30°}[1 - 0.0035 \cdot (T_{amb} + 0.02 \cdot G_{30°} - 25 - 2.42 \cdot (v-1)] \cdot [0.12$$
$$+ (Y-2005) \cdot 0.00416] \cdot \{15.42346$$
$$- 78.379 \cdot [0.12 + (Y-2005) \cdot 0.00416]$$
$$+ 131.2 \cdot [0.12 + (Y-2005) \cdot 0.00416]^2\} \cdot 0.905 \cdot [1$$
$$- 0.005 \cdot (Y_i - Y)]$$

(9)

Finally, it is important to mention that most calculations have been made using the Rstudio software. The results exposed in this work can be more disaggregated. However, the results are exposed in a level of aggregation that protects information from individual PV systems, as it is legally required in Spain.

4. Results and discussion

The PV production in the Canary Islands (2017) is composed by 1582 registered PV plants. 59 PV plants are new installations still pending for starting producing electricity at the end of 2017, and 29 PV

Renewable and Sustainable Energy Reviews 107 (2019) 123–132

Table 1
Voltage connection range for the PV systems injecting electricity to the power grid in 2017.

V (kV)	No. of PV plants
0.22–0.24	43
0.38–0.40	1138
15	10
20	288
66	15

plants are registered but currently not producing electricity. 99.0% of the PV plants are connected to the distribution grid (Table 1), as 66 kV and above is considered voltage for transmission in Spanish insular power systems [39]. However, as the voltage for households is 220 ~ 240 V, the results show a very low penetration of PV systems at the household's level. Also, many of the multi-MW power plants are composed by 100 kWp PV plants connected at 0.38–0.4 kV because of the specific regulation in force when the owners planned the investment, as it is discussed below. Then, a distinction between PV plants considering if they are connected to the transmission or the distribution grid gives a misleading first-step information about the characteristics of the ownership and nameplate capacity of the PV plants.

In this work, data of PV capacity is related to PV plants injecting electricity to the grid instead of installed capacity (as it is usually reported) because there are differences between both quantities. Only a small PV capacity was injecting electricity to the grid before 2005 because of the expensiveness of the technology in that time and the lack of feed in tariffs offered by the Government to make it profitable (Figs. 2a and 2b). However, the approval in 2004 of a Royal Decree (RD 436/2004) with an attractive feed in tariff [40] and the reduction in PV prices over the following years [1] produced a sharp increase in PV capacity (Fig. 2). This sharp increase in capacity was not homogeneously distributed across the 7 power grids that compose the Canary Islands power system [41]. On the contrary, most of the 166,750 kW PV

capacity (64.0%) is currently injecting energy to the grid in Tenerife because of the quick adaptation of the territorial rules about land use for producing PV energy in this island. Other power grids have almost no PV installations or far below Tenerife mostly because political decisions that will be discussed below.

As the investment cost for new PV systems continued falling, two new Royal Decrees were approved for achieving a substantial reduction on the installation of new PV plants: (i) RD 661/2007, reducing the feed in tariff (2007) [42]; and (ii) RD 1578/2008, establishing a new procedure for defining the value of the feed in tariffs applied in future (2008) [43]. Consequently, the rate of new PV systems installed was reduced for a short period, but the PV sector increased its activity again in 2010. The activity of installing PV plants (and any renewable power plant) finally decreased dramatically after the approval of an unexpected Royal Decree-Law (RDL) in 2013 (RDL 1/2013) [44] suspending any feed in tariff for new installations. A small number of investors still connected new and significant amounts of PV capacity in 2013 and 2014 due to previous compromises to the approval of this RDL. In parallel and for the first time, decreases in PV capacity are detected in some insular power grids, mostly attributed to the fact that the newest PV plants were not capable to operate without feed in centives. The approval of a new Ministerial Order (IET/1459/2014) [45] with a new feed in tariff for the PV systems installed in the Canary Islands further decreased the retribution. Finally, a new Royal Decree (RD 900/2015) was approved in 2015 regulating self-consumption and net metering [46]. This new regulation was justified by the government as a requisite for aligning this option with customers that should continue contributing to the maintenance of the power system with a similar cost than before the installation of the PV system. The result was to reach the lowest new installed PV capacity in the period under study. However, in 2017 the installation of new PV capacity started growing again, attributed to the fact that the cost of the PV plants continued falling in international markets and started being profitable also without feeding tariffs and under the current regulation. The evolution of PV capacity injecting energy to the grid and the date when new

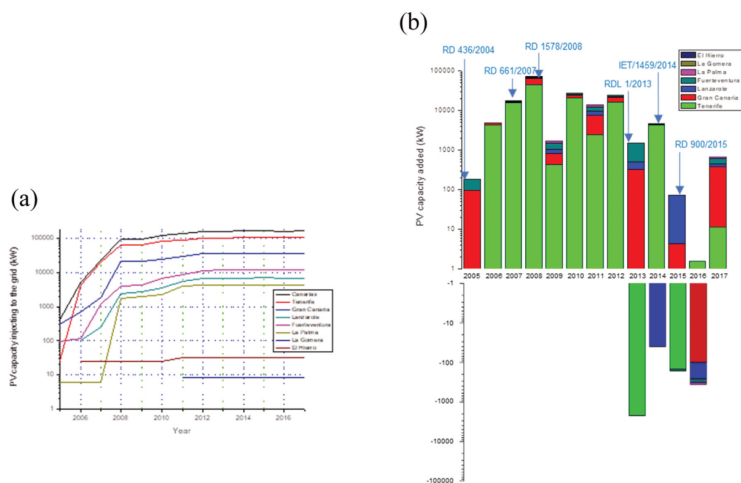

Fig. 2. Evolution of the (a) cumulative PV capacity injecting electricity to the different power grids of the Canary Islands in the period 2005–2017, and (b) the new PV capacity annually added to the power grids in the same period.
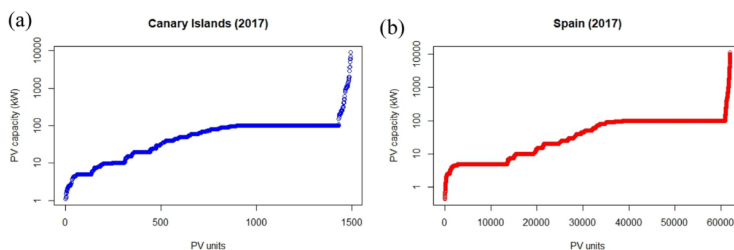
126

116

Fig. 3. PV plants injecting electricity (a) in the Canary Islands, and (b) Spain in 2017 sorted by PV capacity.

regulation was in force are illustrated in Fig. 2b.

Thus, in the Canary Islands most PV plants (503) have a capacity of 100 kW (Fig. 3a) due to the higher feed in tariff introduced in 2004 [40] for PV plants with 100 kW nameplate capacity and below. Most of these 100 kW PV plants are integrating multi-MW PV plants, and none of these PV plants were injecting electricity before 2006. A similar scenario is produced considering the PV plants installed in Spain (Fig. 3b).

As it is exposed in Fig. 4, the distribution of the PV capacity along the municipalities of the Canary Islands is related to the solar resource but mostly to the availability of land with high irradiance and suitable for the location of the PV systems. Only Tenerife and the municipality of Arico had in 2008 a regulation in force and a simple procedure for placing renewable power plants on ground in unproductive areas. In other islands the regulation for placing these power plants was slow and tedious. Also, there were islands that initially rejected the installation of PV systems because of environmental concerns (Fuerteventura) or because the government was mainly interested in supplying most electricity with a new hydro-wind power station (El Hierro). The price of the land, the configuration of the power grid and the insular power capacity were other factors influencing the process but in a much lesser extent.

As a first approximation for defining the upper limit of the monthly specific yield in a PV plant located in the Canary Islands, the simple clear sky model referenced in the experimental section [21] has been used. The month of maximum solar radiation (July) at the sea level at the most southerly area of the Canary Islands (27°35′N) and default parameters of the model were used, resulting 261.1 kWh/m^2 month. For the upper limit of the specific yield in a PV plant located in the Canary Islands we consider the 24 h average temperature in the area in July from 2005 to 2016 (22.1 °C), a windy area (6.5 m/s), and a PV module produced in 2017 (i.e., 17% conversion efficiency and no losses of power per years of operation). To this value, the 9.5% PV system losses also described in the experimental section are introduced, resulting an upper limit of the specific yield in a PV plant located in the Canary Islands as 199.2 kWh/kW$_p$·month. We also consider that there are a few PV plants with tracking systems. These PV plants could produce about a 40% more energy in a 2-axis tracking configuration [17]. Then, the upper limit of the specific yield in a 2-axis tracking PV plant located in the Canary Islands is 278.9 kWh/kW$_p$·month. The specific yield results 1964.9 kWh/kW$_p$·yr for a PV plant with no tracking system and an average annual temperature of 19.8 °C, 6.5 m/s average wind speed and 9,5% PV system losses (2750.9 kWh/kW$_p$ yr for a 2-axis tracking system).

Then, all raw data of specific yield above 199.2 kWh/kW$_p$ month have been analyzed to check if the values are consistent with the expected annual output on a monthly basis for the plant capacity. The analysis shows that most specific yields above 199.2 kWh/kW$_p$ month can be attributed to meter readings larger than in a monthly basis and compensated by other meter readings in shorter periods of time in the previous or following months. These data are about a very small percentage of the total. Only a small amount of the specific yield above 199.2;kWh/kW$_p$ month can be attributed to PV plants with tracking systems but giving service in the tracking configuration in periods shorter than a year.

The monthly specific yields of all PV plants in the Canary Islands are exposed in Fig. 5. The 0 kWh values have been removed because one of the objectives of this work is to estimate the specific yield of the PV plants when they are injecting electricity to the grid, and some of the zeros are not directly related to technical failures of the plants but to administrative issues or legal conflicts, as discussed above. The average monthly output in the Canary Islands is 129.6 kWh/kW$_p$ month (1555.2 kWh/kW$_p$ yr) and the highest specific yield in the Canary Islands is 205.8 kWh/kW$_p$ month (2469.6 kWh/kW$_p$ yr) for a single PV plant placed in Fuerteventura. This value is the only one for all Canary Islands PV systems not consistent with the clear sky limit exposed above for a 2-axis tracking system (278.9 kWh/kW$_p$ month). There are a few other PV plants that should also be considered not consistent with the clear sky limit as almost all PV plants in the Canaries have no tracking system. However, this PV system located in Fuerteventura has reached higher specific yields than the 2-axis clear sky limit up to 6 months on a continuing basis in the period 2009 – 2013 (a record of 382.3 kWh/kW$_p$ month in May 2013). It is important to mention that the new feeding tariff (IET/1459/2014) [45] mainly based on PV capacity (and not energy produced) was approved in 2014. Then, the incentives were substantial before 2014 but they are at present very small for producing energy higher than a value defined in the new regulation. Therefore, this PV plant should be analyzed on-site to better understand this unexpected behavior. As it is the Spanish Energy Commission, and not the utilities, operators or traders the entity in charge of remunerating the feed in tariffs to PV producers, there are no incentives to these companies for inspecting the PV production above clear sky limits, compared to inspecting activities for detecting fraud in connections to the grid. Thus, an alert system based on clear sky values is recommended to be adopted by the Spanish Energy Commission.

We have also determined the average specific yield of the Canary Islands PV plants in the period 2005–2017 considering only the PV plants injecting power to the grid the whole year (Fig. 6a). A substantial increase in performance is detected in the period 2008–2009, attributed to a huge increase in new PV capacity injecting power to the grid, based on new technology placed in best locations and orientations. A soft decrease in performance is detected after 2009, attributed to the expected power loss per year of operation. We have also determined the 8 municipalities with higher PV specific yield (Fig. 6a) in the period 2005–2017 (an overall performance in the period has no sense because the fleet of PV plants has not been constant in the period). We have not included municipalities with less than 5 PV systems and combined PV capacities lower than 40 kW for avoiding errors due to the small amount of samples. It can be observed that there exists a good
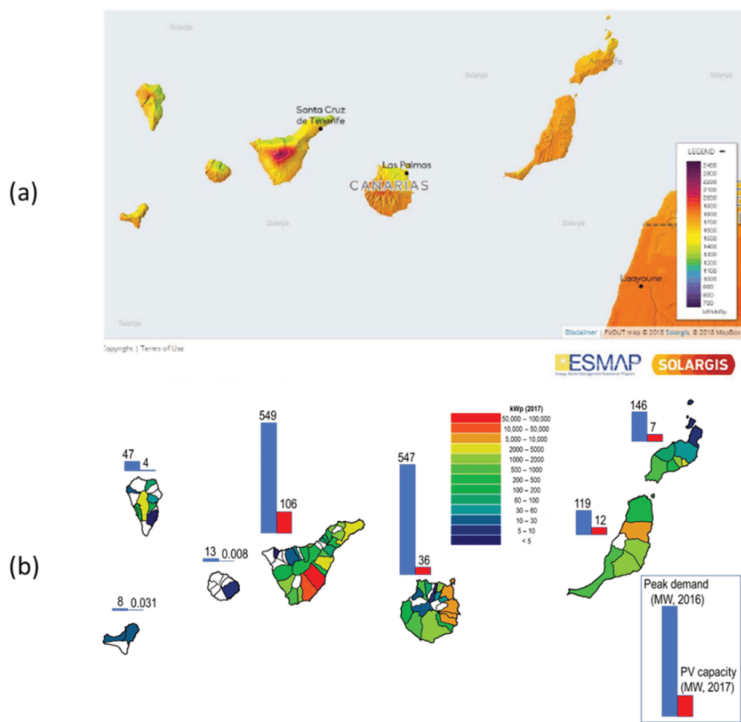
Fig. 4. (a) Irradiance map of the Canary Islands [52]; and (b) distribution of PV capacity in the municipalities of the Canary Islands in 2016, and power peak demand in 2016 (last data available).

agreement between solar irradiation (Fig. 4a), municipalities with the largest share of PV capacity installed (Fig. 4b), and municipalities with the largest specific yield (Fig. 6b).

From Fig. 6 we can analyze if the performance has any dependence with the capacity of the PV plant. For analyzing a homogeneous sample we select PV plants with different capacities but installed in the same municipality and starting injecting electricity to the grid the same year. For this case Telde and the year 2010 is the most appropriate choice, as this municipality reached the highest amount of new installed PV systems in a single year (50 new PV plants ranging from 6.9 to 450 kW$_p$, all of them mounted on roof). The results show (Fig. 7) that no clear relation exists between the performance and the capacity of the PV plant for this range of capacities. A larger amount of samples with a higher range of capacities is required for a more accurate conclusion.

Also the PV specific yield has been studied to find any dependence with the placement of the PV plant on roof or on the ground, as there are some conflicts about uses of available land and how to develop renewable energy [47]. For this objective, we have considered Agüimes because in 2009 this municipality installed 5×100 kW$_p$/each PV plants and placed on the ground, and there are other 6 PV plants totaling 466 kWp placed on roof which started injecting electricity to the grid the same year. We avoid using the starting year for the calculation,

as the PV plants started injecting in different months. The results show an average 11.8% higher specific yield for the PV plants placed on the ground in a period of 5 years (Table 2) compared to systems installed on roof. These differences in specific yield may be attributed to an improved design and management, optimal location and a well-defined nameplate capacity for the PV systems placed on ground. However, analysis in-site is needed for supporting these hypotheses. An alternative analysis could be to consider that PV plants connected to the grid at 20–66 kV are on the ground, and the remaining on roof, studying only some defined municipalities to guarantee that the same solar resource is being used. However, it has been confirmed that there are multi-MW PV plants divided in 100 kW$_p$ plants that are connected to the grid at $0.38 \sim 0.40$ kV each.

Also, it is interesting to consider a multi-MW PV plant located in a defined municipality composed by many 100 kW$_p$ plants with individual counters, starting into operation the same year and checking the average production. Results show that the 100 kW$_p$ PV plant with the highest specific yield is 8.6% higher than for the plant with the lowest specific yield in a period of 9 years. In another case the difference is a 10.0% (Table 2). These differences in specific yield could be partly attributed to PV plants that have PV modules from different manufacturers in the same multi-MW PV plant, a fact that has been
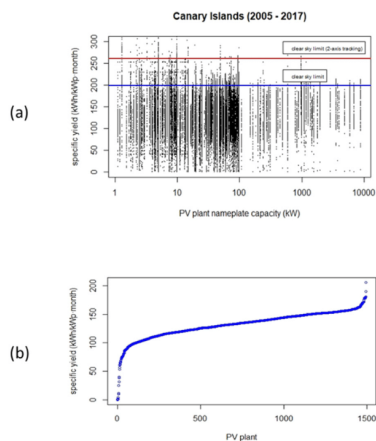
128

118

Fig. 5. (a) Specific yields of all PV systems located in the Canary Islands in the period 2005–2017, and (b) average specific yield for all PV systems located in the Canary Islands.

confirmed after inspecting the PV plants.

We have also determined how the kWh/kW$_p$ should diminish with time using Eq. (9). The first step has been to check which solar irradiance database best fit the specific yield of well managed PV power plants in terms of annual variation. To this purpose we have use the data of all 100 kW$_p$ PV plants of 3 multi-MW plants placed in the

municipalities with the largest PV capacity installed. The first step has been to select the irradiance data base, because we have detected a substantial mismatch in 2016 for the PVGIS-SARAH database compared to the other two databases and the specific yield of the PV plants in different municipalities (Fig. 8). Then, the PVGIS-SARAH database has been discarded. Of course, in case of doubt it is always recommended to evaluate the quality of the satellite data with ideally a set of data from a well-maintained meteorological station in the neighborhood [19], but this is beyond our objectives. Also the PVGIS-ERA5 is discarded for this analysis because the data provided is short in terms of years of irradiance offered for analyzing our dataset.

Fig. 9 shows the real and estimated specific yield for three multi-MW PV plants placed in the municipalities with the best solar resource (Arico and Granadilla). A soft decrease in real specific yield is observed from 2015 for two of the PV plants. This decrease can be considered about a 3% in PV specific yield (Table 2) and can be attributed to a change in the retribution system approved by the Spanish Government in 2014 (IET/1459/2014)[45]. This new retribution system makes unattractive to clean the PV systems, as most of the retribution is based on the capacity instead of the energy produced if a minimum amount of hours producing energy is reached, as it has been described elsewhere [2]. The decrease in performance of a PV plant varies substantially depending on the location and the source of dust [48], and the 3% decrease in specific yield obtained for these multi-MW PV plants is in accordance to similar values obtained for large PV systems in quite similar climatic areas [14,49]. In our case we consider that the decrease in performance is low because the PV plants studied are placed in windy locations where heavy rains are produced occasionally. In one of these two PV plants the lack of cleaning starting in 2015 has been confirmed.

Also, the results exposed in Table 3 show that the multi-MW PV plants studied have specific yields that are very close to the clear sky limit defined above and applied to these municipalities. Then, the procedure developed in this work could offer a good first approximation to evaluate the performance of any PV plant in terms of its location and the age of the PV system. Moreover, the specific yield of these
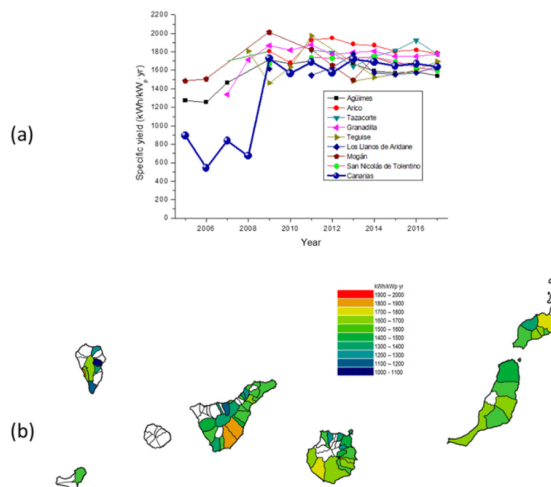

Fig. 6. (a) Evolution of the average specific yield for the Canary Islands municipalities in the period 2005–2017; and (b) colored map of the Canary Islands municipalities in terms of average specific yield in the period 2005–2017.
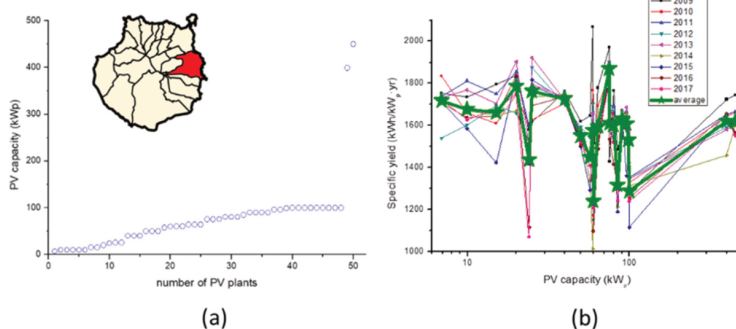
129

Fig. 7. (a) Distribution of PV nameplate capacities installed in Telde in 2010, and (b) average annual specific yield for the PV plants installed in 2010 in terms of nameplate capacity ranges.

Table 2
Dependence of the PV specific yield in relation to PV plants placed on ground or on roof, and differences in PV specific yield for $100\,kW_p$ plants placed in a multi-MW PV plant.

Municipality	Characteristics	Difference	map
Agüimes	on ground / on roof	+ 11.8%	
Granadilla and Arico	Multi-MW $100\,kW_p$ PV plants	8.6–10.0%	

Fig. 9. Real and simulated specific yields for three multi-MW PV plants.

Table 3
Difference in specific yield between the expected values obtained in Eq. (9) and the real data.

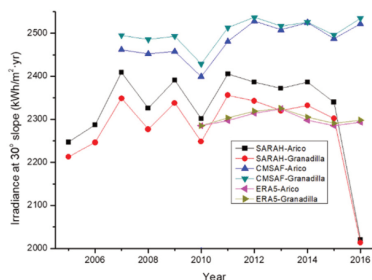|	PV1	PV2	PV3
2009		− 1,87%	− 2,24%
2010		1,15%	− 1,67%
2011	2,02%	2,84%	− 0,75%
2012	− 0,49%	− 0,71%	− 6,27%
2013	− 0,59%	1,71%	− 4,12%
2014	− 1,35%	1,91%	− 3,48%
2015	− 2,90%	− 0,55%	− 4,87%
2016	− 5,55%	− 2,01%	− 5,33%

Fig. 8. Average annual irradiance for two different locations obtained from the PVGIS SARAH, CMSAF and ERA5 [17].

multi-MW PV plants can act as references for calibrating our procedure and for comparisons with the specific yields of other PV plants.

We can also determine the slope of PV plants using the simulation and irradiance database. There are works in literature introducing algorithms for determining the location and orientation of PV plants, but based on the daily energy profile [50], and with 5 min resolution for increasing precision [3]. As we can check in Fig. 10, the multi-MW PV3 plant match with a higher accuracy the 10° slope than the more
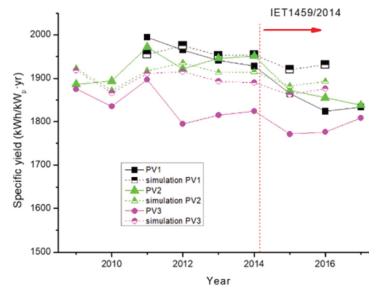
efficient 30° slope defined for this latitude. It means that PV3 could have about a 3.6% higher output with the 30° slope configuration but because of the reasons exposed above the 10° slope configuration is applied in many PV systems in the Canary Islands. However, in future, with a high penetration of PV in the power system, it is expected that the price of electricity in the whole market will decrease with the increase of irradiance, as it is currently happening in California [51]. Then, a higher energy production in months of low irradiance is expected to have a higher retribution than in summer.

We have also determined the monthly losses in capacity by zeros in specific yield for PV plants in operation at defined capacity ranges

130

120

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
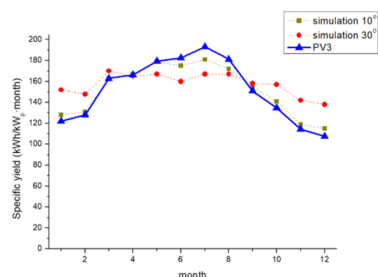Fecha: 02/06/2021 14:57:08

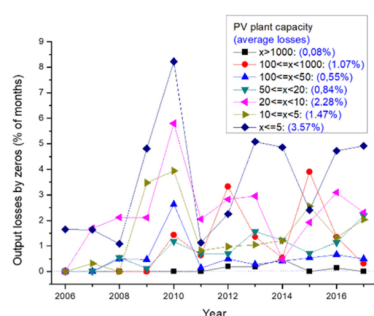**Fig. 10.** Specific yield per month in 2010–2016 for the PV3 plant.

Fig. 11. Output losses due to monthly zeros injected in the grid for different PV capacity ranges.

(Fig. 11). For this purpose, we have only counted zeros that are produced in months after other months where the PV plant has been producing energy. We have not included 2005 as we have no values from 2004 to check which PV plants were producing electricity before January 2005. It is observed that PV plants with lower capacity tends to fail more injecting electricity to the grid, and average losses are below 3.6% for all capacity ranges. If counting zeros gives no enough information about the failure of a PV plant, an alternative could be to count monthly specific yields below certain values (instead of zeros) that could be considered as periods where the performance of the PV plant was below a minimum expected in the area.

5. Conclusions

In general, PV plants are expected to be installed where there is enough solar radiation to make the technology attractive, but in Canary Islands the potential irradiation is being considered in combination with other more important parameters, mostly the easiness of the regulation in force for installing PV systems in the area.

The equation for determining the potential specific yield of any PV plant in the Canary Islands in terms of the location and only using geographic information available has been validated and offers a very precise approximation to the real performance of well managed PV plants. This equation can be also very useful to prevent unrealistic specific yields and a lack of regular maintenance in the PV system as, for example, lack of regular cleaning of the PV modules because of soiling, or to determine the slope of the PV modules.

The evolution of the PV capacity in the Canary Islands has been mostly defined by the regulation approved at the country level and it is observed to be quite similar to the whole country. A defined relation between performance of the PV plant and the PV nameplate capacity is not found. A slightly higher specific yield for PV plants installed on ground compared to PV plants on roof are expected. Also, a quite similar variability in specific yield is observed for PV plants with similar characteristics in a multi-MW PV plant. Additionally, it is observed that PV plants with low nameplate capacity tends to fail more injecting electricity to the grid than large ones.

Finally, an alert system based on clear sky values is recommended to be adopted by the Spanish Energy Commission. It is also necessary to mention that it will be very useful if power system operators can provide the research institutions with disaggregate information of PV electricity produced in defined power systems, because similar and helpful analyses could be produced in other areas around the world.

References

[1] ISE F. Photovoltaics REPORT. Freiburg, Germany; 2018.
[2] Guerrero-Lemus R, González-Díaz B, Ríos G, Dib RN. Study of the new Spanish legislation applied to an insular system that has achieved grid parity on PV and wind energy. Renew Sustain Energy Rev 2015;49. https://doi.org/10.1016/j.rser.2015.04.079.
[3] Haghdadi N, Copper J, Bruce A, MacGill I. A method to estimate the location and orientation of distributed photovoltaic systems from their generation output data. Renew Energy 2017;108:390–400. https://doi.org/10.1016/J.RENENE.2017.02.080.
[4] Lorenz E, Scheidsteger T, Hurka J, Heinemann D, Kurz C. Regional PV power prediction for improved grid integration. Prog Photovolt Res Appl 2011;19:757–71. https://doi.org/10.1002/pip.1033.
[5] Passey R, Watt M, Bruce A, MacGill I. Who pays, who benefits? The financial impacts of solar photovoltaic systems and air-conditioners on Australian households. Energy Res Soc Sci 2017;39:198–215. https://doi.org/10.1016/J.ERSS.2017.11.018.
[6] Wirth G, Lorenz E, Spring A, Becker G, Pardatscher R, Witzmann R. Modeling the maximum power output of a distributed PV fleet. Prog Photovolt Res Appl 2015;23:1164–81. https://doi.org/10.1002/pip.2513.
[7] Leloux J, Narvarte L, Trebosc D. Review of the performance of residential PV systems in Belgium. Renew Sustain Energy Rev 2012;16:178–84. https://doi.org/10.1016/J.RSER.2011.07.145.
[8] Bertrand C, Housmans C, Leloux J, Journée M. Solar irradiation from the energy production of residential PV systems. Renew Energy 2018;125:306–18. https://doi.org/10.1016/J.RENENE.2018.02.036.
[9] Leloux J, Narvarte L, Trebosc D. Review of the performance of residential PV systems in France. Renew Sustain Energy Rev 2012;16:1369–76. https://doi.org/10.1016/J.RSER.2011.10.018.
[10] Kelsey N, Meckling J. Who wins in renewable energy? Evidence from Europe and the United States. Energy Res Soc Sci 2018;37:65–73. https://doi.org/10.1016/J.ERSS.2017.08.003.
[11] Li Y. Incentive pass-through in the California Solar Initiative – an analysis based on third-party contracts. Energy Policy 2018;121:534–41. https://doi.org/10.1016/J.ENPOL.2018.07.015.
[12] World electric power plants database: global market data and price assessments - Platts n.d. ⟨https://www.platts.com.es/products/world-electric-power-plants-database⟩ [Accessed 15 September 2018].
[13] NASA. Atmospheric science data center. surface meteorology and solar energy. n.d. ⟨https://eosweb.larc.nasa.gov/⟩ [Accessed 24 August 2018].
[14] IEA Pvps Task-13. PV Performance modeling methods and practices.
[15] Gueymard CA, Ruiz-Arias JA. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol Energy 2016;128:1–30. https://doi.org/10.1016/J.SOLENER.2015.10.010.
[16] Wilcox S. National solar radiation database 1991–2010 Update: User's Manual; 1991.
[17] uropean Commission n.d. ⟨http://re.jrc.ec.europa.eu/pvgis.html⟩ [Accessed 19 July 2018].
[18] Huld T, Müller R, Gambardella A. A new solar radiation database for estimating PV performance in Europe and Africa. Sol Energy 2012;86:1803–15. https://doi.org/10.1016/J.SOLENER.2012.03.006.
[19] IEA Pvps Task-13-12. Uncertainties in PV system yield predictions and assessments. 2018.
[20] Narvarte L, Lorenzo E. Tracking and ground cover ratio. Prog Photovolt Res Appl 2008;16:703–14. https://doi.org/10.1002/pip.847.

131

121

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

[21] Bird R. Bird Clear Sky Model. ⟨https://www.nrel.gov/grid/solar-resource/clear-sky.html⟩ [Accessed 3 August 2018]; 1981.

[22] Guerrero-Lemus R, Vega R, Kim T, Kimm A, Shephard LE. Bifacial solar photovoltaics - a technology review. Renew Sustain Energy Rev 2016:60. https://doi.org/10.1016/j.rser.2016.03.041.

[23] Dolara A, Leva S, Manzolini G. Comparison of different physical models for PV power output prediction. Sol Energy 2015;119:83–99. https://doi.org/10.1016/J.SOLENER.2015.06.017.

[24] King DL, Boyson WE, Kratochvill JA SAndia report Photovoltaic array performance model. n.d.

[25] Faiman D. Assessing the outdoor operating temperature of photovoltaic modules. Prog Photovolt Res Appl 2008;16:307–15. https://doi.org/10.1002/pip.813.

[26] Ross RGJ, Smokler MI. Flat-plate solar array project. Eng Sci Reliab 1986;6.

[27] Skoplaki E, Palyvos JA. Operating temperature of photovoltaic modules: a survey of pertinent correlations. Renew Energy 2009;34:23–9. https://doi.org/10.1016/J.RENENE.2008.04.009.

[28] Luketa-Hanlin A, Stein JS. Improvement and validation of a transient model to predict photovoltaic module temperature. World Renew. Energy Forum, WREF 2012, Denver, CO; 2012.

[29] Evans DL. Simplified method for predicting photovoltaic array output. Sol Energy 1981;27:555–60. https://doi.org/10.1016/0038-092X(81)90051-7.

[30] Nobre AM, Severiano CA, Karthik S, Kubis M, Zhao L, Martins FR, et al. PV power conversion and short-term forecasting in a tropical, densely-built environment in Singapore. Renew. Energy 2016;94:496–509. https://doi.org/10.1016/J.RENENE.2016.03.075.

[31] Segado PM, Carretero J, Sidrach-De-Cardona M. Models to predict the operating temperature of different photovoltaic modules in outdoor conditions; 2014.

[32] DTU. Global wind atlas n.d. ⟨https://www.globalwindatlas.info/⟩ [Accessed 14 August 2018].

[33] Danish wind industry association. roughness and wind shear n.d. ⟨http://drømstørre.dk/wp-content/wind/miller/windpowerweb/en/tour/wres/shear.htm⟩ [Accessed 14 August 2018].

[34] Ross R. Interface design considerations for terrestrial solar cell modules. In: Proceedings of the 12th IEEE Photovolt. Spec. Conference, Baton Rouge, LA, USA; 1976.

[35] Green MA, Emery K, Hisikawa Y, Warta W. Solar cell efficiency tables (version 30). Prog Photovolt Res Appl 2007;15:425–30. https://doi.org/10.1002/pip.781.

[36] Green MA, Hishikawa Y, Dunlop ED, Levi DH, Hohl-Ebinger J, Ho-Baillie AWY. Solar cell efficiency tables (version 52). Prog Photovolt Res Appl 2018;26:427–36. https://doi.org/10.1002/pip.3040.

[37] Fanney AH, Davis MW, Dougherty BP. Short-term characterization of building integration photovoltaic panels. ASME Sol. 2002 Int. Sol. Energy Conf. American Society of Mechanical Engineers; 2002. p. 211–21.

[38] Kratochvil JA, Boyson WE, King DL. Photovoltaic array performance model; 2004.

[39] Ministerio de Industria Energía y Comercio. Resolución de 24 de julio de 2012, de la Secretaría de Estado de Energía, por la que se aprueba la modificación de los procedimientos de operación del Sistema Eléctrico Peninsular (SEP) P.O.-3.1; P.O.-3.2; P.O.-9 y P.O. – 14.4 y los procedimientos de operación de los Sistemas eléctricos Insulares y Extrapeninsulares (SEIE) P.O. SEIE-1 P.O. SEIE-2.2; P.O. SEIE-3.1; P.O. SEIE-7.1; P.O. SEIE-7.2; P.O. SEIE-8.2; P.O. SEIE – 9 y P.O. SEIE-2.3 para su adaptación a la nueva normativa eléctrica. Resolución; 2012.

[40] Real Decreto 436/2004, de 12 de marzo, por el que se establece la metodología para la actualización y sistematización del régimen jurídico y económico de la actividad de producción de energía eléctrica en régimen especial. Real Decreto; 2004.

[41] Consejería de Economía, Industria C y C. Anuario Energético de Canarias. 2016; 2016.

[42] Real Decreto 661/2007, de 25 de mayo, por el que se regula la actividad de producción de energía eléctrica en régimen especial. vol. 126. Real Decreto. doi:10.3892/mmr.2011.655; 2007.

[43] Real Decreto 1578/2008, de 26 de septiembre, de retribución de la actividad de producción de energía eléctrica mediante tecnología solar fotovoltaica para instalaciones posteriores a la fecha límite de mantenimiento de la retribución del Real Decreto 661/; 2008.

[44] Real Decreto-ley 1/2013, de 25 de enero, por el que se prorroga el programa de recualificación profesional de las personas que agoten su protección por desempleo y se adoptan otras medidas urgentes para el empleo y la protección social de las personas des. Real Decreto Ley; 2013.

[45] Ministerio de Industria Energía y Turismo de España. Orden IET/1459/2014, de 1 de agosto, por la que se aprueban los parámetros retributivos y se establece el mecanismo de asignación del régimen retributivo específico para nuevas instalaciones eólicas y fotovoltaicas en los sistemas eléctricos de los territ. Order; 2014.

[46] Real Decreto 900/2015, de 9 de octubre, por el que se regulan las condiciones administrativas, técnicas y económicas de las modalidades de suministro de energía eléctrica con autoconsumo y de producción con autoconsumo. Boe:27548–27562. doi:10.1016/j.jns.2003.09.014; 2015.

[47] Späth L. Large-scale photovoltaics? Yes please, but not like this! Insights on different perspectives underlying the trade-off between land use and renewable electricity development. Energy Policy 2018;122:429–37. https://doi.org/10.1016/J.ENPOL.2018.07.029.

[48] Maghami MR, Hizam H, Gomes C, Radzi MA, Rezadad MI, Hajighorbani S. Power loss due to soiling on solar panel: a review. Renew Sustain Energy Rev 2016;59:1307–16. https://doi.org/10.1016/J.RSER.2016.01.044.

[49] Kimber A, Mitchell L, Nogradi S, Wenger H The Effect of Soiling on Large Grid-Connected Photovoltaic Systems in California and the Southwest Region of the United States. In: Proceedings of the 4th World Conference Photovolt. Energy Conference, IEEE; 2006, p. 2391–5. doi:10.1109/WCPEC.2006.279690.

[50] M.K W, Kerrigan SL, Thornton A. Automatic detection of PV system configuration. World Renew. Energy Forum, WREF 2012, Denver, CO: 2012, p. 1933–1937.

[51] U.S. Department of Energy. Renewable energy data book 2018. ⟨https://www.nrel.gov/news/press/2018/data_book_shows_continued_growth_of_renewable_electricity.htm⟩ [Accessed 15 August 2018]; 2016.

[52] SOLARGIS. Solar resource maps and GIS data for 180+ countries | Solargis n.d. ⟨https://solargis.com/maps-and-gis-data/download/⟩ [Accessed 20 August 2018].

132

122

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annex II: Validation of All-Sky Imager Technology and Solar Irradiance Forecasting at Three Locations: NREL, San Antonio, Texas, and the Canary Islands, Spain

123

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

124

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

*applied
sciences*

MDPI

Article

Validation of All-Sky Imager Technology and Solar Irradiance Forecasting at Three Locations: NREL, San Antonio, Texas, and the Canary Islands, Spain

Walter Richardson Jr. [1,*], David Cañadillas [2], Ariana Moncada [1], Ricardo Guerrero-Lemus [2], Les Shephard [3], Rolando Vega-Avila [4] and Hariharan Krishnaswami [5]

[1] Department of Mathematics, University of Texas at San Antonio, San Antonio, TX 78219, USA; arianam.moncada@gmail.com

[2] Departamento de Física, Universidad de La Laguna, S/N 38206 S/C de Tenerife, Spain; dcanadil@ull.edu.es (D.C.); rglemus@ull.edu.es (R.G.-L.)

[3] Department of Civil Engineering, University of Texas at San Antonio, San Antonio, TX 78219, USA; Les.Shephard@utsa.edu

[4] CPS Energy, San Antonio, TX 78219, USA; RVega-Avila@cpsenergy.com

[5] Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78219, USA; harikrismn@gmail.com

* Correspondence: Walter.Richardson@utsa.edu; Tel.: +01-210-458-4760

check for
updates

Featured Application: Validation of all-sky imager technology at 3 locations: NREL, Joint Base San Antonio, and the Canary Islands, Spain.

Abstract: Increasing photovoltaic (PV) generation in the world's power grid necessitates accurate solar irradiance forecasts to ensure grid stability and reliability. The University of Texas at San Antonio (UTSA) SkyImager was designed as a low cost, edge computing, all-sky imager that provides intra-hour irradiance forecasts. The SkyImager utilizes a single board computer and high-resolution camera with a fisheye lens housed in an all-weather enclosure. General Purpose IO pins allow external sensors to be connected, a unique aspect is the use of only open source software. Code for the SkyImager is written in Python and calls libraries such as OpenCV, Scikit-Learn, SQLite, and Mosquito. The SkyImager was first deployed in 2015 at the National Renewable Energy Laboratory (NREL) as part of the DOE INTEGRATE project. This effort aggregated renewable resources and loads into microgrids which were then controlled by an Energy Management System using the OpenFMB Reference Architecture. In 2016 a second SkyImager was installed at the CPS Energy microgrid at Joint Base San Antonio. As part of a collaborative effort between CPS Energy, UT San Antonio, ENDESA, and Universidad de La Laguna, two SkyImagers have also been deployed in the Canary Islands that utilize stereoscopic images to determine cloud heights. Deployments at three geographically diverse locations not only provided large amounts of image data, but also operational experience under very different climatic conditions. This resulted in improvements/additions to the original design: weatherproofing techniques, environmental sensors, maintenance schedules, optimal deployment locations, OpenFMB protocols, and offloading data to the cloud. Originally, optical flow followed by ray-tracing was used to predict cumulus cloud shadows. The latter problem is ill-posed and was replaced by a machine learning strategy with impressive results. R^2 values for the multi-layer perceptron of 0.95 for 5 moderately cloudy days and 1.00 for 5 clear sky days validate using images to determine irradiance. The SkyImager in a distributed environment with cloud-computing will be an integral part of the command and control for today's SmartGrid and Internet of Things.

125

1. Introduction

The Department of Energy (DOE) estimates that PV power will grow to 14% of the electricity supply by 2030 as the price of solar electricity reaches a point at which it is cost-competitive with cogen ($0.06/kwh by 2020). It is imperative that power grid reliability and stability be maintained under this high penetration of low carbon energy [1]. Organizations such as North American Electric Reliability Corporation (NERC) and California Energy Commission (CEC) [2] have formulated several requirements needed in a "grid-friendly" PV power plant. For instance, CEC has developed a set of several smart inverter functionalities such as dynamic volt/var operation and ramp rate control. Existing PV plants do not have these functionalities even though the inverters are capable, due to the lack of communications standards and dynamic control. For PV power plants to participate in energy markets and ancillary services markets, they need to be considered "dispatchable" power plants.

High penetration of PV systems can be achieved if PV inverters [3] participate in the grid frequency regulation by active power control. Currently, frequency disturbances in the grid are handled by load curtailment. The disadvantage of this methodology is that it can cause voltage stress on the distributed generation. The alternative is to operate the PV system below its MPP (Maximum Power Point) to provide active power control. This can be done by modifying the MPP algorithm in such a way that it can track the next MPP point while working in the reduced power mode (RPM). A critical component of coordinated inverter control is forecasted solar power output or forecasted MPP at the array level. Having accurate intra-hour solar forecasts can enable implementation of a coordinated inverter control strategy capable of regulating a set-point, which may be a signal from a utility requiring either power curtailment or frequency regulation. The electric utility industry has yet to see an integrated solution to the dispatchability problem of PV plants, a system solution that effectively integrates intra-hour solar forecasting and smart control of inverters to achieve not only a grid-friendly plant, but also one that provides monetary and efficiency benefits to the PV plant operator.

Microgrids lack the stabilizing effects present in a large urban macrogrid that itself is joined to an interconnect; these issues are critical when a microgrid is operated in islanded mode. The Energy Management Systems (EMS) that balance PV output, load, and battery storage require accurate intra-hour irradiance forecasts to solve the control problem by shedding non-critical load when power generated is predicted to drop significantly below the load. An increasingly important problem for utilities is optimal scheduling and dispatch of a microgrid [4–7], both when connected to the macrogrid and when operated as an island. This task is divided into Day-Ahead Scheduling, which finds optimal schedules for the next operating day and focuses on energy markets, and Intra-day Dispatching and Scheduling in which schedules are continuously updated during the current day. Both cases follow these steps: (1) forecast day-ahead load, (2) forecast day-ahead renewable power (solar, wind), (3) Micro-Grid Management System (MGMS) optimizes the day-ahead plan, produces schedules for flows within the microgrid and to the PCC, (4) MGMS transmits schedules to utility control center.

This article describes a four-year research effort to develop hardware and software with an aim to solve the intra-hour solar forecasting problem for electric utilities. It was a collaborative effort between many groups, including national labs and research institutes (NREL and the Texas Sustainable Energy Research Institute TSERI), two universities (The University of Texas at San Antonio and Universidad de La Laguna in the Canary Islands), public and private utilities (CPS Energy, ENDESA, and Duke Energy), and a private company, Siemens-Omnetric. It serves as a case study in managing research in a joint theatre of operations and integrating the efforts of researchers and engineers who came from very different university and industrial cultures. Details of our research and technology development have been presented in journal articles [8,9], conference proceedings [10–13], and technical reports [14].

While this paper gives a detailed overview of that research, our primary goal is to describe how the UTSA SkyImager was validated at three geographically diverse locations, the pitfalls encountered, the lessons learned, and the outlook for future research efforts.

The SkyImager evolved from a realization that existing all-sky imaging systems were too expensive to be deployed in large numbers, suffered from data-loss issues caused by the shadow band and camera arm, and used proprietary software. The Raspberry Pi single board computer ($35) and programable high-resolution Pi-Cam ($20) with a fisheye lens ($20) formed the heart of the new system. The most expensive component was the all-weather security camera enclosure ($350). In addition, General Purpose IO (GPIO) pins would allow a variety of external sensors to be connected to the Pi. Low cost and ease of use were essential if the SkyImager were to be deployed in a rural sustainable development microgrid. In contrast to some commercial all-sky imaging systems, only open source software would be used. The Pi accommodates several operating systems (OS) including Raspbian, a Linux-based derivative of Debian. It can be operated with a monitor or in "headless" mode, and once deployed can be accessed remotely with SSH/SFTP. Code for the SkyImager would be written in Python and allow calls to libraries such as Open Computer Vision, Scikit-Learn, SQLite, and Mosquito. In the summer of 2014 it was an open question whether the proposed imager could deliver the functionality of much more expensive systems and be thoroughly tested before its deployment.

As part of the DOE microgrid INTEGRATE program, the first deployment of the SkyImager occurred in Fall 2015 at the ESIF building at NREL. INTEGRATE aggregated sustainable generation and loads into microgrids controlled by an Energy Management System with the OpenFMB protocol. In 2016 a second SkyImager was installed at the CPS Energy microgrid at Joint Base San Antonio. A multi-year collaboration between CPS-UTSA and the Universidad de La Laguna resulted in the deployment of two SkyImagers in the Canary Islands. These utilize stereoscopic images to determine heights of the bases of cumulus clouds. Deployment of SkyImagers in three diverse locations provided not only big data, but operational experience in harsh extremes of climate. This resulted in improvements and additions to our original design: weatherproofing, new environmental sensors, the need for scheduled maintenance, optimal positioning of the camera, communications with OpenFMB publish-subscribe protocols, and using WiFi and cloud computing. The SkyImager will be an integral part of the command and control for microgrids, both as part of the larger SmartGrid in an urban environment or in an islanded mode in a military or rural setting.

Solar forecasting is widely considered a key means of integrating solar power efficiently and reliably into the electric grid. For a utility to meet projected customer demand with electricity from sustainable resources, high-accuracy global horizontal irradiance (GHI) forecasts must be available over widely different time and space scales. A convenient separation of this forecasting problem into two parts is as follows: (1) *intra-hour forecasts* of the ramp events that are caused when cumulus clouds move between the sun and the solar panels, and (2) *day-ahead forecasts* for 12, 24, and perhaps 36 h into the future. There is overlap between the two parts, but this taxonomy is convenient not only because the physics and forecasting techniques are generally different, but also the way in which the utility makes use of the forecasts. A single ramp event on a microgrid powered primarily by PV arrays can result in over/under voltages, as well as frequency deviations and may require secondary spinning reserves to be brought on line. Day-ahead irradiance forecasts are useful in predicting surplus/deficit generation capacity that can then be augmented or sold in the day-ahead electricity market.

1.1. Day-Ahead GHI Forecasting

For day-ahead GHI forecasting, both numerical weather prediction (NWP) [15] and satellite imagery provide effective tools for forecasting irradiance. The National Center for Environmental Prediction (NCEP), a part of NOAA, runs two versions of the Rapid Refresh (RAP) numerical weather model to predict environmental data. The first version generates weather data on a 13-km (8-mile) resolution horizontal grid and the second, the High-Resolution Rapid Refresh (HRRR), generates data on a 3-km (2-mile) grid. RAP forecasts use multiple data sources: commercial aircraft weather data,

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

balloon data, radar data, surface observations, and satellite data to generate forecasts with hourly resolution in time and forecast lengths of 18 hours. For further details, consult the RAP website [16]. RAP data are available for download through the National Model Archive and Distribution System (NOMADS). Several benefits accrue from using NWP for irradiance forecasting: NOAA incurs much of the computational burden and these models incorporate first-principles physics such as the Navier-Stokes equations, thereby allowing for the dynamic formation of clouds. Satellite technology is advancing rapidly with GOES-16 (Geostationary Operational Environmental Satellite) pictures being updated every 5 min with maximum resolution of 5000 by 3000 pixels. Figure 1 shows such an image cropped to the central Texas region. The temporal sampling of the data is still insufficient to support optical flow predictions 15-min ahead. Continued improvements in GOES-R (geostationary satellite with high spatial and temporal resolution) may well make the satellite approach to minutes-ahead irradiance prediction more attractive in the future [17–20]. Statistical methods based on historical time-series data and climatology are also useful for day-ahead PV forecasting.
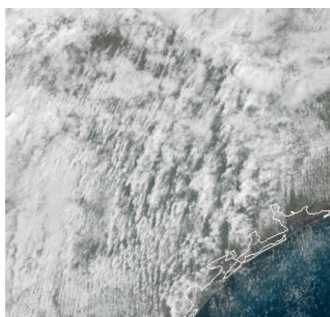

Figure 1. GOES high-resolution satellite image cropped to show the central Texas region.

1.2. Intra-Hour Solar Forecasting

The energy alliance between UTSA and CPS Energy has as one of its primary goals the development of new solar forecasting technologies that combine inexpensive all-sky imaging cameras with sophisticated image processing techniques and artificial intelligence software to produce high-accuracy 15-min ahead solar irradiance forecasts. GHI consists of two components, the Direct Normal Irradiance (DNI) caused by sunlight traveling in a direct path from sun to PV array and the Diffuse Horizontal Irradiance (DHI), background illumination that is due to secondary reflections and absorption/re-radiation. The formula is $GHI = \cos(\theta_z) \, DNI + DHI$ where θ_z is the zenith angle. Shadows cast by low-level cumulus clouds significantly impact the DNI but have little effect upon the background DHI. While it is possible to predict DNI separately [21], for verifying PV power output forecasts GHI is used. Figure 2 displays the three quantities: GHI, DNI, and DHI, on the date 27 October 2015 at the NREL ESIF facility in Golden, CO. It shows that moderately cloudy conditions occasion multiple ramp events.

128

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
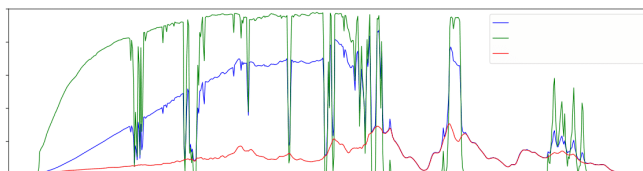Fecha: 02/06/2021 14:57:08

Figure 2. Daily solar irradiance (W/m^2) for 27 October 2015, Golden, CO.

Figure 3 displays a sequence of eight pictures taken by the SkyImager at the NREL site, one every two minutes starting (upper left) at 12:31pm MST. At that time the sun is not obscured, but cumulus clouds are moving in from the left. At 12:35 the cloud begins to enter the solar disk and by 12:37 the sun is completely occluded. This continues until 12:44 when the cloud has moved past the sun and the DNI recovers. This ramp event is seen in the DNI oscillations that occur around the noon hour in Figure 2. While this example considers a single ramp event, it strongly suggests that the correlation between measured GHI and the presence of clouds obscuring the sun in the SkyImager pictures could be learned by AI models.
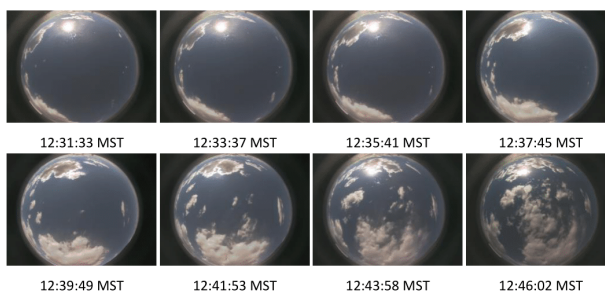

Figure 3. SkyImager sequence (27 October 2015, Golden, CO) showing a ramp event.

The challenge in short term prediction of PV power is simply "Where will cumulus cloud shadows be fifteen minutes from now?" Our approach incorporates as much of the physics as possible, but is an idealization necessitated by the requirement to produce GHI forecasts in real time for the MGMS. Solving Navier-Stokes for the true dynamics of the atmosphere is not feasible on a Raspberry Pi. If GHI can be accurately forecast, then predicting PV power output is straightforward. The evolution of clouds and irradiance shown in Figures 2 and 3 is even more striking when video of the images is viewed, confirming that SkyImager pictures are highly correlated with the observed GHI time series. Moreover, it suggests the camera sensor could be used to measure irradiance as well as predict it. Once GHI has been accurately forecast, it is usually straightforward to assign a corresponding PV power output, which is what the MGMS requires.

Figure 4 shows the relationship between GHI ($Watts/m^2$) and PV power (Watts) from the RSF2 PV arrays located at NREL. The relationship is almost linear with a slight hysteresis effect that reflects the differences in morning versus afternoon irradiance. The task of predicting PV power output over multiple temporal and spatial scales, and from a variety of different equipment is a challenging one [22–24].

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
             UNIVERSIDAD DE LA LAGUNA

             María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
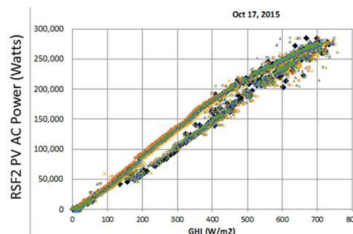   *UNIVERSIDAD DE LA LAGUNA*

Figure 4. Relationship between GHI (W/m^2) and PV Power (Watts) determined at NREL.

State of the Art in Solar Forecasting

As photovoltaics achieve greater penetration, the SmartGrid will demand accurate solar forecasting hence a network of low cost, distributed sensors to acquire large amounts of image and weather data for input to the forecasting algorithms. Commercial sky imaging systems often prove too costly and have proprietary software, leading several research groups to develop their own systems. The solar forecasting research group at UC San Diego [25–29] has done pioneering work in this field for many years. For one example, Coimbra et al. [30] proposed DNI forecasting models using images from a Yankee TSI 880 with hemispherical mirror as inputs to Artificial Neural Networks (ANN). The TSI has high capital and maintenance costs, uses a shadow band mechanism, and requires proprietary software. The UCSD Sky Imager described in Yang et al. [31] captures images with an upward-facing charge-coupled device (CCD) Panasonic sensor and a 4.5 mm focal length fisheye lens. Compared with the TSI it has higher resolution, greater dynamic range, and lossless PNG compression. The Universitat Erlangen-Nurnberg [32] group used a five-megapixel C-mount camera equipped with a fisheye lens. They implemented the Thirions "'daemons'" algorithm for image registration and cloud-motion estimation similar to optical flow. In Australia, West et al. [33] used off-the-shelf IP security cameras (Mobotix Q24, Vivotek FE8172V) for all-sky imaging. Inexpensive compared to the TSI, the cost of such systems still is ~800 euros. Rather than a feature-tracking strategy, they used dense optical flow to estimate cloud movement. See also Wood-Bradley [34]. Several research groups in China are working on the irradiance forecasting problem [35,36] generally using a TSI imager, but in one case Geostationary Statellite data [37]. As mentioned before dramatic improvements in GOES-R technology and resolution (spatial and temporal) will make this approach more attractive for intra-hour forecasting. See also the historical review [38] of irradiance and PV power forecasting that was produced using text mining and machine learning.

A recurring theme in the INTEGRATE project was that while all-sky imaging was a critical component of microgrid stability and control, it could not be developed in a stand-alone fashion but must be fully integrated into the microgrid management system (MGMS). Uriate et al. [39] discuss the importance of Ramp Rates (RR) on the inertial stability margin of a microgrid deployed at the Marine Corp Base at Twentynine Palms. They show that a large ramp in PV power can destabilize frequency when the PV load is suddenly transferred to the cogen. The inertia constant *H* of a generator is the ratio of stored kinetic energy to system capacity. Microgrids usually have $H < 1$ s compared to 2–10 s for large generation plants. Frequency stability is defined by the condition $|\Delta f_{pu}| < \Delta f_{max}^{pu}$ where allowable frequency deviation Δf_{max}^{pu} in p.u. is typically 0.01–0.05 per unit. In [39] the authors derive the ODE $\omega_{mech} J d\omega_{mech}/dt + \omega_{mech}^2 D = P_{accel}$ where ω_{mech} is the mechanical speed of the generator in rad/s, J the moment of inertia (kg·m^2), D a damping coefficient (Nm/s), and P$_{accel}$ the power imbalance exerted on a generator rotor, to model the microgrid stability control problem. The NREL microgrid has a 300 kW Caterpillar diesel generator, whereas JBSA has no cogen. However, the same issues of stability and frequency control apply when there are no spinning resources. Some electric

130

codes are specifying ancillary control must be added to the EMS in order to handle ramp events of a certain magnitude and duration. See [40–42] for details.

1.3. Climatology and Microgrid Architectures at the Three Locations

As shown in Figure 5 the UTSA SkyImager has been deployed at 3 geographically diverse locations: Golden, Colorado on the rooftop of the ESIF building at NREL, in San Antonio, Texas at the CPS Energy microgrid facility at Joint Base San Antonio (JBSA) and the Engineering Building at UTSA, as well as in the Canary Islands, Spain at Tenerife and Caleta de Sebo. Each location presented unique challenges in terms of local climate, physical and cyber access, and microgrid design, equipment, operation, and customer needs.
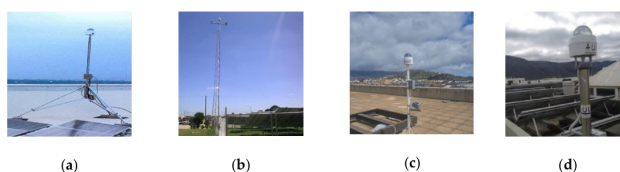

(**a**)	(**b**)	(**c**)	(**d**)

Figure 5. SkyImager at (**a**) NREL, (**b**) JBSA, (**c**) Tenerife and (**d**) Caleta de Sebo.

The UTSA SkyImager was first conceived as a technology for providing accurate intra-hour irradiance forecasts as inputs to a microgrid management system that would then provide the utility with command and control of the microgrid in either connected or islanded mode. The Department of Energy INTEGRATE project [43] lasted for 18 months beginning on 6 March 2015 and partnered NREL, Omnetric-Siemens, CPS Energy, Duke Energy, and UTSA. The project goal was to increase the capacity of the electric grid to incorporate renewables by upgrading and optimizing architectures for control and communication in microgrids. There were three major components: (1) OpenFMB, a reference architecture that allows real time interaction among distributed intelligent nodes, (2) optimization with the Spectrum Power Microgrid Management System based on the Siemens SP7 Platform, and (3) PV and Load Forecasting using UTSA's applications for both intra-hour and day-ahead irradiance and building load forecasts. The OpenFMB framework leverages existing standards such as IEC's Common Information Model (CIM) semantic data model and the Internet of Things (IoT) publish/subscribe protocols (DDS, MQTT, and AMQP) to allow flexible integration of renewable energy and storage into the existing electric grid. The OpenFMB standard was ratified by the North American Energy Standards Board (NAESB) in March of 2016 and allows communication between diverse grid devices–meters, relays, inverters, capacitor bank controllers, etc. It allows federated message exchanges with readings such as kW, kVAR, V, I, frequency, phase, and State of Charge (SOC) published every 2 seconds as well as data-driven events, alarms, and control in near-real-time.

1.3.1. SkyImager at National Renewable Energy Laboratory in Golden, CO

The site of the first SkyImager deployment was NREL in the Rocky Mountains. Golden's high elevation and mid-latitude interior continent geography results in a cool, dry climate. There are large seasonal and diurnal swings in temperature. At night, temperatures drop quickly and freezing temperatures are possible in some mountain locations year-round. The thin atmosphere allows for greater penetration of solar radiation. As a result of Colorado's distance from major sources of moisture (Pacific Ocean, Gulf of Mexico), precipitation is generally light in lower elevations.

Eastward-moving storms from the Pacific lose much of their moisture falling as rain or snow on the mountaintops. Eastern slopes receive relatively little rainfall, particularly in mid-winter. The SkyImager enclosure came equipped with a heater/fan that performed well at NREL. Given the

climate, it proved useful in keeping frost off the plastic dome. It adds to the expense and complexity of the technology and may not be required at other locations. Most installations of the security camera enclosure would be facing downward and perhaps under a building overhang. Used facing upward and exposed to the sky, there were issues with water getting inside the enclosure. A simple solution was silicon caulk applied at the base of the dome. In a typical security installation, a green tinted plastic dome is used with the enclosure to protect components from UV radiation. For all-sky imaging a clear plastic dome is a necessity. With any plastic material on a bright sunny day there can be issues with glare caused by the dome, but this was minor. The alternative is a glass dome but that has it own set of problems.

As shown in Figure 6, the microgrid at NREL was already well established and the process of deploying the SkyImager went relatively smoothly. Denver International Airport is located some 36 miles from Golden; this distance introduces some error in the Cloud Base Height for the ray-tracing algorithm originally used in the SkyImager. The ESIF building at NREL had the infrastructure necessary for easy installation of both the SkyImager and the Hardkernel Odroid C1 single board computer (SBC) used for load and day-ahead PV forecasts. Information was transferred using a Wi-Fi network on a LAN system. NREL also provided un-interruptible 120 VAC power, ample Ethernet connectivity, and excellent on-site weather and irradiance data. In addition to solar PV arrays, generation included a 500 kW wind power simulator and a 300 kW caterpillar diesel. A 300 kWh battery system provided energy storage and the load was separated into a controllable component (250 kW) and a critical load (250 kW).
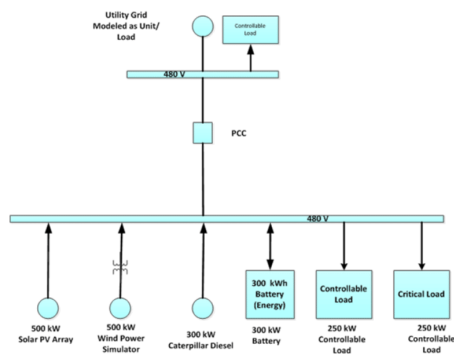

Figure 6. Microgrid at NREL ESIF Building where the first SkyImager was deployed in 2015.

1.3.2. SkyImager at San Antonio, TX, USA

Texas produces more electricity than any of the other 49 states, and as a result has its own interconnect ERCOT. In 2017, power statewide was generated by a variety of sources: natural gas (45%), coal (30%), wind (15%), and nuclear (9%). In 2014, wind replaced nuclear as the third-largest source of power and Texas now produces more wind power than any other state. Solar generation is increasing, but still relatively small for a state with abundant annual sunshine. Located in central Texas some 200 miles from the Gulf of Mexico, San Antonio is home to almost 1.5 million people and several military bases. CPS Energy serves San Antonio and is the nation's largest public power, natural gas and electric company. They are committed to renewables, funding a 400 MWac project with multiple PV plants (Alamo 1–7) close to San Antonio, and wind farms in West and South Texas. CPS Energy is among the top public power wind energy buyers in the nation and number one in Texas for solar

generation. In keeping with this commitment, TSERI was formed in 2001 as an alliance between CPS Energy and UTSA.

For San Antonio, the most significant local weather issue is low-level Gulf stratus [44]. Elevations of the terrain increase from sea level at the Gulf coast to almost 800 ft at San Antonio, and a moist air mass over the Gulf of Mexico will cool adiabatically to saturation as it moves upslope. Nocturnal radiational cooling causes cloud formation before midnight, resulting in a ceiling of 500–1000 feet. A solid cloud deck will cover much of central Texas and remain in place until late morning when the sun burns off the stratus and cumulus clouds begin to form. Forecasting Gulf stratus is an important problem for aviation; it is a matter of accurately predicting low-level wind flow (<5000 ft) with the most favored wind direction for stratus formation from 90° to 180°. It is important to address these local weather conditions that occur below the spatial and temporal resolution of NWP, but are crucial for both inter-hour and day-ahead irradiance forecasts. Use of machine learning using local datasets and climatology will allow the information and intelligence of a study such as [44] to be incorporated in site-specific irradiance forecasts.

The Fort Sam Houston Library location at JBSA presented several unique challenges for the deployment of the UTSA hardware and software, challenges that provide valuable insights for other researchers. Many of the issues that arose were heavily dependent on the specific location. At JBSA, the Sky Imager was deployed using an edge-computing configuration with a wired Ethernet connection for cyber security. The JBSA microgrid is shown in Figure 7 and includes the Base Library building, solar arrays, inverters, and the pod housing the battery energy storage system (ESS). The need for accurate on-site meteorological observations necessitated installation of a complete MET Station atop a 10m antenna tower. A Campbell Scientific weatherproof instrument box at the tower base contained a National Instruments MyRio computer, a transformer, backup battery, and an Odroid C2 single board computer (SBC) for calculating the day-ahead load/PV forecasts. Atop the tower sat the SkyImager, a WXT520 Vaisala weather transmitter, and a pyranometer.
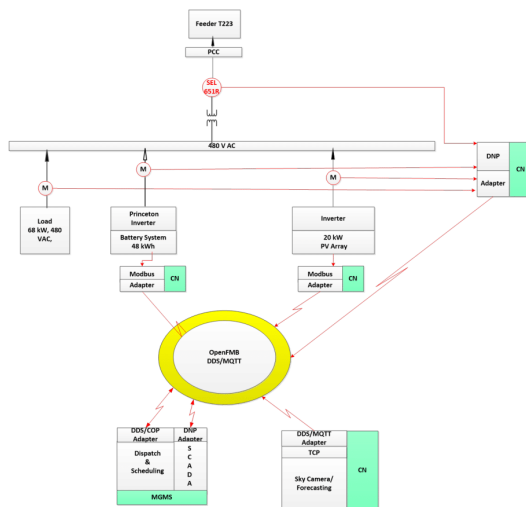

Figure 7. The Microgrid at Joint Base San Antonio, TX, USA.

In July of 2018 another SkyImager was deployed in San Antonio at the location of a university PV generation project. Funded by a DOE-SECO grant [45] in 2014, solar panels were installed on the Engineering Building, HEB University Center III, and Durango buildings at UTSA. In addition, equipment was installed to record measurements from 4 Combiners, 4 Inverters, 2 Kipp & Zonen CMP11 pyranometers, and a WXT520 Vaisala Weather Transmitter at the UCIII. Figure 8a displays the SkyImager and PV panels and Figure 8b shows combiners/inverters atop the Engineering Building. The only ingredient lacking to make this a research microgrid was energy storage.

(**a**)   (**b**)

Figure 8. (**a**) PV panels and SkyImager at UTSA; (**b**) Inverters & combiners, Engineering Bldg.

1.3.3. SkyImager in the Canary Islands

Six insular power grids comprise the electrical network in the Canary Islands. Conventional generation costs more here than PV technologies, and savings can be shared between the PV system owners and the Spanish utility ENDESA. Penetration of renewables varies among these grids, from a high of 60% penetration of wind energy in El Hierro (after Gorona del Viento hydro-wind power plant is operational) to Lanzarote-Fuerteventura which achieves single-digit integration of renewables because of strong environmental regulations, a weak power grid, and an unstable regulatory environment in Spain for renewable energy infrastructure during the period 2011-15. However, new regulatory policies will provide a more attractive framework for investment. The Canary Islands Government plans to avoid ground-based renewable facilities with a large environmental footprint in favor of smaller rooftop plants close to electricity users. As in Hawaii, the existing distribution grid is not prepared for a large penetration of residential PV systems with resulting reverse flows, voltage and frequency instabilities, and drops at the end of long lines. ENDESA has built a testbed smart grid in a village at the north end of the Fuerteventura-Lanzarote insular power system (La Graciosa).

La Graciosa is the smallest island in the Canary Archipelago with a surface area of 29 km^2. It is in a marine nature reserve north of Lanzarote and home to about 700 people in the island capital of Caleta de Sebo. Average global irradiation is 5.157 kWh/ kW·day (1883 kWh/kW·yr) while the average monthly high temperature is 20.8 °C. Located a few kilometers away from the African coast, its proximity to the Sahara Desert gives to La Graciosa particularly stable atmospheric characteristics due to a quasi-permanent subsidence thermal inversion. Constant north trade winds, along with the high content of aerosols and dust in the atmosphere, have a large influence over the cloud dynamics and therefore, the irradiance in the region. As shown in Figure 9, the La Graciosa grid is supplied by three 20/0.4 kV transformers (600, 400, and 400 kVA) and tied by a 20 kV seabed cable to Lanzarote. The island has two PV generation plants (5 kW and 30 kW), but recently La Graciosa PV capacity was increased, enhancing the attractiveness of a smart grid energy management system.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
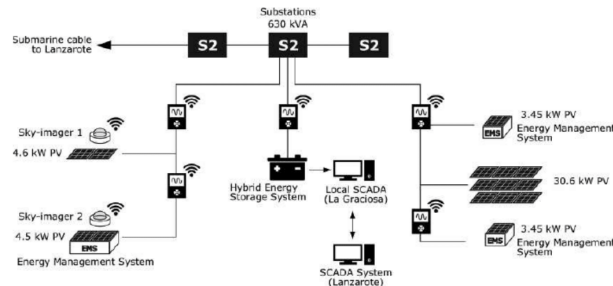Fecha: 02/06/2021 14:57:08

Figure 9. Microgrid in the Canary Islands.

One of the main differences between La Graciosa project and the prior two experiences in the USA, is that in the island, a system composed of two sky-imagers was installed, as can be seen in Figure 9. The reason behind this was to give the forecasting system the ability to estimate cloud base height (CBH) making use of stereoscopic techniques as in [46,47]. This provides the system with added value in terms of functionality and gives extra data to incorporate in the next steps of the image processing and forecasting pipeline. A recent paper comparing the use of different instruments to measure the CBH concluded that using a pair of inexpensive cameras was the most cost-effective alternative in comparison with other methods such as a ceilometer or LIDAR [48]. In fact, cloud base height is quite important to estimate the position of the shadows if a ray tracing approach is taken and it can also be included as a feature if Machine Learning methods are preferred, as it correlates with the position of the clouds in the image and the recorded irradiance or PV production.

In this case, the device falls away from the Internet of Things (IoT) concept, since two cameras are involved in the system, and some computation must be done either in one of the devices or (as it was done in the project) on a dedicated server. Of course, there are some advantages and drawbacks for using either method, but we found particularly easy the connection between the sky-imagers and the server, and we could exploit the higher computational capabilities of the dedicated server. The main requirement to work this way is to have a robust internet access, which fortunately was granted by the owners of the buildings where the sky-imagers were installed. The network speed can also influence the way of operating, as it can act as a bottleneck in the data stream (due to the relatively large size of images compared to other types of files). Also, in the future the use of Machine Learning algorithms could be done on the server, which is expected to perform better than computing directly on the device.

2. Materials and Methods

In the original configuration of the SkyImager, a security camera enclosure housed a Raspberry Pi single board computer with programmable Pi camera. The enclosure contained a small circuit board with heater and fan that runs off a supplied 24V AC power supply, standard with many security cameras. The 12V AC output from this board is input to an AC-DC converter which supplies 12V DC to a TOBSON converter which supplies 5VDC at 3A for the Raspberry Pi.

2.1. SkyImager Hardware

Figure 10a displays the original SkyImager hardware. At NREL it was found necessary to add an extra SBC for increased computational power–the Odroid C1 by Hardkernel. Heat dissipation is an issue with SBC in Texas summers. One C1 was destroyed by heat and as result a cooling fan was added to the design. The new C2 Odroid has a heat sink to eliminate the overheating issue.

135

Acquiring and fusing the 3-exposure images could be done with just the Raspberry Pi 2. The Pi 3 model is 50% faster than its predecessor; careful optimization of the workflow will allow acquisition, processing, and forecasting with just a Pi 3. This would reduce cost and greatly simplify network connections. Figure 10b shows this configuration with a single Pi 3, plastic case, cooling fan, camera, and WeatherBoard. Images could be pushed to the cloud for processing, however, the necessary bandwidth would be substantial. The sky imagers in La Graciosa are built upon a Raspberri Pi 3 model B with no ancillary boards, and a super wide fish-eye lens (field of view over 180°). An inexpensive mini PV module was added to record irradiance at the camera locations there.
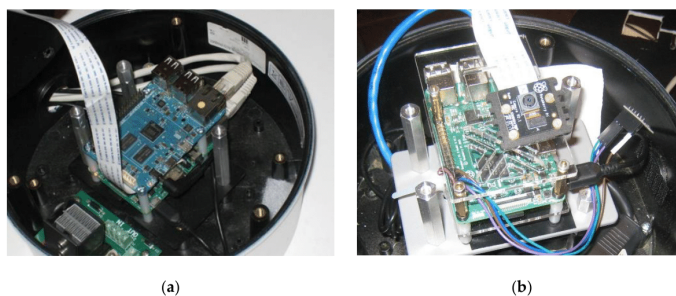

(**a**)   (**b**)

Figure 10. (**a**) Original configuration Raspberry Pi 2/Odroid C1, (**b**) Current hardware with a single Pi 3 Model B.

Figure 11 shows the equipment deployed on the MET tower at Ft. Sam: a Kipp&Zonen CMP11 pyranometer, the UTSA SkyImager, and a Vaisala WXT520 weather transmitter. Each of the commercial devices costs several thousand dollars; this cost prompted us to search for low-cost substitutes.

Figure 11. CMP11 pyranometer, UTSA SkyImager, Vaisala Weather Station.

Several inexpensive alternatives to a commercial pyranometer exist [49]. Devices can be added to the GPIO pins on the Raspberry Pi. The Hardkernel Weather-Board 2 shown in Figure 12a can take not only temperature, humidity, and pressure readings (bme280 Application-Specific Integrated Circuit ASIC), but also measures light in the Visible, Infra-Red, and Ultra-Violet bands (si1132 ASIC). There is a Python interface for data retrieval. After calibration and conversion of Lux to W/m^2, this provides irradiance measurements and limited weather data in real time. Another ancillary device that will

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
            UNIVERSIDAD DE LA LAGUNA

            María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
            UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

be useful during initial deployment of the SkyImager is a GPS locator. At $20, it looks like a small mouse for a desktop computer and plugs into a USB port. The Linux programs gpsmon and cgps can be installed on Raspbian and used to take readings of the exact position using the latest GPS satellite data. These are just two of many environmental sensors that can be connected to the Raspberry Pi. Figure 12b shows the new PiNoIR camera, which captures infrared light as well as visible. This would allow for increased contrast between low-level cumulus and high-level cirrus clouds composed of ice crystals. The SkyImager can be used for additional tasks such as air quality monitoring. Figure 12c shows the $25 MQ-131 ozone detection sensor, for example.
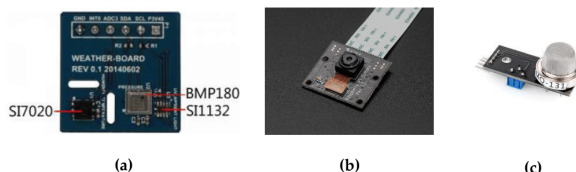

Figure 12. (**a**) Hardkernel's WeatherBoard, (**b**) PiNoIR camera, and (**c**) MQ 131 Ozone Sensor.

2.2. Image Processing Pipeline

Several additional external inputs were required for our forecasting algorithms: distortion parameters for the fish eye lens, zenith angle, True North, and most importantly the cloud base height (CBH). These inputs are used in the image processing pipeline (Figure 13) to output real-time GHI forecasts for the MGMS. A summary of the pipeline is included here, for details see [8]. (1) Distortion Removal (due to fish eye lens), (2) Cropping and Masking, (3) Calculation of "Red-to-Blue Ratio" (RBR), (4) Apply Median Filter to remove impulsive noise, (5) Thresholding to determine cloud presence, (6) Compute Cloud Cover percentage (clear/moderately-cloudy/overcast), (7) Project Clouds to height of CBH, (8) Use Optical Flow to move clouds forward in time, (9) Ray-Tracing to locate cloud shadows, and finally (10) Calculate GHI using shadow locations. Although physically correct, Step (9) Ray-Tracing is an inverse problem mathematically, hence "ill-posed". Small errors in locating shadows can produce significant errors in the forecast irradiance. To address this issue, we investigated using artificial intelligence and neural works to predict GHI values directly from forecast cloud locations.

At NREL another raw image was acquired every 15-seconds. The pipeline described above must be fine-tuned if the processing SBC is to achieve the necessary throughput. An SBC is much more limited than a desktop server as regards CPU speed and available memory/storage, which is provided by a 32 Gb micro-SD card. The usual tradeoffs between keeping a large array in memory versus writing it to disk, are still present even though there is no disk. Efficient programming constructs are required if the goal of low cost is to be achieved. The EMS may run on a military grade RuggedCom server but the SkyImager software is constrained run on an ARM architecture. Steps in the pipeline that have little effect on the overall forecast accuracy can be eliminated. Profiling/timing runs on the optical flow algorithms will show bottlenecks that can be addressed. This is important whether a ray-tracing approach or a machine learning strategy is employed.

The goal in intra-hour solar forecasting is real time PV power predictions. Those forecasts result from a two-step process: predicting cumulus cloud locations 15-min in the future and using projected cloud locations to forecast irradiance. Each step introduces errors. Work is ongoing for *Step 1 - Optical Flow*: compute error metrics of the 15-min ahead image versus the actual image. *Step 2 - Machine Learning* takes the predicted image and computes GHI. This approach separates optical flow [50] from machine learning (ML) and allows *GHI to be predicted directly from the image itself*. The SkyImager can be used as a pyranometer for measuring/observing irradiance.

The training datasets for the neural networks are region-specific, if not site-specific, and require many all-sky images taken on moderately cloudy days. The weights determined for Golden, CO, will have to be fined tuned for deployment in the San Antonio area for example. Training is computationally expensive, whereas the inference or prediction is very fast and can be handled by the Pi. Research on massive deep learning networks requires Graphical Processing Units (GPU) for training and software such as Theano, Keras, or Tensorflow.
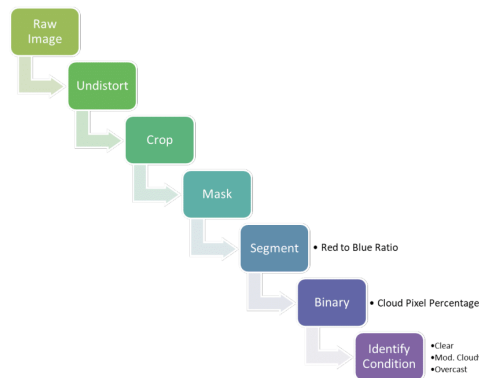
Figure 13. Image Processing Pipeline.

2.3. Machine Learning for Irradiance Forecasting

Machine Learning (ML) is now ubiquitous in all areas of engineering and data science. It has been used in many different ways to help solve the solar power forecasting problem, as described in [51,52]. Another area where ML is widely used is forecasting building load [53,54] which includes methods that are physics-based, statistics-based (Gaussian Process, Linear Regression), and use machine learning (Artificial Neural Network, Support Vector Machine, Deep Learning). Classic references for Deep Learning include [55–57] and for Convolutional Neural Networks, [58].

AI software for data mining has evolved dramatically over the last few years. In data analytics Python is the premier programming language [59] and this fully validated our decision to use it for the SkyImager project. Rapidminer (Version 8.1.001) [60] is a machine learning platform with a point and click interface. As shown in Figure 14, a data flow pipeline is established that permits the user to input a data set, select attributes to analyze, determine target and predictor variable roles, partition the data into training, validation, and sometimes testing subsets, create a logical fork to apply different subprocess models such as Random Forests or Deep Learning, run the model(s), and assess error metrics and overall performance. It is a proprietary package, but a version with somewhat reduced functionality is available for educational use. For some models Rapidminer utilizes the H2O machine learning modules (Version 3.8.2.6) [61,62]. Specifically, our Deep Learning (DL) model is found in H2O as the Python function H2ODeepLearningEstimator(). An open source Python package Scikit-Learn (Version 0.19.0) [63] allows a user to prototype and compare a variety of classification, clustering, and regression models. Neural networks and deep learning has seen the evolution of specialized software such as Keras, Theano, and Google's Tensorflow, which recently became open source. The computational demands of training networks on big data are extreme, and this has resulted in a hardware evolution from central processing units (CPU) to graphical processing units (GPU) to special purpose tensor processing units (TPU).
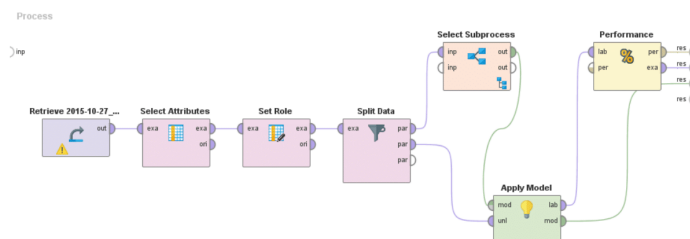
138

Figure 14. Point and click GUI for Rapidminer allows easy model development.

Before predicting irradiance for cloudy days, consider the much simpler problem of forecasting on a clear day. The Haurwitz analytic model performs well on days with no clouds. Using physics, one can derive a closed-form functional relationship [64]: $GHI_{clr} = 1098 [\cos\theta_z \exp(-0.057/\cos\theta_z)]$, where θ_z is the solar zenith angle. Can neural networks learn this relationship, given a large enough dataset on which to train? The actual forecasting problem on a cloudy day is of course much harder. Information in each all-sky image is used to locate and track low-level cumulus clouds as they move between the sun and PV-arrays. It is the difference between using ML to recognize machine-written (or even hand-written) digits versus recognizing and identifying faces in a crowd of people moving down the street. Work is ongoing to identify the best features to extract from the images, to efficiently solve the intra-hour solar forecasting problem and to predict very short-term ramp events.

A critical component of any machine learning strategy is deciding which features or input variables are most strongly correlated with the labels or output variables. A second aspect involves finding a representation of the data that is compressed or sparse in some basis. This dimensionality reduction [65] can be achieved through principal component analysis (PCA) or by the simple process of discarding unimportant features in the inputs. In our studies, the label or target variables were scalars: GHI values measured atop the ESIF Building at NREL. Other choices are possible such as the value $GHI_{clr} - GHI_{mea}$, the deviation from the clear sky value.

Each input or example is a 3-channel RGB image from the Pi camera. In the current configuration, 3 images taken 5 seconds apart at low, medium, and high exposure times are fused using the Mertens algorithm into one raw image. As mentioned, this approach reduces over-exposure and washout in the circumsolar region. The 1024×768 JPEG forms the basic input measurement for both the optical flow and machine learning algorithms. Low level cumulus clouds between the sun and the PV arrays have the greatest effect on the DNI, hence on GHI. For that reason, and to satisfy the need for dimensionality reduction, the first preprocessing step is to locate the area in the image that surrounds the sun and extract a 128×128 subimage.

Several approaches to locating the sun in the image are possible. Calculating the zenith angle from the SOLPOS program, finding true North, and then mapping physical to pixel coordinates would require extensive calibration. It was decided to use a simple robust image processing approach that finds the maximum intensity in the image. On a clear day, this always locates the sun, but occasionally when the sun is totally obscured by broken clouds, the brightest point in the picture is actually sunlight reflected off of a nearby cloud. This can be observed in a time lapse video clip in which for a few frames the sun is not at the center of the sub-image. While the cause of some transient errors, it never lasts long and does not happen when the sun is totally obscured by a uniform cloud deck without breaks. Figure 15 shows the low exposure image used to locate and center the sun and the resulting raw fused image that will ultimately become the input to the neural networks. Lastly the subimage will be resized to a point (8×8) where the neural networks will train in a reasonable amount of time. In supervised learning, the neural networks require labeled training examples: ordered pairs (x, y)

139

where x is the input vector, in this case the extracted subimage *img*, and y is the measured irradiance in units of Watts/m^2 at the time the picture was taken. Since new images are fused every 15 seconds, GHI values were treated as constant on a 60 second interval for the purposes of assigning labels for the neural network images.

Figure 15. Sub-images of circumsolar region: (**a**) low exposure and (**b**) raw fused.

Many different metrics are used in data analysis and ML for evaluating model performance. Table 1 lists common ones for solar forecasting. A_t is the actual value, F_t the forecast value, and μ the mean; the summation over t can be over all observations in the ML context or over the values of a time series. The metrics provide a posteriori error bounds upon which utilities can make economic decisions. MAPE is relative L$_1$ error, normalized for number of observations and converted to percent. When the denominator of a fraction is close to zero, μ is used instead of A_t, a common practice when predicting spot electricity prices. Some metrics such as L$_1$ are more robust–less sensitive to outliers–than the classical RMS error norms.

Table 1. Regression Metrics: A actual, F forecast, μ mean.

Metric	Definition
Mean Squared Error	$\text{MSE} = (1/n)\sum_t (A_t - F_t)^2$
Normalized RMS Error	$nRMSE = \sqrt{\overline{MSE}}/\rho, \rho = (A_{max} - A_{min})$
Explained Variance	$R^2 = 1 - \sum_t (A_t - F_t)^2 / \sum_t (A_t - \mu)^2$
Mean Absolute Error	$MAE = (1/n)\sum_t |A_t - F_t|$
Mean Absolute Percent Error	$\text{MAPE} = \frac{100\%}{n}\sum_t |(A_t - F_t)/A_t|$

2.4. Stereographic Method for CBH Estimation

Obtaining CBH from paired sky images has been done by several authors in the past. For instance, the method proposed in [46] generated blocks of clouds which are computed to make a 3D reconstruction of the clouds. Then, the authors were able to obtain the height of the clouds by using geometric computation. On the other hand, the method in [66] used the cross correlation of non-projected saturation images to find all possible combinations that yield feasible heights, selecting the most correlated one (or the one with the minimum error).

In the GRACIOSA project, a pure geometric method based on the relative position of the sky-imagers and the clouds was implemented. First, the algorithm looks for the same cloud feature in the images coming from the two sky-imagers. This task can be extremely difficult due to the

140

chaotic nature of clouds and slight changes in image properties such as luminosity. But fortunately, a Scale-Invariant Feature Transform (SIFT) algorithm [67], can handle this situation. This algorithm performs excellently, identifying features in a constantly changing shape (as clouds), since it considers possible changes in scale and orientation. SIFT is applied to pairs of simultaneous images from both cameras (Figure 16). Once the features from both images have been paired up, the best matches are selected to continue the calculations. Valid features are then transposed from the image (pixels) to real space (azimuth and zenith). With real space coordinates defined and the projection matrix of the lenses known, geometric computation is used to obtain the length of the vectors containing each feature and the geographical position of each camera in real space, from which the height of the evaluated feature can be derived.
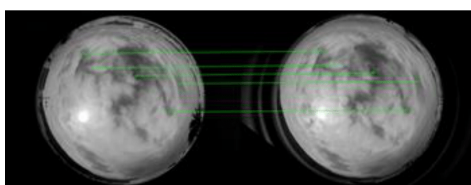

Figure 16. Feature matching by SIFT algorithm in a pair of images from both cameras.

3. Results

Almost a terabyte of image data was collected at NREL from 15 October 2015 through 16 April 2016. The measured DNI values were recorded using NREL's CHP1-L pyranometer with units of Watts/m^2. Days were grouped into three categories: (1) Clear Sky; consisting of predominantly clear days with little or no cloud cover; (2) Overcast; large masses of clouds that obscure the sun for most of the day; and finally (3) Moderately Cloudy; characterized by large variation in irradiance values and multiple ramp events. For Clear Sky and Overcast conditions there is no forecasting to do –persistence can't be beat. Other cloud cover classifications are possible. One could use unsupervised ML clustering algorithms on the raw irradiance data to find other breakdowns. Standard METAR cloud classification separates clouds into low, middle, and high level. All clouds affect measured irradiance [68]. We focus on cumulus clouds because they have the greatest effect on ramp events.

Prediction of intra-hour GHI can be partitioned into several distinct sub-tasks. (1) Acquiring a time series of all-sky images. Every 15 s a new raw image is fused [69] from three different exposure times to allow for High Dynamic Range (HDR) [70]. (2) Using the recent past images and optical flow to extrapolate cloud locations 15 min into the future. It is possible to enhance the algorithm, but it must not hamper production of real time forecasts. (3) Using the predicted image and the weights from training the neural network, a predicted GHI value is output to the microgrid management system.

In the original configuration, our software used optical flow to track movement of cumulus clouds and then ray tracing to predict cloud shadow locations. A better methodology would utilize artificial intelligence (AI) to classify the reduction in GHI that will result when cumulus clouds are predicted in the circumsolar region. It is expensive to train neural networks, but this is done offline for a given location. Once optimal weights are determined, the calculation of a single GHI value is very fast–amounting to an inner product. If the optical flow calculation requires too much time, a second single board computer can be added to the hardware as was employed at NREL.

Details of our research on machine learning to predict solar irradiance are described in the paper [9]. Our intent here is to provide the reader with a concise summary of that research. A critical outcome was verification that the SkyImager with the Pi camera can measure GHI in real time. Lacking the accuracy of an expensive pyranometer, this approach would use the image sequence acquired to solve the forecasting problem, in order to simultaneously estimate GHI. This data could be

incorporated with readings from the WeatherBoard to provide additional inputs to the MGMS using MQTT or DDS protocols. Variables that are considered deterministic should be treated as stochastic random variables. Convolutional neural networks [58] which preserve spatial information offer the best performance for image datasets. Expensive offline training of the networks is normally done one time but it is possible to do continuous learning where the networks use feedback in the form of newly acquired data to refine the learned weights.

In the field of machine learning, there are standard ways of visualizing both the input dataset consisting of vectors in a high dimensional space, as well as targets and predicted outputs. The UTSA SkyImager collected all images in this study on the ESIF building rooftop at NREL as part of the INTEGRATE project. During 147 days from October, 2015 until May, 2016 there were 14 days of no data collected (technical reasons) and 27 days of partial data collection. The remaining 106 days of no missing data formed the inputs for training and testing the neural networks. This yielded 156,495 observations (examples or rows) for input to the neural networks. We used the standard split $(70\% - 30\%)$ of the data into training and testing subsets: 109,547 examples for training and 46,948 for testing.

3.1. Comparing 4 Different ML Models

Each input example is uniquely associated with one of the SkyImager pictures taken every 15 s. The normalized pixel values for the Red-Green-Blue (RGB) channels of an 8×8 resized subimage centered about the sun are flattened into one row vector. Note that other color spaces such as HSV or HSL could also be used. Resizing provided dimensionality reduction and reduced runtimes, but with substantial computer resources the 128×128 sub-images could be used for training. Average values of each channel were included as additional features for a total of $3 \times 64 + 3 = 195$ features. The first entry in each row is the measured GHI in W/m^2. To show how well random variables X and Y are correlated, one uses a scatter diagram. Figure 17 shows scatter diagrams of measured GHI versus predicted GHI for four ML models: Multi-Layer Perceptron (MLP), Random Forest (RF), Deep Learning (DL), and Gradient Boosted Trees (GBT). While all models perform well (tight clustering around $y = x$, perfect correlation), DL and GBT have fewer outliers and visually outperform the other two models.

Note that the MLP and RF models were run using the Scikit-Learn ML software package while DL and GBT were run on the Rapidminer platform. This validated our results on different ML packages—results should depend on the algorithms, not the platform on which they are implemented. Currently, Scikit-Learn does not offer a deep learning model. Using *Rapidminer* is very convenient on a powerful desktop PC, our ultimate goal is to use the trained weights on a Raspberry Pi 3 computer for real time forecasting. Scikit-Learn with its open source Python interface should prove valuable for that task.

MAE, MAPE, nRMSE, and R^2 error metrics are given in Table 2 and run times in minutes. The explained variance (R^2) in the last column is very significant. Both DL and GBT achieve values of 0.87, while MLP and RF are 0.1 less. The other error metrics closely track the R^2 values. In DL the extra accuracy is at the expense of much longer run times, but GBT gets the highest accuracy and is very fast: 10 min faster than RF.

Table 2. Evaluation of four machine learning models.

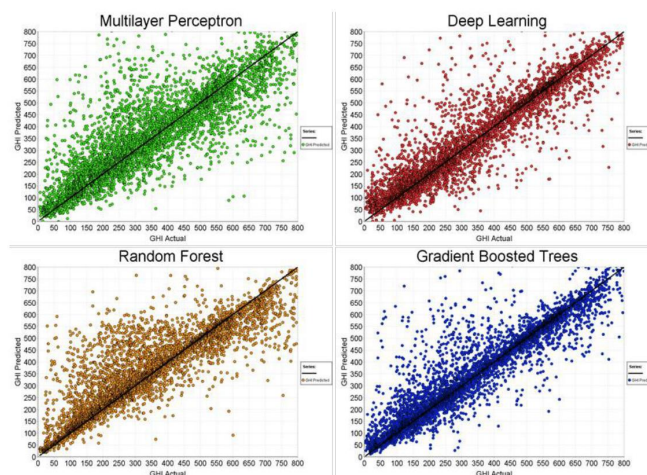| ML Model	Platform	Runtime (min)	MAE	MAPE	nRMSE	R^2
Multilayer Perceptron	Scikit-Learn	3:05	81.21	33.05%	32%	0.71
Random Forest	Scikit-Learn	14:53	66.86	28%	29%	0.76
Deep Learning	Rapidminer	43:52	50.992	27.13%	21.6%	0.871
Gradient Boosted Trees	Rapidminer	4:50	47.072	22.73%	21.1%	0.875

142

Figure 17. GHI actual vs GHI predicted on 24 October 2015 using different AI models.

A time series is another approach to visualizing the results: GHI values are plotted on the *y*-axis and time on the *x*-axis. Figure 18 compares measured and predicted GHI for one day, 3 October 2015. Although there are differences in the two curves, they track each other well. More illuminating is a time series display for the entire group of testing days shown in Figure 19, where actual GHI is blue and forecast values are red. It is difficult to distinguish the two curves because they track each other so closely. For both figures the deep learning model was used for prediction. Observe in the center of Figure 19 a group of five consecutive clear sky days that are easy to predict, as are completely overcast days.
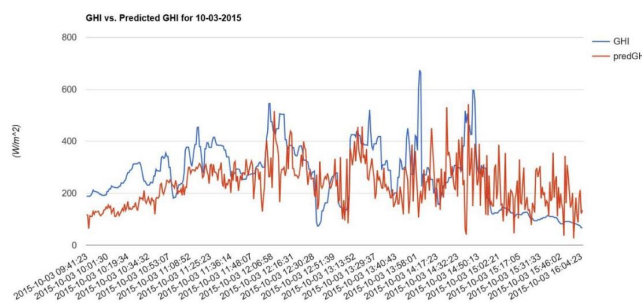
Figure 18. Time series of actual versus predicted GHI for single day of data.

The two curves differ most on moderately cloudy days with air mass cumulus clouds present.
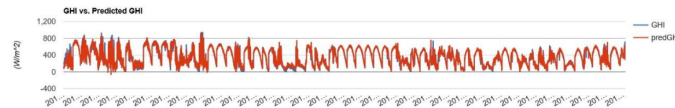
143

Figure 19. Time series of measured and predicted GHI for all testing days.

3.2. Different Deep Learning Model Results

Machine learning algorithms have many hyper-parameters that can be optimized to improve accuracies and reduce run times. Table 3 shows how changing the number of hidden layers for the DL model, nodes in the layers, and number of epochs (complete passes through the training dataset) affects results. Model 1 has 2 hidden layers each with 50 nodes; it requires ten epochs to train with a run time of ~2 min and $R^2 = 0.815$. Model 2 also has two hidden layers (195,195) and 10 epochs, but runs 3 times longer and only improves R^2 to 0.824. To achieve $R^2 = 0.871$ Model 3 (195,195,195) needs 500 epochs and ~45 min. A point of diminishing returns is reached: Model 4 (195,195,97, 195,195) takes more than an hour to run on a desktop PC. Further improvements in accuracy would require larger input images or tuning DL parameters.

Table 3. Comparison of 4 Deep Learning Models.

Hid. Layer	Nodes per H. Layer	# Epochs	Run Time	MAE	MAPE	nRMSE	R^2
2	50,50	10	1:55 min	65.807	32.73%	25.8%	0.815
2	195,195	10	7:10 min	66.872	40.76%	25.6%	0.824
3	195,195,195	500	43:52 min	50.992	27.13%	21.6%	0.871
5	195,195,97,195,195	100	67:42 min	48.519	26.09%	21.6%	0.871

3.3. Cloudy Versus Clear Sky Days

From each 1024×768 fused raw image, the algorithm extracts a 128×128 pixel subimage centered on the sun. Using the *transform.resize* function from Skimage, this is resized to 8×8 pixels to achieve dimensionality reduction. This idea is critical to successful machine learning: each example in the dataset is a vector in a high-dimensional space and there are many examples. Principal Component Analysis and Linear Discriminant Analysis are other techniques for reduction, but our approach is simple and has proved to be effective.

In the following case study ten days of SkyImager data acquired at NREL were used to synthesize two datasets. Five moderately cloudy days comprised the first dataset, October 16, 17, 18, 19, 20 in 2015. Five clear sky days of data from November 10, 11, 12, 13, 14 of 2015 made up the second dataset. The neural networks were fed 32×32 pixel resized images and 4 ML models from Scikit-Learn were compared: Generalized Linear Regression Model (GLM), Multi-Layer Perceptron (MLP), Random Forest Regressor (RFR), and Gradient Boosted Trees (GBT). Table 4 and Figure 20 show the results. GLM and GBT have much shorter runtimes in both cases, while MLP and RFR achieve higher R^2 values. Maximum accuracies are achieved with MLP but at a cost of increased runtimes. Extreme accuracy in R^2 values for clear sky days (0.97, 1.0, 1.0, 0.99) indicates the networks are learning the analytic form of the Haurwitz clear-sky GHI model well.

144

Table 4. Results for two datasets of 32 × 32 resized sub-images.

Model Name	5 Moderately Cloudy Days				5 Clear Sky Days			
	GLM	MLP	RFR	GBT	GLM	MLP	RFR	GBT
MAPE	39.75	13.53	15.79	20.93	12.0	2.44	2.66	3.79
Explained Variance R^2	0.70	0.95	0.93	0.90	0.97	1.00	1.00	0.99
Mean Absolute Error	86.43	30.66	33.30	44.68	20.43	5.30	5.84	8.57
Elapsed Time (Sec)	0.1	513.1	69.9	12.4	0.1	512.8	60.5	12.0
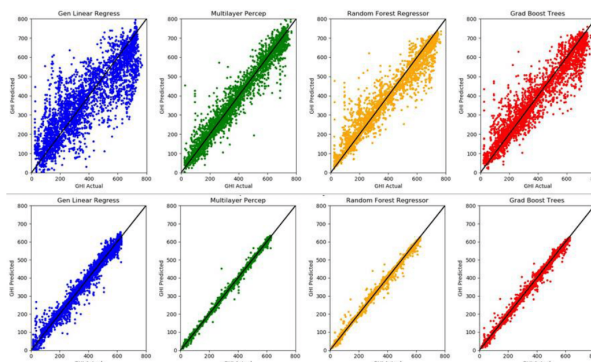
Figure 20. Scatterplots of Actual GHI versus Predicted GHI for 4 ML Models on moderately cloudy days (top row) and clear sky days (bottom row).

Using still finer sampling for the resized sub-images should yield better results at the cost of larger data files and runtimes. At some point however, statistics suggests diminishing returns. In addition to detailed descriptions of ML models and software our article [9] presents another case study. It uses only one moderately cloudy day (17 October 2015) of observations and runs the four ML models with 8 × 8, 32 × 32, and 64 × 64-pixel sub-images. The size of the CSV input data file increases quickly: 3 Megabytes, 111 Mb, and 442 Mb, as do runtimes 139 s, 444 s, and 1309 s. Accuracies improve, but not beyond a certain point.

3.4. JBSA Microgrid Data

The JBSA microgrid was built as a testbed for the CPS Energy Grid Modernization Laboratory. Management and control of a microgrid must address many factors including cybersecurity, data acquisition, data management, real-time computation, storage, bandwidth, interoperability, and usability requirements. In addition to the data acquired by the UTSA equipment–SkyImager, WXT520 Vaisala weather station, and pyranometer; this includes 36-hour ahead hourly weather forecasts scraped from the web, the day-ahead Load/PV forecasts, battery State-Of-Charge (SOC) readings, actual load for the base library building, and control data from the Siemens MGMS. The goal is to use all available data in order to refine site-specific solar irradiance forecasts for improved operation and control of the microgrid. Non-UTSA data from the JBSA microgrid was acquired from Itron MV-90 xi meters. This is a system used for the collection and management of interval data consisting of time stamped readings taken every x minutes where x can be 5, 15, 30, or 60. A large electric utility may acquire a billion interval readings in a single year and use them in a variety of ways

145

including billing (demand response, real time pricing, curtailable rates), open market operations, and load/market research.

Analyses of the JBSA MV90 data, all of which were taken at 15 min intervals, demonstrated that much finer temporal resolution would be required to capture details of ramp events and provide accurate irradiance forecasts to the MGMS. While 1 min resolution was provided by the UTSA equipment such as the pyranometer and Vaisala weather station, the cost of this equipment precluded widespread deployment in a distributed environment. Similar equipment at the UTSA solar testbed provided a wealth of 1 min data, but we envisioned a network of hundreds of low cost SkyImagers spread across the city of San Antonio, and for this scenario low cost was an essential requirement. As previously discussed, there are a plethora of low-cost sensors that can be connected to the GPIO pins on a SBC, including the WeatherBoard2 (WB2), air quality and ozone sensors, and even devices to measure dust. Costing tens of dollars they provide a cost-effective way of environmental sensing that can easily be incorporated with the SkyImager. Figure 21 shows observations of irradiance, temperature, humidity, and pressure taken every minute on 26 November 2018 with the WB2 sensor mounted on a SkyImager.
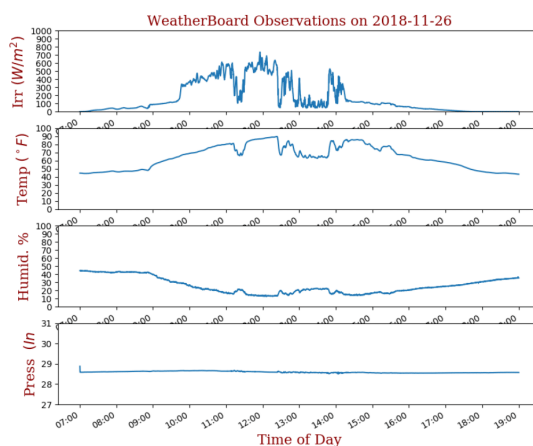

Figure 21. Data from the WeatherBoard sensor on the UTSA SkyImager.

The mini photovoltaic module used at ULL is an off-the-shelf PV module, made of c-Si, with open circuit voltage of 6V and a short circuit current of 200 mA, with a maximum DC output of 1.1 W. The module was connected to a resistive load to dissipate the heat, and the data was registered by a INA219 DC Current Sensor able to measure very small currents, attached to the Raspberry Pi 3 model B.

3.5. One Second Minimodule Data from La Graciosa

The Universidad de La Laguna in the Canary Islands provided 1 sec data from the minimodule at the La Graciosa microgrid operated by ENDESA. How does the spectral content of the voltage signal change when moving from 1 s, 5 s, 15 s, to 60 s sampling? Figure 22 shows the effect of sub-sampling on the time-series. Some of the noise present in the data might be due to voltage fluctuations or seagulls (the location was the Fisherman's Guild building). Certainly the area of the minimodule is very small, so it behaves as a point measurement where small occluding objects can drastically the

146

affect irradiance measurements. Still, the analysis strongly suggests that to resolve the frequency and voltage swings that occur during a sudden ramp event, PMU measurements may be a necessity.
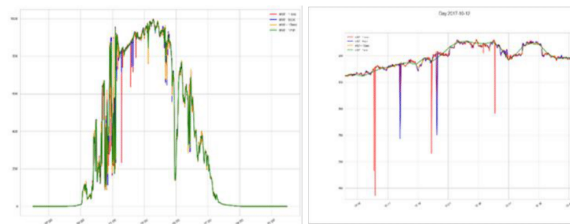

Figure 22. Effect of subsampling on the 1-sec minimodule data from La Graciosa.

3.6. CBH Estimations

The results of the CBH estimations are presented in Figure 23. The left boxplot shows the statistical distribution of the heights obtained with the stereographic method, while the right boxplot shows the statistical distribution of a weather station located in Arrecife, Lanzarote, which belongs to the network of the University of Wyoming.
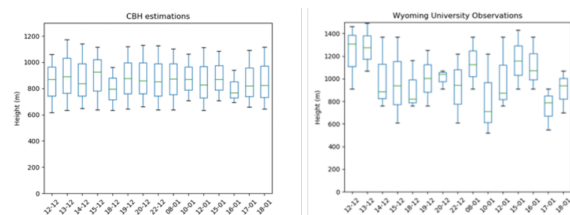
Figure 23. Statistical distribution of the heights obtained by two-camera stereographic method (**left**) and from a weather station located in Arrecife, Lanzarote (**right**).

There are several reasons for the apparent mismatch in the data. First, the weather station is located 30 km south of the position of the cameras, which undoubtedly has a significant effect taking into account how the atmospheric conditions develop in the region (with the thermal inversion steadily rising its level from the Sahara Desert). Second, the data of the weather station is obtained using a punctual measurement such as LIDAR, while the CBH estimations of the cameras cover a larger area of the sky (mainly the central part of the fish-eye image, since the distortion on the borders makes it almost impossible to compute the height). Finally, the temporal resolution of the weather station data is up to 1 hour, while the estimation of the CBH by the stereographic approach is done every minute. Likely the most important conclusion here is that the estimations made by the systems are coherent with the previous knowledge of the atmospheric conditions, with a stable thermal inversion ranging from 600 to 2000 m depending on the season of the year, which prevents clouds to rise over a certain height.

4. Discussion

Our future research efforts will be directed in several areas. Optical flow is a critical area for the success of intra-hour solar forecasting. IoT and cyber-security also form a critical component. Solving

147

the AC Optimal Power Flow equations with GHI forecasts from the SkyImager will be important for solving the energy storage and microgrid control problems. Using the Raspberry Pi additionally as a multiple-sensor platform will be investigated. Environmental studies of the effects of dust and bird feces on the solar panels may well utilize SkyImager technology.

It was on the INTEGRATE project that a synergism developed between researchers and engineers at the national lab, universities, utilities, and private industry that continued after the project ended. Management styles are quite different in academia and private industry with national labs somewhere in between. Software version control was critical, as was careful documentation of all work. Both for the utility where engineers would use the hardware/software and for the university where graduate students and faculty would move on to other projects this was very important. Our decision to use Python was an excellent one. Increasingly, both documentation, tutorials, and example programs are being delivered in the form of IPython notebooks (.IPNB files) as, for example, Google's Tensorflow. Even a package such as Open Computer Vision [71,72] that is written in C++ for efficiency has Python bindings that allow easy access to routines for image fusion and optical flow.

4.1. Lessons Learned at the 3 Deployment Locations

4.1.1. SkyImager at NREL

The NREL microgrid was located at a specialized research facility, but every attempt was made to simulate conditions at a utility. The communications network was well established; there was abundant state-of-the-art ancillary equipment such as pyranometers and an on-site weather station, and the process of deployment went relatively smoothly. Still there were important lessons to be learned. The SkyImager was initially configured with a single Raspberry Pi 2, which proved insufficient for both acquiring images and processing them through the pipeline to produce irradiance forecasts. This problem was solved by adding an Odroid C1, but this made the design more complex and required bridging between the two SBC using a USB-internet connection. The plethora of operating systems for both the SBC and EMS servers provided still another challenge. There are differences in the way open source packages such as OpenCV and Mosquitto install and operate on Raspbian, Ubuntu 14, and Ubuntu 16. In some cases, there are compiled binaries available and in others software must be compiled from source files. From a solar forecasting perspective, NREL was where the best, most complete data was acquired: over six months of daily images and ground truth pyranometer observations. It took researchers over a year to analyze the data and the process is ongoing.

4.1.2. SkyImager at San Antonio, TX, USA

Two critical applications of islanded microgrids are remote installations in developing countries and power systems for the military that must be entirely stand alone. While this made a military base the perfect site for testing a microgrid EMS, it also meant that obtaining base access for UTSA researchers was an issue. For safety purposes, it requires at least two people to lower the 10 m MET tower. Beyond the initial installation, access to the tower is required every month to inspect the instruments and clean the surface of the plastic dome that covers the camera enclosure. The cost of the pyranometer and weather station exceeded that of the SkyImager by a factor of 20 and prompted adding low-cost sensors to the SBC for measuring temperature, humidity, and light.

Occasional loss of power to the SBC as the battery went through its initial testing phase was a minor issue which required tinkering so that the forecasting software was immediately brought back on line during a reboot. Initially the Mosquitto MQTT broker was chosen for the UTSA software, but because of compatibility issues with the MyRio software, HiveMQ proved to be a better choice. Direct internet access to our SBC from outside the corporate network was available only through a WebEx session that required close coordination between the utility and university personnel. Lesson learned: when initially transitioning hardware/software from a research environment to a production one, it is imperative to have physical/cyber access to the equipment and network. Several Odroid's were

damaged by high temperatures in the enclosure boxes and had to be replaced. MV90 meter readings every 15 min are clearly insufficient for the intra-hour forecasting problem. We are currently working with a group in Austin to add inexpensive Phasor Measurement Units (PMU) to acquire detailed frequency and phase information in conjunction with weather, irradiance, and sky images. Cloud computing [73] and 5G networks offer unique opportunities to move most of the computations off the Raspberry Pi to the cloud.

4.1.3. SkyImager at La Graciosa, Canary Islands

Some valuable lessons learned from the Graciosa project have to do with the performance and durability of the sky-imagers in a harsh, dusty, and salty environment such as La Graciosa. The closeness of the island to the Sahara Desert makes the dust content in the atmosphere high. Quality of the images is severely affected by the deposition of dust on the enclosure, as seen in Figure 24, showing that scheduled cleanings of the enclosure are necessary to ensure the proper function of the devices. Our first estimation is that cleaning is necessary at least twice a year, but it is highly dependent on the climatologic and atmospheric conditions. Besides the deposition of dust on the enclosure, water infiltrations in the interior of the enclosure have been a frequent problem, even if the equipment was specially selected to have a high degree of protection to water and dust (IP67). Closeness to the sea, as well as the strong rains that occurred in December of 2017, appear to be the main source of this problem.
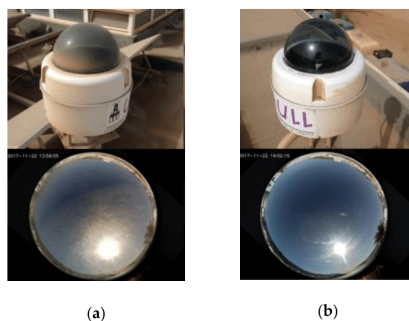

(**a**)   (**b**)

Figure 24. (**a**) SkyImager/image with dust on enclosure, (**b**) after cleaning it.

5. Conclusions

In March of 2018 the California Energy Commission mandated that beginning in 2020, all new home and apartment construction must include solar generation. When this level of distributed generation becomes part of the electric grid, ISO are faced with new challenges in terms of frequency and voltage control. Indeed, in Hawaii these issues resulted in a temporary hold on rebates for new residential solar installations.

In a macrogrid extending over hundreds of square kilometers, integrating a mix of generation conventional power plants, solar, and wind, and functioning as part of a larger interconnect such as ERCOT, there is an inherent inertia that works on the side of the utility. In microgrids, however, this inertia is lacking, and the control problem becomes much more difficult to solve when in islanded mode [4,7]. There are also different temporal scales involved. Optimal day-ahead scheduling and control of a microgrid is a distinct problem from hour-ahead control to ensure frequency and voltage do not vary outside prescribed limits. Using equipment such as the OP4500 RT-LAB/RCP/HIL real-time

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

power grid digital simulator by Opal RT, we hope to use Hardware-in-the-Loop equipment to analyze the microgrid at JBSA. PMU measurements also need to be incorporated.

All-sky imaging technology will be a critical component in the overall solution strategy to predict solar irradiance 15 min ahead, and to take corrective measures during ramp events. It must, however, be fully integrated with NWP and satellite-based approaches for day-ahead load forecasting and optimal control of a microgrid. Optimal use of this technology will encompass a diverse group of specializations, including IoT and edge-computing, cyber-security, machine learning, and image processing.

For example, the characteristics and statistics of the all-sky imager must be included in the stochastic optimization programming for risk neutral and risk adverse operational control of a microgrid [6]. An holistic R&D approach is required. While a Raspberry Pi is the essence of plug-N-play and it is relatively straightforward to build a SkyImager, integration into the IoT and field deployment will remain an active area of research. Imagers will range the gamut in cost, accuracy, and interoperability. MGMS will integrate forecasts from imagers, NWP, and satellites, as well as hundreds of other meters and devices to solve the microgrid control problem. While physics-based methodology will continue to be important, machine learning and IoT technology will play an increasingly critical role. Development of standards such as OpenFMB for interoperability of thousands of devices will also be a necessary component.

6. Patents

A provisional US patent application "distributed solar energy prediction imaging" has resulted from the work reported in this manuscript.

References

1. Guerrero-Lemus, R.; Shephard, L. *Low Carbon Energy in Africa and Latin America: Renewable Technologies, Natural Gas and Nuclear Energy (Lecture Notes in Energy)*, 1st ed.; Springer: Berlin, Germany, 2017.
2. Available online: https://uc-ciee.org/downloads/appendixA.pdf (accessed on 16 February 2019).
3. Janssen, T.; Krishnaswami, H. Voltage and Current Control of a Multi-port NPC Inverter Configuration for a Grid-Connected Photovoltaic System. In Proceedings of the IEEE 17th Workshop on Control and Modeling for Power Electronics, Trondheim, Norway, 27–30 June 2016.
4. Olivares, D.; Lara, J.; Canizares, C.; Kazerani, M. Stochastic-Predictive Energy Management System for Isolated Microgrids. *IEEE Trans. Smart Grid* **2015**, *6*, 2681–2693. [CrossRef]
5. Canizares, C.; Palma-Behnke, R.; Olivares, D.; Mehrizi-Sani, A.; Etermadi, A.; Iravani, R.; Kazerani, M.; Hajimiragha, A.; Gomis-Bellmunt, O.; Saeedifard, M.; et al. Trends in Microgrid Control. *IEEE Trans. Smart Grid* **2014**, *5*, 1905–1919.
6. Farzan, F.; Jafari, M.; Masiello, R.; Lu, Y. Toward Optimal Day-Ahead Scheduling and Operational Control of Microgrids Under Uncertainty. *IEEE Trans. Smart Grid* **2015**, *6*, 499–507. [CrossRef]
7. Michaelson, D.; Mahmood, H.; Jiang, J. A Predictive Energy Management System Using Pre-Emptive Load Shedding for Islanded Photovoltaic Microgrids. *IEEE Trans. Ind. Electron.* **2017**, *64*, 5440–5448. [CrossRef]

150

8. Jr, W.R.; Krishnaswami, H.; Vega, R.; Cervantes, M. A low cost, edge computing, all-sky imager for cloud tracking and intra-hour irradiance forecasting. *Sustainability* **2017**, *12*, 482.
9. Moncada, A.; Richardson, W., Jr.; Vega-Avila, R. Deep Learning to Forecast Solar Irradiance Using a Six-Month UTSA SkyImager Dataset. *Energies* **2018**, *11*, 1988. [CrossRef]
10. Nummikoski, J.; Manjili, Y.S.; Vega, R.; Krishnaswami, H. Adaptive Rule Generation for Solar Forecasting: Interfacing with A Knowledge-Base Library. In Proceedings of the IEEE 39th Photovoltaic Specialists Conference (PVSC), Tampa, FL, USA, 16–21 June 2013.
11. Cervantes, M.; Krishnaswami, H.; Richardson, W.; Vega, R. Utilization of Low Cost, Sky-Imaging Technology for Irradiance Forecasting of Distributed Solar Generation. In Proceedings of the IEEE GreenTech Conference, Kansas City, MO, USA, 6–8 April 2016. [CrossRef]
12. Cañadillas, D.; Richardson, W., Jr.; González-Díaz, B.; Shephard, L.; Guerrero-Lemus, R. First Results of a Low Cost All-Sky Imager for Cloud Tracking and Intra-Hour Irradiance Forecasting serving a PV-based Smart Grid in La Graciosa Island. In Proceedings of the IEEE PVSC-44, Washington, DC, USA, 25 June 2017.
13. Richardson, W.; Krishnaswami, H.; Shephard, L.; Vega, R. Machine Learning versus Ray-Tracing to Forecast Irradiance for an Edge-Computing SkyImager. In Proceedings of the 19th International Conference on Intelligent System Application to Power Systems (ISAP), San Antonio, TX, USA, 17–20 September 2017.
14. Waight, J.; Grover, S.; Laval, S.; Shephard, L.; Boston, J.; Lui, R.; Mathew, J.; Bradley, D.; Lawrence, D.; Sparkman, M.; et al. NREL Integrate: RCS -4-42326: Topic Area 3 OpenFMB Reference Architecture Demonstration Final Report, Minneapolis, 2017.
15. Mathiesen, P.; Kleissl, J. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. *Solar Energy* **2011**, *85*, 967–977. [CrossRef]
16. Available online: https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/rapid-refresh-rap (accessed on 20 December 2018).
17. Xia, S.; Mestas-Nuñez, A.M.; Xie, H.; Vega, R. An Evaluation of Satellite Estimates of Solar Surface Irradiance Using Ground Observations in San Antonio, Texas, USA. *Remote Sens.* **2017**, *9*, 1268. [CrossRef]
18. Mueller, R.; Trentmann, J.; Träger-Chatterjee, C.; Posselt, R.; Stöckli, R. The Role of the Effective Cloud Albedo for Climate Monitoring and Analysis. *Remote Sens.* **2011**, *3*, 2305–2320. [CrossRef]
19. Urbich, I.; Bendix, J.; Muller, R. A Novel Approach for the Short-Term Forecast of the Effective Cloud Albedo. *Remote Sens.* **2018**, *10*, 995. [CrossRef]
20. Perez, R.; Ineichen, P.; Moore, K.; Kmiecik, M.; Chain, C.; George, R.; Vignola, F. A New Operational Model for Satellite-Derived Irradiances: Description and Validation. *Solar Energy* **2002**, *73*, 307–317. [CrossRef]
21. Law, E.; Prasad, A.; Kay, M.; Taylor, R. Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting: A review. *Solar Energy* **2014**, *108*, 287–307. [CrossRef]
22. Raza, M.Q.; Nadarajah, M.; Ekanayake, C. On recent advances in PV output power forecasting. *Sol. Energy* **2016**, *136*, 125–144. [CrossRef]
23. Barbieri, F.; Rajakaruna, S.; Gosh, A. Very short-term photovoltaic power forecasting with cloud modeling: A review. *Renew. Sustain. Energy* **2017**, *75*, 242–263. [CrossRef]
24. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de_Pison, F.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Solar Energy* **2016**, *136*, 78–111. [CrossRef]
25. Marquez, R.; Coimbra, C. Intra-hour DNI forecasting based on cloud tracking image analysis. *Sol. Energy* **2013**, *91*, 327–336. [CrossRef]
26. Marquez, R.; Coimbra, C. Proposed Metric for Evaluation of Solar Forecasting Models. *J. Sol. Energy Eng.* **2013**, *135*, 011016. [CrossRef]
27. Gohari, S.; Urquhart, B.; Yang, H.; Kurtz, B.; Nguyen, D.; Chow, M.; Kleissl, J. Comparison of solar power output forecasting performance of the Total Sky Imager and the University of California, San Diego Sky Imager. *Energy Procedia* **2014**, *49*, 2340–2350. [CrossRef]
28. Urquhart, B.; Kurtz, B.; Dahlin, E.; Ghonima, M.; Shields, J.; Kleissl, J. Development of a sky imaging system for short-term solar power forecasting. *Atmos. Meas. Tech. Discuss.* **2014**, *7*, 4859–4907. [CrossRef]
29. Chow, C.W.; Urquhart, B.; Lave, M.; Dominguez, A.; Kleissl, J.; Shields, J.; Washom, B. Intro-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Sol. Energy* **2011**, *85*, 2881–2893. [CrossRef]
30. Chu, Y.; Pedro, H.; Coimbra, C. Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning. *Solar Energy* **2013**, *98*, 592–603. [CrossRef]

53. Dong, B.; Li, Z.; Rahman, S.; Vega, R. A hybrid model approach for forecasting future residential electricity consumption. *Energy Build.* **2016**, *117*, 341–351. [CrossRef]
54. Li, Z.; Rahman, S.; Vega, R.; Dong, B. A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting. *Energies* **2016**, *9*, 55. [CrossRef]
55. Salakhutdinov, R.; Hinton, G. Deep Boltzmann Machines. In Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS'09), Clearwater Beach, FL, USA, 16 April 2009.
56. Bengio, Y.; LeCun, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444.
57. Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
58. Mallat, S. Understanding deep convolutional networks. *Phil. Trans. R. Soc. A* **2016**, *374*, 20150203. [CrossRef]
59. Raschka, S. *Python Machine Learning, Birmingham*; PACKT Publishing: London, UK, 2016.
60. Hofmann, M.; Klinkenberg, R. (Eds.) *RapidMiner—Data Mining Use Cases and Business Analytics Applications*; Chapman and Hall/CRC: New York, NY, USA, 2013.
61. Available online: http://www.h2o.ai/wp-content/themes/h2o2016/images/resources/DeepLearningBooklet.pdf (accessed on 16 February 2019).
62. Available online: http://h2o-release.s3.amazonaws.com/h2o/rel-turchin/3/docs-website/h2o-docs/booklets/GBM_Vignette.pdf (accessed on 16 February 2019).
63. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach.Learn. Res.* **2011**, *12*, 2825–2830.
64. Reno, M.; Hansen, C.; Stein, J. *Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis*; SANDIA: Albuquerque, NM, USA, 2012.
65. Hinton, G.; Salakhutdinov, R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]
66. Nguyen, D.; Kleissl, J. Stereographic methods for cloud base height determination using two sky imagers. *Solar Energy* **2014**, *107*, 495–509. [CrossRef]
67. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
68. Li, Q.; Lu, W.; Yang, J.; Wang, J. Thin Cloud Detection of All-Sky Images Using Markov Random Fields. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 1545–1558. [CrossRef]
69. Mertens, T.; Kautz, J.; van Reeth, F. Exposure fusion. In Proceedings of the 15th Pacific Conference on Computer Graphics and Applications (PG'07), Maui, HI, USA, 29 October–2 November 2007; pp. 382–390.
70. Traonmilin, Y.; Aguerrebere, C. Simultaneous High Dynamic Range and Superresolution Imaging without Regularization. *SIAM J. Imaging Sci.* **2014**, *7*, 1624–1644. [CrossRef]
71. Available online: https://opencv.org/ (accessed on 16 February 2019).
72. Available online: https://docs.opencv.org/3.0-beta/opencv2refman.pdf (accessed on 16 February 2019).
73. Nagothu, K.; Kelley, B.; Jamshidi, M.; Rajaee, A. Persistent Net-AMI for Microgrid Infrastructure Using Cognitive Radio on Cloud Data Centers. *IEEE Syst. J.* **2012**, *6*, 4–15. [CrossRef]

153

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

154

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por:	David Cañadillas Ramallo	Fecha: 02/06/2021 14:51:11
	UNIVERSIDAD DE LA LAGUNA	
	María de las Maravillas Aguiar Aguilar	07/06/2021 16:06:22
	UNIVERSIDAD DE LA LAGUNA	

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

Annex III: EDA-based optimized global control for PV inverters in distribution grids

155

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

156

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
UNIVERSIDAD DE LA LAGUNA

María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
UNIVERSIDAD DE LA LAGUNA

ORIGINAL RESEARCH PAPER

EDA-based optimized global control for PV inverters in distribution grids

David Cañadillas[1] | **Hamed Valizadeh[3]** | **Jan Kleissl[3]** | **Benjamín González-Díaz[2]** | **Ricardo Guerrero-Lemus[1]**

[1] Departamento de Física, Universidad de La Laguna, Avenida Astrofísico Francisco Sánchez, S/C de Tenerife, Spain

[2] Departamento de Ingeniería Industrial, Universidad de La Laguna, Avenida Astrofísico Francisco Sánchez, S/C de Tenerife, Spain

[3] Center for Energy Research, Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, California, USA

Correspondence
David Cañadillas, Departamento de Física. Universidad de La Laguna. Avenida Astrofísico Francisco Sánchez, S/N, 38206, S/C de Tenerife, Spain.
Email: dcanadil@ull.edu.es

Abstract

Operating distribution grids is increasingly challenging due to the increasing penetration of photovoltaic systems. To address these challenges, modern photovoltaic inverters include features for local control, which sometimes lead to suboptimal results. Improved communication infrastructure and photovoltaic inverters favour global control strategies, which receive information from all the systems in the grid. An estimation of distribution algorithm is used to optimize a global control strategy that minimizes active power curtailment and use of reactive power of the photovoltaic inverters, while maintaining voltage stability. Optimized global control outperforms every other local control evaluated in terms of apparent energy used for control (9.9% less usage compared to the second best alternative in all scenarios studied) and ranks second in terms of voltage stability (with a 0.14% of total time outside the voltage limits). Two new indicators to compare control strategies are proposed, and optimized global control strategy ranks best for both efficiency index (0.98) and average apparent power use (0.48 kVA).

1 | INTRODUCTION

Photovoltaic (PV) systems connected to medium-voltage and low-voltage distribution grids pose several engineering challenges to distribution system operators (DSO) who ensure the quality of the electric supply. The most common problems are overvoltage at the end of the feeders and overloading of transformers or lines due to active power injection. Traditionally, these problems have been addressed by reinforcing the grid [1], which is a rather expensive solution.

Another effect of including distributed generation (DG) on distribution grids is a change in the traditional operational paradigm of power systems. The electric system has traditionally been operated in a radial way [2], where large generators supply the power to the loads in an unidirectional manner. With a large number of DG, power flow can be bidirectional and come from the demand-side depending on the load conditions. Active power injection in an electric grid increases the voltage especially in low and medium voltage distribution grids, due to their low X/R ratio (the lines are more "resistive", meaning that

active power injections/absorptions has a larger effect on the voltage).

When the DG are PV systems, the variability of the solar resource caused by passing clouds becomes an important issue. Short-term solar forecasting has been identified as a key strategy for integration of PV systems in the grid, but high accuracy solar forecasts remain elusive. However, results may be difficult to obtain unless dedicated equipment such as sky-imagers are used [3]. The randomness of especially household loads is another source of uncertainty. The mismatch between high PV production and high demand periods exacerbates the issues in distribution grids [4].

Conventional voltage regulation elements on the distribution grid consist of on-load tap changers (OLTC) in the transformers and voltage regulators along the lines. All the above-mentioned issues make it difficult to find a fixed set of operational conditions that ensure the quality and stability of the electric supply. Without DG, these problems have been easily addressed using approaches such as transformer tap changes or voltage regulators in the middle of the distribution grids. However, the

complexity resulting from adding DG at the distribution level requires different solutions.

Nowadays, solar inverters are capable of delivering PV power with a variable power factor (PF), which allows them to perform voltage control depending on different variables, such as voltage or power. The most common local control strategies supported by commercial PV inverters are reactive power compensation (operation with flexible PF) and curtailment of active power [1]. Increasingly countries are adopting new grid codes which require PV system inverters to participate in the regulation of the distribution grid [5]. It may be conceivable that solar inverters become the main source of voltage regulation in distribution grids.

Although, under some circumstances adequate voltage regulation can only be achieved by curtailing the active power injected to the grid, one of the objectives should be to curtail the least amount of clean energy. For that reason, reactive power compensation is a preferred strategy.

Voltage regulation through reactive power control is based on the reactance of the line, ignoring its resistance [6]. Hence, the effectiveness of reactive power compensation is determined by the X/R ratio of the lines which differs depending on the voltage level of the grid, increasing as the voltage increases [7]. For that reason, it is expected that reactive power capabilities are less effective in low voltage circuits, since their lines are more resistive (lower X/R ratio) than in MV and HV networks.

The position of each PV system in the feeder is a critical factor in the behaviour of local controllers [8]. For instance, inverters located near the transformer are less likely to suffer from voltage problems, while those at the end of the line are more exposed to voltage problems. This fact may result in uneven and unequitable use of the solar inverters depending on their position on the feeder. A solution to uneven use includes different configurations for the local controllers (different set points on the curve) depending on the distance of the PV system to the substation [9].

Distribution systems are unbalanced systems. Although the DSO evenly distributes the capacity on each phase, shifts in the timing of loads by individual households can lead to imbalance especially as the size of the distribution grid diminishes. In unbalanced three-phase four-wire systems the power injections in one of the phases effects the other two, which is caused by neutral point shifting [10].

Due to the factors listed above typical local inverter control strategies can lead to solutions far from the optimum. Reactive power usage and active power curtailment may be managed suboptimally and control actions may be distributed suboptimally among the systems. Centralized control is expected to improve the control actions since the cost function takes into account all variables.

The evolution of the communication infrastructures and capabilities of PV systems promotes the introduction of global optimization methods. Global optimization methods can retrieve information from every system on the grid and analyze them and their operational conditions in real time to find optimal operating points for all the systems to maintain power quality [11]. The critical enabler of global optimized controls is the increased communication of the devices in the power grid [12].

EC is a family of heuristics inspired by nature that tries to emulate natural selection, through the generation of a population of candidate solutions, which are evaluated with an objective function (also referred to as cost or fitness function). The best individuals are then selected to reproduce, "spreading their genes" or properties to the next generations of candidate solutions.

However, as the number of PV system increases, so does the number of variables to be tracked and optimized, along with the complexity of the task of monitoring and control. Therefore, fast optimization algorithms are needed to find near optimal solutions in a reduced timeframe. There are multiple families of algorithms and heuristics, such as evolutionary computing (EC), that have started to emerge as useful tools to confront the computational problems derived from high complexity tasks.

Inside the EC family, there is a large set of algorithms and heuristics, and some are well known and commonly implemented in optimization problems in power systems [13–17]. One interesting family of algorithms that has not been widely applied to power systems is an estimation of distribution algorithms (EDAs) [18]. Unlike other EC algorithms, EDAs build a probabilistic model from the best individuals and sample the subsequent generations from it, instead of randomly generating mutations in the genes of the individuals, crossover, or reproducing them with some other arbitrary criteria. By using a probabilistic model, the search space of the algorithm narrows at each iteration, approaching the region with optimal solutions. The EDA will converge faster to the near-optimal solutions than other EC algorithms.

Several references in the literature evaluate the EDA performance against other algorithms, mainly genetic algorithms (GA) due to its similarities. In [19], EDA performance is contrasted with other metaheuristics in an electric vehicle charging problem. In that study, EDA performs best in terms of computational time and quality of the solution. In [20], a node allocation problem in a network is solved using GA and EDA. Again, EDA outperforms GA with these settings, especially as the population grows. Finally, in [21], an in-depth analysis of different metaheuristics is made for a series of NK landscapes. In this study, different strategies for EDA implementation are analyzed. UMDA and an advanced EDA (hierarchical Bayesian optimization algorithm, hBOA) are contrasted with GA. EDAs strategies prove to be better in terms of the number of evaluations.

EDAs have been used in various applications: as an optimizer for a controller for networked dc motor systems [22], system-level synthesis for chip designs [23], carbon emissions reduction and project makespan minimization [24], dynamic deployment of near space communication systems [25], energy-efficient scheduling of cloud computing services [26], or calibrating the parameters for microscopic traffic models [27]. To the best knowledge of the authors, there is only one paper on DG, where an EDA is used to intelligently charge electric vehicles [19] and there are no studies on distribution system optimization with PV using EDAs. EDAs are well suited for distribution system optimization as they can solve large-scale, complex, non-linear
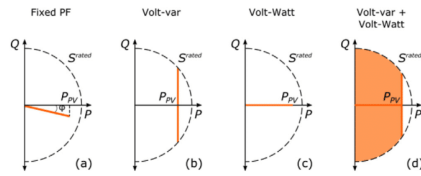
FIGURE 1 Operational regions for the inverter depending on the type of control selected. P_{PV} is the active power from the PV panels. (a) Fixed PF function, (b) VV function, (c) VW function and (d) Combined VV and VW

optimization problems and are immune to local minima. They converge faster to near-optimal solutions and their computational costs are lower compared to other EC alternatives [19].

In this paper, we implement an EDA-based optimization to obtain near-optimal solutions to the set of operational parameters of all PV inverters in a medium-voltage distribution grid to evaluate if they are capable of maintaining the voltage stability and the power quality by themselves. There are two main contributions of this paper. The first contribution is the implementation of an EDA-based optimization global control which optimizes the parameters (active power curtailment and reactive power support) of all the PV systems in a modified IEEE123 medium-voltage distribution grid. This kind of control could be easily deployed in actual grids with the technology currently available in commercial PV inverters. Second, to the authors' knowledge, for the first time, the viability of using solar inverters as the only voltage regulators in the distribution grid is examined.

2 | STANDARD LOCAL CONTROLS OF PV INVERTERS

Although smart inverters can provide a variety of additional functionalities for grid stability or responses to fault events [28], we will focus on voltage support. There are a variety of standard local control strategies that modern smart inverters have implemented. We compare our proposed EDA-based optimized global control (OptGC) to four of these local control methods: the fixed PF function, the Volt-var (VV) function, the Volt-Watt (VW) function, and a combination of VV and VW (VV+VW). The operational regions on the QP diagram for each of the local controls studied are depicted in Figure 1 and described further below.

2.1 | Fixed power factor

Modern PV inverters can operate at PFs different than the unity. Fixed PF control can be used to limit the injections of active power while providing a reactive absorption or injection throughout the day. Reactive power output increases proportionally to the PF, as the PV active power increases [29]. Fixed PF control is usually suboptimal since it implies a continuous curtailment of active power and inefficient use of reactive power

irrespective of grid conditions, but it is easy to implement and its effects can be easily observed.

2.2 | Volt-var function

The VV function is a local controller that reacts to voltage measurements at the point of common coupling of the solar inverter with the feeder [9]. There are several VV configurations: different set points and slopes for the curves, the existence, or lack thereof, of a "deadband", the possibility of including hysteresis, etc. In this study, two different curves for VV control with and without deadband are used, which are shown in Figure 2. Without a deadband (top-left), reactive power compensation starts as soon as the voltage magnitude differs from unity. When the deadband is included, (top-right), the control does not absorb nor inject reactive power until the voltage approaches the set limits (in this case, at 0.97 p.u. and 1.03 p.u.), which in turn, should reduce reactive power usage.

2.3 | Volt-Watt function

The VW function is used to reduce the amount of active power delivered to the system (power curtailment) as a function of the voltage at the PCC of the solar inverter. An operation curve defining the voltage set point where the control should start to act is configured in the inverter [29]. The VW curve used for this study is shown in Figure 2 (bottom). VW is only useful in overvoltage conditions (that is, it does not respond to undervoltage situations) since it does not make use of reactive power.

2.4 | Simultaneous volt-var and volt-Watt function

Another strategy for local control is the combination of both VV+VW controls. As in the only VW control, this strategy is only useful when experiencing overvoltage situations, otherwise, it would behave as an only VV function. A priority of which function dominates over the other has to be defined [29]. If VW has priority, the inverter will first curtail active power before using reactive power compensation. If VV has priority, the inverter will first dispatch reactive power, and if the further reduction of voltage is required, active power will be curtailed. In general, the objective should be to curtail as little active power as possible. For that reason, in our simulations, the VV function has priority. That means that active power may be curtailed only if reactive power utilization it at its limit, allowing a larger share of the inverters apparent power to be used for reactive power compensation.

3 | OPTIMIZATION PROBLEM FORMULATION

The objective of the optimization problem is to find a set of near-optimal points of operation for each PV inverter in a distribution grid that maintains voltage within the limits, while minimizing the curtailment of active power and the use of reactive

159
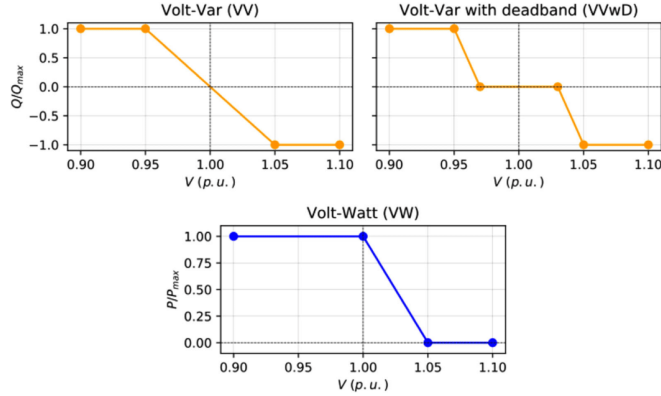
FIGURE 2   Local control function curves. Top-left: VV curve without dead band. Top-right: VV curve with dead band (VVwD). Bottom: VW curve
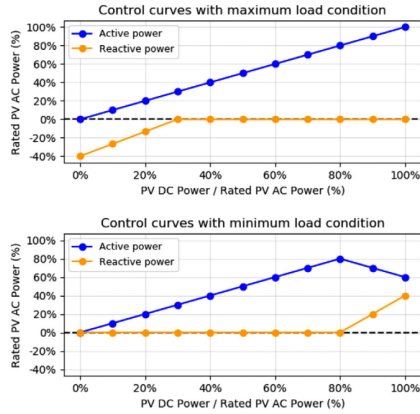
FIGURE 3   Example of the control curves output by the optimization algorithm and used by OptGC. These are the splines for one PV system resulting from one optimization step of the algorithm. The top and bottom figure shows results for a run with maximum and minimum load conditions, respectively. Positive values of reactive power correspond to inductive power and negative values are capacitive power

power. The optimization outputs are computed every 15 minutes and consist of the functions of optimal active power and reactive power output versus PV active power output capability (based on dc power output from solar irradiance) for every PV system and for 11 PV active power output levels from 0% to 100% (Figure 3). The optimization explores a linear space (or spline) [10] discretized in equal steps (in this case, 11 points

from 0% to 100% PV power output, both included). Each point of the spline is called knot. It has to be mentioned that, although the optimization is run for a discretized linear space, the control algorithm uses the whole spline to find the actual values i.e. it interpolates the values between the knots. This behaviour is better explained in Section 5.

The optimization ensures that all knots satisfy the constraints of the optimization problem. The constraints of the problem are handled via large penalty functions. Otherwise, every individual must be individually evaluated for constraint compliance before being included in the model, which will lead to additional computational costs for finding individuals that satisfy the constraints. The use of penalty functions means that if a constraint is not satisfied, a large penalty value is added to the objective function to rank the solution as suboptimal by the optimization algorithm. To drive the optimization algorithm to regions of the solution space closer to the true optima, the penalty for the constraint violation is obtained using distance functions. These distance functions are defined for each constraint and measure how far each solution is from the optimal region as explained further below.

The first constraint states that, for each PV inverter p, the total complex power output cannot exceed the rated apparent power $|S_p^R|$ of the PV inverter at each knot k:

$$\left(P_{p,k}^{\mathrm{PV}} - P_{p,k}^{\mathrm{curt}}\right)^2 + \left(Q_{p,k}^{\mathrm{PV}}\right)^2 \leq \left(\left|S_p^R\right|\right)^2 \qquad (1)$$

where $P_{p,k}^{\mathrm{PV}}$ is the actual PV active power produced by the PV system; $P_{p,k}^{\mathrm{curt}}$ is the curtailed active power; $Q_{p,k}^{\mathrm{PV}}$ is the reactive power absorption/injection.

Let $c_{p,k}^S$ be a distance function to the feasible region for each PV system p at each power output level k. The distance function for this constraint is a summation of all the $c_{p,k}^S$ multiplied by a

160

penalty value, p_v, large enough to dominate over the rest of the elements of the objective function if the constraint is violated. If the constraint is satisfied by a particular knot, the distance value of that knot is removed from the scaled distance function $d_{SR}(\bar{x})$ using the max operator

$$c_{p,k}^S = \left(P_{p,k}^{\text{PV}} - P_{p,k}^c\right)^2 + \left(Q_{p,k}^{\text{PV}}\right)^2 - \left(\left|S_p^R\right|\right)^2 \tag{2}$$

$$d_{SR}(\bar{x}) = p_v \cdot \sum_p^{n_{PV_s}} \sum_k^{n_{kts}} \max\left\{c_{p,k}^S, 0\right\} \tag{3}$$

where $\bar{x}$ is a vector representing one individual or solution generated by the EDA.

Another obvious constraint is that the curtailed active power of the PV system p at each power output k cannot exceed the actual produced PV power at the same p and k

$$P_{p,k}^{\text{curt}} \leq P_{p,k}^{\text{PV}}. \tag{4}$$

Again, the distance function is a summation of the linear function multiplied by a large penalty value

$$d_{P^{\text{curt}}}(\bar{x}) = p_v \cdot \sum_k^{n_{kts}} \max\left\{\left(P_{p,k}^{\text{curt}} - P_{p,k}^{\text{PV}}\right), 0\right\}. \tag{5}$$

The most important constraints (and the most difficult to satisfy) involve voltage limits. The voltage magnitude $|V_{n,k}|$ of each node n of the circuit at every power output level k has to remain within the normative limits $|V^{\text{min}}|$ and $|V^{\text{max}}|$ at all times. In this study, the voltage limits defined in the ANSI C84.1 norm ($|V^{\text{min}}| = 0.95$ p.u. and $|V^{\text{max}}| = 1.05$ p.u.) are used

$$\left|V^{\text{min}}\right| \leq \left|V_{n,k}\right| \leq \left|V^{\text{max}}\right|. \tag{6}$$

To evaluate the voltages that each solution generates, a power flow should be run. However, running 11×1000 (the number of candidate solutions generated by the EDA times the number of knots in the spline) power flows for each generation of the optimization routine would be computationally infeasible. Instead, a linearization of the power flow is used to estimate the voltage magnitude at every node, consistent with [10, 30, 31]

$$\left|V_{i,k}(\bar{x})\right| = \left|V_i^{\text{base}}\right| + \sum_{n=1}^{n_{\text{nodes}}} \left(s_{i,n}^P \cdot \left(P_{p,k}^{\text{PV}} - P_{p,k}^{\text{curt}}\right)\right.$$
$$\left. + s_{i,n}^Q \cdot \left(Q_{p,k}^{\text{PV}}\right)\right). \tag{7}$$

In this formula, $|V_{i,k}(\bar{x})|$ is the voltage magnitude in node i at knot k for the solution $\bar{x}$. $|V_i^{\text{base}}|$ is the voltage magnitude of node i excluding the effect of PV systems, and $s_{i,n}^P$ and $s_{i,n}^Q$ are respectively the sensitivity of voltage magnitude at node i to active and reactive power injections/absorptions at node n. Sensitivity matrices are obtained using the perturbation-observation
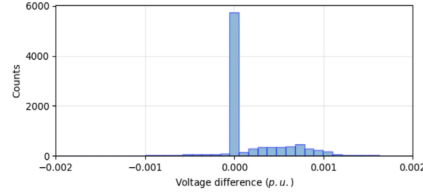
FIGURE 4 Histogram showing the distribution of voltage differences between the linearized power flow and the iterative Newton-Raphson method power flow returned by OpenDSS

method, by making a small change in network state and measuring the effect of that change [32].

A comparison between the linear power flow and the traditional iterative Newton-Raphson method shows that the linearization errors are small (Figure 4). Most of the voltages from the linearized power flow are within 10^{-4} p.u. compared to the power flow performed with the iterative method. All errors are less than 2×10^{-3} p.u.

The distance functions for the voltage constraints are the following slightly modified linear functions:

$$c_k^{V\text{min}} = V^{\text{min}} - \min\left(\left|\overline{V_k(x)}\right|\right) \tag{8}$$

$$d_{V\text{min}}(\bar{x}) = p_v \cdot \sum_k^{n_{\text{knots}}} \max\left\{\left(c_k^{V\text{min}}\right), 0\right\} \tag{9}$$

$$c_k^{V\text{max}} = \max\left(\left|\overline{V_k(x)}\right|\right) - V^{\text{max}} \tag{10}$$

$$d_{V\text{max}}(\bar{x}) = p_v \cdot \sum_k^{n_{\text{knots}}} \max\left\{\left(c_k^{V\text{max}}\right), 0\right\}. \tag{11}$$

The objective problem is finally given by

$$\min_{Q_{p,i}^{\text{PV}}, P_{p,i}^{\text{curt}}} \sum_{k=1}^{n_{\text{knots}}} \sum_{p=1}^{n_{PV_s}} \left(w\left(Q_{p,k}^{\text{PV}}\right)^2 + (1-w)\, P_{p,k}^{\text{curt}}\right) \tag{12}$$

$$\text{s.t.} \left(P_{p,k}^{\text{PV}} - P_{p,k}^{\text{curt}}\right)^2 + \left(Q_{p,k}^{\text{PV}}\right)^2 \leq \left(\left|S_p^{\text{Rated}}\right|\right)^2, \tag{1}$$

$$P_{p,k}^{\text{curt}} \leq P_{p,k}^{\text{PV}}, \tag{4}$$

$$\left|V^{\text{min}}\right| \leq \left|V_{n,k}\right| \leq \left|V^{\text{max}}\right|, \tag{6}$$

where w is a weighting factor which has been set to 0.01 to favour the use of reactive power over active power curtailment.

The dimensionality of the problem increases with the number of nodes in the circuit (n_{nodes}), the number of PV systems (n_{PV_s}) and the number of knots (n_{knots}). All individuals in the population have ($2 \times n_{\text{knots}} \times n_{PV_s}$) variables, where the 2 comes from active and reactive power.

4 | ESTIMATION OF DISTRIBUTION ALGORITHM IMPLEMENTATION

Unlike other EC algorithms, EDA relies on the construction of a probabilistic model with the best individuals from the current generation to sample the population of the next generation. This is based on the reasonable assumption that good solutions for a determined problem should have a similar structure [19]. The general pseudo-code of an EDA is shown in Algorithm 1.

Algorithm 1. General EDA pseudo-code

1    $P_0 \leftarrow$ Generate a random initial population P_0, of λ individuals $\overline{x} = (x_1, x_2, x_3, \dots, x_v)$, where v is the number of variables.

For $g = 1, 2, 3 \dots$ in number of generations:

2    $P_{g-1}^{Se} \leftarrow$ Evaluate population and sort them according to their fitness value and select Se best individuals

3    $p_g(x) = p(x P_{g-1}^{Se}) \leftarrow$ Build the probabilistic model: estimate the probability distribution of an individual being among the selected individuals

4    $P_g \leftarrow$ Sample a new population from $p_g(x)$

EDAs are assumption-free approximators, allowing them to be directly implemented in a wide variety of problems without the complex tuning. EDAs provide flexibility in constructing the statistical distributions, which can be tuned to fit a specific problem. Unlike many other EC algorithms, EDAs have very few parameters that need to be adjusted, although the complexity of the model will always depend on the complexity of the problem being solved.

EDAs can be classified according to the relationships and interdependencies between their input variables. EDAs can have either independent variables or variables with multivariate dependencies [18]. It is reasonable to assume that in our optimization problem, the variables may have multivariate dependencies, since power injections or absorptions in one phase have an effect on the other phases. However, variables corresponding to different knots should be independent. It is worth mentioning that the probabilistic model is only used for searching the solutions of the optimization problem. The grid is modelled in a deterministic manner, where the load, the irradiance curves for the PV systems and the structure of the grid are predetermined. The only variables allowed to vary are those defined for the optimization problem: the active power curtailed and the reactive power injection/absorption of each PV system.

To determine whether univariate or multivariate algorithms are more suitable for the proposed problem, estimation of multivariate normal algorithm (EMNA$_{global}$) and univariate marginal distribution algorithm for continuous domains (UMDA$_c$) (both from [18]) were compared. An in-depth review and taxonomy showing the advantages and disadvantages of the different EDA strategies are presented in [33]. Generally, UMDA$_c$ ignore feature dependencies and have worse performance for deceptive problems (a class of challenging problems which usually mislead the search to some local optima rather than the global optimum [34]). But on the other side,

they are the fastest among the other EDAs, they are well suited for high dimensionality problems and are easily scalable. When implementing EMNA$_{global}$ to our problem, the construction of the covariance matrix was adapted to capture just the relationships between variables belonging to each knot. In other words, 11 different covariance matrices (one for each knot) are constructed. In UMDA$_c$ variables are treated independently, so a different distribution is computed for each variable at each generation.

Algorithm 2. UMDA$_c$ for the optimization problem

1    $P_0 \leftarrow$ Generate a random initial population P, of individuals $\overline{x} = (x_1, x_2, x_3, \dots, x_n)$, where n is the number of variables ($n_{knots} \times 2n_{PVs}$), according to a multivariate normal distribution with mean μ_0 and covariance matrix Σ_0 such as:

$$X \sim \mathcal{N}(\mu_0, \Sigma_0), \ \mu_0 = \mathbf{0}$$
$$\Sigma_0 = \begin{bmatrix} \sigma_{0,i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{0,j}^2 \end{bmatrix}$$
$$\sigma_0^2 = large\ constant\ value$$

For $g = 1, 2, 3 \dots$ in number of generations:

2    $P_{g-1}^{Se} \leftarrow$ Evaluate population, sort them according to their fitness value and select Se best individuals

3    $p_g(x) = p(x P_{g-1}^{Se}) = \mathcal{N}(\mu_g, \Sigma_g) \leftarrow$ Build the probabilistic model: estimate the probability distribution of an individual being among the selected individuals, where the mean and the covariance matrix are computed as follows:

$$\mu_g = \bar{X} = \frac{1}{N} \sum_{r=1}^{N} x_{i,r}, \quad i = 1, \dots, n$$
$$\Sigma_g = \begin{bmatrix} \sigma_{i,i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{j,j}^2 \end{bmatrix}$$
$$\sigma_{i,i}^2 = \frac{1}{N} \sum_{r=1}^{N} (x_{i,r} - \bar{X}_i)^2 \quad i = 1, \dots, n$$

4    $P_g \leftarrow$ Sample a new population from $p_g(x)$

Taking dependencies into account (EMNA$_{global}$), increased the computational expense of computing the covariance matrices, but did not reduce the cost function compared to UMDA$_c$. One would expect that taking dependencies into account, while increasing computational expenses, would reduce the number of iterations of the optimizer, since the algorithm would build better distributions to sample from. But EMNA$_{global}$ neither reduced the number of iterations nor improved the quality of the solution. Therefore, the UMDA$_c$ algorithm was implemented to solve the optimization problem. The pseudo-code for our implementation of the UMDA$_c$ is shown in Algorithm 2. When sampling the first generation, a large value for the variance of each variable σ_0^2 is set to favour the exploration of the solution space.

5 | SIMULATIONS

A modified version of the IEEE123 node test feeder was selected to perform the analysis. The IEEE123 node test feeder is a medium-size system with multiple voltage regulators and
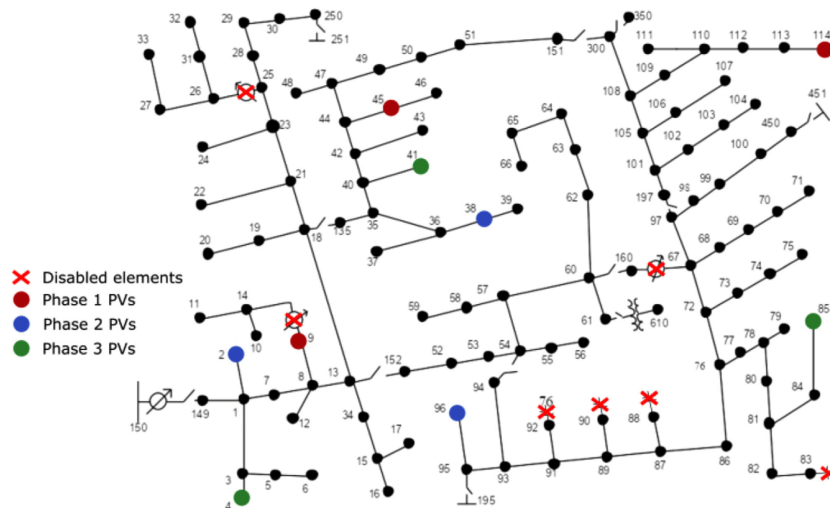
FIGURE 5 Layout of the modified IEEE123 circuit used for this study. Nine PV systems were added to the circuit and all voltage regulators and shunt capacitors were disabled

shunt capacitors, and a nominal voltage of 4.16 kV. It is characterized by overhead and underground lines with different X/R ratios, and considerable imbalance. IEEE123 is a well behaved-circuit with little convergence problems, and it was recommended for the evaluation of new control strategies [35].

The main modifications consist of disconnecting shunt capacitors (nodes 83, 88, 90 and 92) and voltage regulators (lines 9–14, 25–26 and 67–160), and adding PV systems as shown in Figure 5. The modifications allow evaluating if the control provided by the PV inverters is sufficient to overcome voltage problems without the help of conventional equipment. The main effect of the proposed modifications is an overall decrease of the line voltages, especially pronounced in the last nodes of the feeder. This fact leads to uncommon large undervoltages in situations of maximum load. The 5% and 95% quantiles of the line voltages shift from 0.956 to 1.037 with shunt capacitors and voltage regulators to 0.929 to 0.990 without shunt capacitors and voltage regulators.

The 9 PV systems were evenly distributed among the three phases and located at the beginning, middle, and end of the line for each phase. The rated power of each system was selected at 1,000 kVA for the total PV rated power be comparable to the total load of the feeder. To evaluate the proposed optimization heuristic in a realistic scenario where the active power injections are non-uniform among the phases (as a result of the control) and will affect the balance of the phases, the PV systems are modelled with single-phase inverters. The peak loads for phase 1, 2 and 3 are respectively 1,420 kW, 915 kW, and 1,155 kW. The

PV penetration level is 225% given

$$PV_{level} = \frac{S_{PV}}{S_{load}} \quad (13)$$

where S_{PV} is the total rated apparent PV power and S_{load} is the total peak apparent power of the loads.

The optimization problem is solved for 6 different scenarios designed to be representative of the range of situations the optimization algorithm may encounter during actual operation. These 6 scenarios are a combination of 3 irradiance curves (clear, partly cloudy and overcast conditions) and 2 load curves (high and low demand periods) as shown in Figure 6. Although the analysis ignores seasonal changes or week days versus weekend days, these 6 scenarios were considered to be enough to evaluate the performance and robustness of the algorithm.

For each scenario, a power flow is performed every minute for the whole day (1440 minutes, Figure 7). To reduce the computational cost, the optimization algorithm calculates the splines (Figure 3) only every 15 minutes. For the next 15 minutes, the operating points (active and reactive power output) are selected from the spline according to the actual PV power output of each PV system. Assuming that points lying between the calculated knots also satisfy the constraints, the actual operational points of the PV inverters are linearly interpolated between the knots and yield the actual power flow.
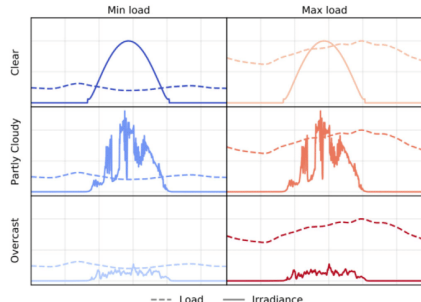
163

FIGURE 6 Normalized load and irradiance curves selected for each of the six proposed scenarios. Dashed lines represent load curves, and solid lines represent irradiance curves
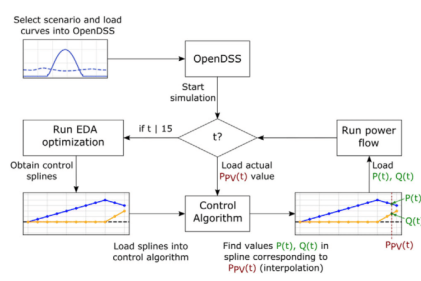

FIGURE 7 Selection process of the operational points for each PV inverter at each time step the simulation is run

The power flows at each time step are solved using OpenDSS. Python scripts automate the process and the optimization routine. The proposed EDA was coded and implemented by the authors from scratch for reasons of flexibility and control over the different aspects of the algorithm

The results of the EDA-based OptGC are compared to six local control strategies: (i) without control (NC), (ii) fixed PF (PF0.98), (iii) VV function with the first curve in Figure 2 (VV), (iv) VV function with a dead band following the second curve in Figure 2, (VVwD), (v) VW function with the third curve in Figure 2, (VW) and (vi) a combined control of VVwD and VW (VVwD+VW).

6 | RESULTS

The main results of the simulations for all the scenarios are summarized in Figures 8–10 and Figures 11–13 for the minimum and maximum load, respectively. In the following sections, those results are analyzed in depth.
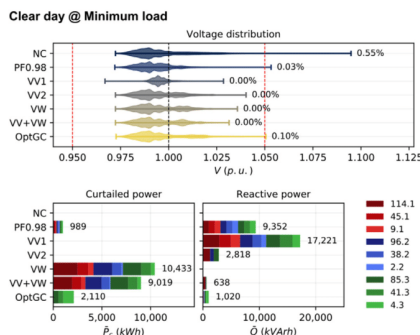

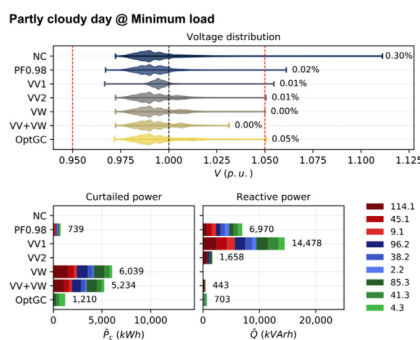**FIGURE 8** Results for clear day at minimum load conditions

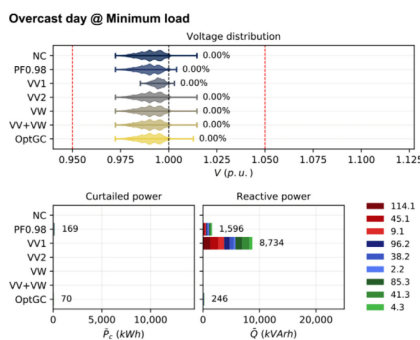**FIGURE 9** Results for partly cloudy day at minimum load conditions

FIGURE 10 Results for overcast day at minimum load conditions
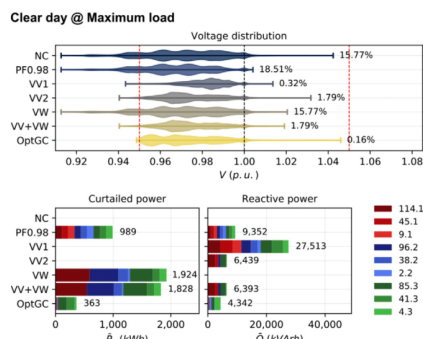
164

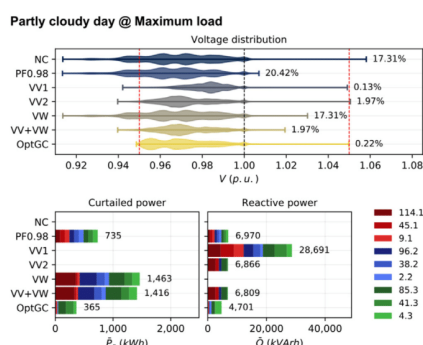FIGURE 11   Results for clear day at maximum load conditions

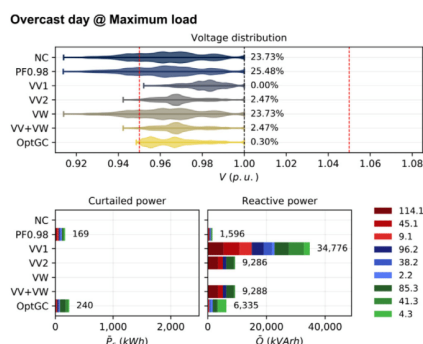FIGURE 12   Results for partly cloudy day at maximum load conditions

FIGURE 13   Results for overcast day at maximum load conditions

6.1 | Minimum load scenarios

When studying solar power integration the most common problems are overvoltages during low load and high PV power injection [36]. As the controls in the substation have been disabled—the OLTC are set to operate at 1 p.u. voltages—the minimum load conditions are found to be less problematic. In a different set-up with the OLTC enabled, the minimum load scenarios would be more problematic. One of our objectives was to evaluate if the PV inverters were able to maintain the voltage within the acceptable limits without support from any other device in the grid, including capacitor banks or OLTC.

The results for the minimum load scenarios with clear, partly cloudy and overcast conditions are shown in Figures 8–10, respectively. In those figures, the top plot represents the distribution of voltage magnitudes throughout the day, where the whiskers of the violin plots include 100% of the values. Bottom left and bottom right plots represent active power curtailment and reactive power used for control, respectively. Different colours represent different phases (red is phase 1, blue is phase 2 and green is phase 3). Different tones of the colours represent different PV systems where darker colours represent PV systems located at a larger distance to the substation. The legend indicates the bus number and the phase number.

In general, OptGC reduces reactive power use and active power curtailment from the PV inverters compared to the local control strategies while maintaining the voltage within the limits most of the time. In minimal load conditions, OptGC operates outside the voltage limits 0.02% of the total time. These voltage violation events are caused by variations in load within the 15 minutes when the same spline is used in the controller. Since the control is designed to only consider changes in PV power output, it can handle solar variations across the full range of PV power output. But large changes in load may lead to situations not considered by the optimization algorithm, causing voltage violations.

Regarding active power curtailment, while the naïve PF0.98 curtails less active power than OptGC, it comes at the price of larger voltage deviations. Regarding reactive power compensation, the OptGC outperforms every other control strategy in terms of amount of reactive energy absorbed/injected, while the percentage of times outside the voltage limits remains similar. VV, VVwD, and PF0.98, all operate 0.01% of the time outside the limits, while VVwD+VW always operates within the limits; albeit VVwD+VW curtails 4.2 times more energy than OptGC (14,253 kWh in total across the three solar scenarios).

Another interesting finding is that OptGC makes a slightly unbalanced use of power curtailment and reactive power compensation in comparison to local control methods. OptGC actuates control preferentially on phase 1 instead of spreading it equally among all the phases. This behaviour is a result of neutral point shifting, where the reactive power injections/absorptions in one phase affect the voltage in the other phases.

For all the controls, the PV systems at the end of the feeder are more active in regulation (darker colours or higher numbers in the legend in Figures 8–10). This is expected, since nodes

located at the end of the feeder are more prone to suffer voltage violations.

The non-zero values of power curtailment and reactive power use for OptGC in overcast condition are a consequence of the nature of the optimization algorithm which should be considered as an approximator. Near-optimal solutions may lead to small, but unnecessary control actions, such as power curtailment.

6.2 | Maximum load scenarios

Although high demand periods are less problematic in PV integration studies, due to the disconnection of the OLTC at the substation and the capacitors banks, the largest voltage violations in our study occur under maximum load conditions.

The results for the maximum load scenarios with clear, partly cloudy and overcast conditions are shown in Figures 11–13, respectively. The structure and colour code of the figures are the same as for minimum load results.

OptGC outperforms every other alternative in terms of reactive power use (except for the PF0.98 control on the overcast day) with a percentage of bad voltage values of 0.2%. Since PF0.98 uses reactive power as a function of the actual PV power delivered, it is expected that on days with low irradiance, PF0.98 will use little reactive power. But PF0.98 fails to maintain the voltages within the allowed range; violations occur 21% of the time. VVwD+VW and VVwD are close in terms of reactive power usage and voltage stability, but inferior to OptGC. VVwD+VW and VVwD use 46% and 47% more reactive power than OptGC, respectively; while the percentages of time outside the voltage limits are 2% for both. For VV control, the amount of reactive power is 5.9 times higher than OptGC and voltages are violated 0.1% of the time. VV is the only control that achieves better voltage control than OptGC, but at the cost of a disproportionate amount of reactive power.

As expected, the curtailed energy in maximum load scenarios is considerably lower than in the minimum load scenarios (up to 10,433 kWh), with maximum values below 2,000 kWh. OptGC again outperforms all the methods that make use of power curtailment—except for PF0.98 on the overcast day. Again, there is a slight residual active power curtailment on the overcast day that can be explained by the algorithm as an approximator.

Practically all local control strategies violate the lower voltage limits as observed in the violin plots in Figures 11–13. The severe undervoltage situations experienced in maximum load scenarios are explained by the modifications made to the test feeder (explained in Section 5), primarily the removal of shunt capacitors. OptGC yields a high density of voltage magnitudes close to the lower limit. Since the objective function does not reward voltages being a safe distance away from the lower limit, small variations in the load can force voltage violations which occur 0.2% of the time. There is a tradeoff between minimizing power used for control and the proximity of voltages to the limit.

PV systems located at the end of the feeder use more reactive power compensation and power curtailment. For OptGC, the
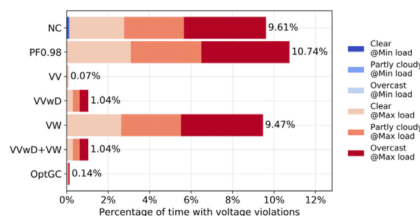

FIGURE 14 Percentage of time with voltage violations (out of voltage limits) split by scenario (each colour represents one scenario). Red colours represent the high load scenarios, where most of the voltage violations take place
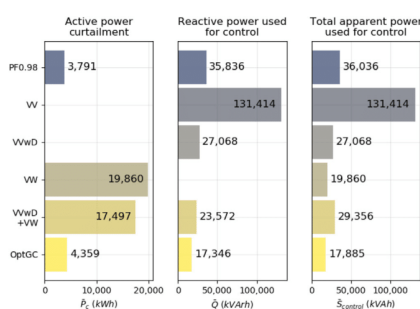
FIGURE 15 Total active power curtailment, reactive and apparent energy used for control in all scenarios

participation of PV systems in the control is unbalanced, with systems in phases 2 and 3 suffering more power curtailment, and systems in phases 1 and 3 using more reactive power compensation. This unbalanced nature of global controllers is due to their ability to inject or absorb reactive power in one phase to avoid voltage violations in the other phases.

6.3 | Summary metrics

Figure 14 summarizes the voltage violation results. VV results in the least voltage violations at 0.07% of the time (or 1,669 minutes). Although the amount of time of OptGC doubles with 0.14% of the time (3096 minutes) outside the voltage limits, both figures are far superior compared to the other strategies. The next controls in the ranking are the VVwD and the VVwD+VW, which are outside of the limits 1.04% of the time. The remaining strategies violate voltages 10% of the time.

Figure 15 shows the total amount of reactive power and active power curtailed for each strategy for all scenarios, as well as the total apparent power. Since some strategies only use reactive power compensation, some rely only on active power curtailment, and some use both, ranking the different controls is

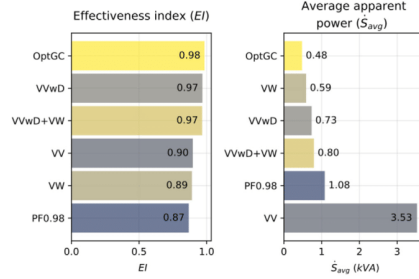**FIGURE 16** Performance indicators for each control, sorted from best (top) to worst (bottom). Left: EI. Right: average apparent power ($\hat{S}_{avg}$)

difficult. We proposed two new metrics to comprehensively evaluate the performance of the controls by linking the total energy used for control (both curtailed active power and reactive power) with the resulting voltages over all scenarios.

The first indicator, $\hat{S}_{avg}$, estimates the average apparent power needed for successful voltage regulation

$$\hat{S}_{avg} = \frac{S_{control}}{t_{V\,good}} \qquad (14)$$

where $S_{control}$ is the total amount of apparent power used by each control in kVAh and $t_{V\,good}$ is the number of hours with voltages within the accepted range achieved with that same control. $S_{control}$ includes the curtailed active power (P_{curt}) and the reactive power (Q) in all scenarios as:

$$S_{control} = \sqrt{P_{curt}^{2} + |Q|^{2}}. \qquad (15)$$

$\hat{S}_{avg}$ has units of kVA and evaluates the apparent power expended to achieve an hour of voltages in the permitted range. A smaller S_{avg} indicates a better performance.

The second indicator, the effectiveness index (EI), quantifies both how effective the control is relative to the usage of apparent power, and its success in voltage regulation. EI is defined as

$$\text{EI} = \left(1 - \frac{S_{control}}{S_{PV}} \right) \left(\frac{t_{V\,good}}{t_{total}} \right) \qquad (16)$$

where S_{PV} is the total available apparent power (i.e. the sum of the inverter rated powers), and t_{total} is the total number of hours evaluated (6 scenarios x 24 hours). The EI ranges from 0 to 1 and a value of 1 indicates maximum effectiveness.

Figure 16 shows the EI and the average apparent power for each strategy. OptGC scores the best for both indicators. Therefore, (i) OptGC makes the most efficient use of apparent power to obtain voltages within the allowable limits; (ii) OptGC is the most effective control method in terms of how voltage viola-

tions are avoided through judicious use of reactive and active power.

6.4 | Benchmarking

In this section, we compare the performance of EDA to two well-known metaheuristics commonly used in optimization problems: GA and particle swarm optimization (PSO).

GA mimics the natural selection by selecting the best individuals in a population for reproduction, generating variance at each generation through crossover and mutation operations. Common GA hyper parameters that need to be tuned are the mutation probability, the crossover probability, the type of crossover, the population size and number of generations. We implemented a GA using Python programming language. Hyper parameters selected were: a mutation probability of 0.5, a crossover probability of 1 (there is always crossover between all selected candidates), a population size of 1,000 (same as EDA) and a number of generations of 150 without any additional stopping criterion nor tolerance. Regarding the type of crossover, as the individuals or vectors representing the solutions are structured in a very specific manner for our problem (i.e. each vector encodes the curtailed power and reactive power for each knot in a defined order), the crossover operation must preserve this order to avoid the exploration of useless areas in the solution space and hence satisfy the constraints. For that reason, a position-based crossover operation [37] is preferred instead of the classic k-point crossover which splits the individuals at some k arbitrary points.

PSO differs in that there is an initial population (swarm) of individuals (particles), but they are not selected for reproduction. Instead, the particles move through the solution space and their position is evaluated at each iteration (iteration is the equivalent of "generation" in EDA and GA). The movement of the particles is initialized randomly. As the solution evolves, the movement for every iteration is based on each particle's best position reached (cognitive coefficient c_1) and global best of all particles reached (social coefficient c_2). The main hyper parameters that need to be tuned for PSO are those related to the movement, and can be observed in Equations (17) and (18)

$$x_i^{t+1} = x_i^t + v_i^{t+1} \qquad (17)$$

$$v_i^{t+1} = \omega \cdot v_i + c_1 \cdot u_1 \cdot \left(x_i^{best} - x_i^t \right)$$
$$+ c_2 \cdot u_2 \cdot \left(x_g^{best} - x_i^t \right) \qquad (18)$$

where x_i^t is the current particle position, v_i^{t+1} is the particle's movement, ω is the inertia of the particle, v_i is the initial velocity of the particle, u_1 and u_2 are uniformly distributed random variables within [0, 1], x_i^{best} is the best position attained by the particle, and x_g^{best} is the global best position attained by all particles. We implemented a generic PSO algorithm using Python, with $\omega = 0.01$, $c_1 = 0.7$, and $c_2 = 0.5$. These hyper parameters were selected empirically by testing a different set of values and evaluating compliance with the problem constraints.
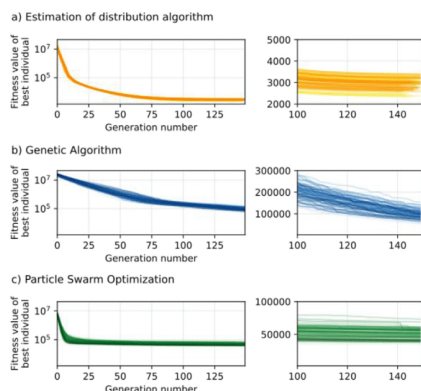
FIGURE 17   Convergence plot for all the metaheuristics implemented for all optimization runs in the clear day scenario under minimum load conditions. Each line represents one optimization run. There are 96 runs (every 15 minutes for 24 hours) during the day

FIGURE 18   Benchmarking results for clear day at minimum load conditions

FIGURE 19   Performance indicators for the benchmarking analysis of the clear day at minimum load conditions scenario

As the main analysis identified that two scenarios are particularly interesting to study, we benchmarked EDA against GA and PSO for minimum load with clear sky conditions and maximum load with overcast conditions. To allow for a fair comparison between the methods, the same number of generations/iterations (150) was defined for each method.

Figure 17 shows the convergences of the three methods for the clear sky day with minimum load. EDA converges rapidly over the first generations, reaching a plateau after generation 100, and it achieves a smaller fitness value than PSO and GA.

GA has a more stable decay without reaching a plateau, indicating that more generations may be needed to improve the fitness. This stable, but slow decrease of the fitness value is consistent with how new individuals are generated in the GA, which is characterized by randomness in its operation. In contrast, EDA effectively reduces the variance by sampling from a normal distribution constructed by the best individuals, actively searching for the best distribution possible.

PSO converges quickly to a local minimum in the first generations, but it is a factor of 20 larger than the minimum attained by EDA. A closer look at the convergence plot (Figure 17, c, right) reveals sudden steps down in some runs in the last iterations for a few runs. This could mean an insufficient number of iterations.

The results of each heuristic in terms of voltage violations, active power curtailed and usage of reactive power, for the clear day with minimum load conditions are shown in Figure 18. The indicators described in the previous section (EI and $\hat{S}_{avg}$) for each method are depicted in Figure 19. It can be seen that, although the other methods obtain better results in terms of voltage violations, this comes at the cost of a 10 times more curtailed active power and 37 times more reactive power usage.
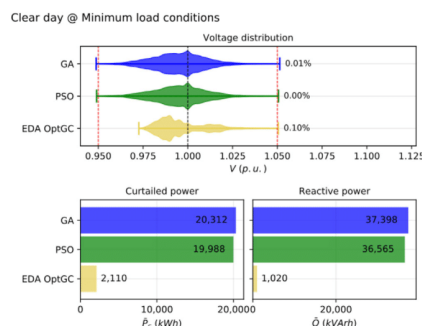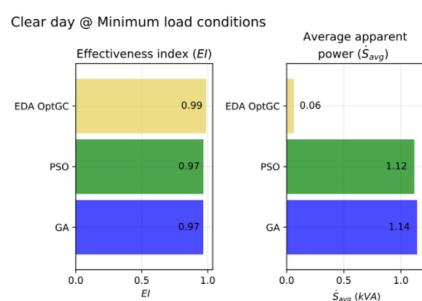
The small improvement for GA and PSO in the voltages violations (less than a 0.1%) does not justify the excessive increase of curtailed power and reactive power usage as shown by the indicators in Figure 19, especially the average consumption of apparent power.

The results are similar for the overcast day with maximum load conditions. The voltage and power results are shown in Figure 20 and the indicators are shown in Figure 21. As it can be seen in Figure 20, again the voltage violations are improved for GA and PSO, but this comes at the expense of an increase in reactive power used by over a factor 7. Once again, EDA scores the best in both indicators, showing that it makes better use of reactive power and power curtailment.

7 | DISCUSSION

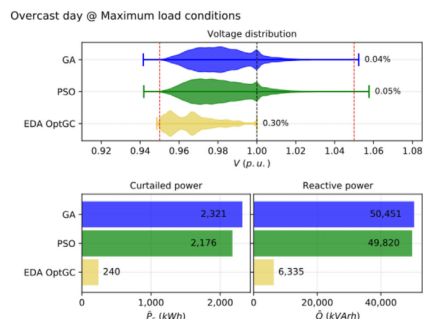The proposed EDA-based optimization has been proven to be an excellent approximator for optimizing voltage regulation

168

Overcast day @ Maximum load conditions

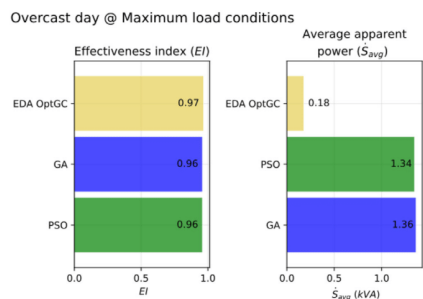**FIGURE 20**   Benchmarking results for overcast day at maximum load conditions

FIGURE 21   Performance indicators for the benchmarking analysis of the overcast day at maximum load conditions scenario

with PV inverters, yielding suitable solutions in a reasonable time. Although, the EDA leads to some suboptimal solutions (cases where there are residual values of P_c and Q even if they are not needed), the reduction in use of power dedicated to voltage regulation compared with other local control strategies is significant.

For the IEEE123 circuit and the specific setup of PV inverters, PV inverters were able to accomplish all the voltage regulation. Although there is a minimal amount of time where the voltage is outside the ANSI limits, these situations may be overcome with different configurations of the optimization algorithm (for example, setting more restrictive voltage limits) or if the Watt-Watt and Watt-var control curves were updated at every time step (instead of updating them every 15 minutes).

Voltage violations mostly result from solar or load variability within the 15-minute updates of the control curves. Solar variability can be significant over time scales of a few minutes. The control curves prescribe changes in the PV Watt and var

output as a function of PV active power. Therefore, solar variability directly affects the control curves and does so at every control time step (1 minute). But the coordination of the control curves across inverters will no longer be optimal given that solar variability causes changes in the state of the grid since the most recent 15-minute control curve time step. Therefore, voltage violations—or more generally suboptimal voltage control—may result from solar variability.

Some voltage violations in the OptGC strategy are caused by the inability of the optimization problem to take into account the load variability. Again, running the optimization algorithm every minute would solve most of these voltage violations. But short-term uncontrolled voltage variability is common in electric distribution systems; load tap changers, which are common voltage control devices, are not capable of handling rapid voltage fluctuations either. Load tap changers only act after the voltage has exceeded the threshold for a certain amount of time, typically 10s of seconds to a minute. The control delay associated with load tap changers suggests that fast load fluctuations are not that common or important in distribution grid operations.

Other findings are aligned with the results expected from the literature. First, there is a clear positive correlation between the distance to the substation and the amount of power used for control purposes [8]. This is an expected behaviour because voltages at the end of the feeder differ more from the substation voltage. Second, in three-phase four-wire configurations, the actuations on one phase affects the voltage of other phases, due to the shifting of the neutral point [30]. OptGC takes advantage of neutral point shifting by optimizing the control to act more on some phases instead of even control among all the phases. Third, the EDA performs well in this problem despite the large number of variables. Since EDAs extract global statistical information from the populations they are immune to potential local minimal, making them good candidates to solve power systems optimization problems [19].

The EDA performance was benchmarked against other state-of-art metaheuristics (GA and PSO). EDA outperformed both methods, in terms of the quality of the solutions. This matches other results in the literature [19], [20], where the more active search strategy of EDAs led to better solutions in less iterations.

This study could be extended as follows:

(i) Using different strategies in the EDA algorithm (e.g. efficiently building a model with multivariate dependencies). Although the EDA-based strategy still needs to be evaluated with a larger number of PV systems in the grid, based on our preliminary analysis we expect a similar performance if the same strategy without variable interdependencies (UMNA) is used.

(ii) Extending the optimization problem to consider variations in load. The solutions input to the global control would then have an extra dimension (that is, they would be planes instead of splines) to represent load variation.

(iii) Adding more PV systems on the grid; or targeting runs for specific load and solar PV value based on probabilistic forecasts; and reducing the computational cost, e.g. by varying

the number of generations being created at each optimization run.

(iv) Unrelated to EDA one could optimize the VV control curves, since it performs quite well in most scenarios.

8 | CONCLUSION

In this study, an EDA-based optimized global voltage control for PV systems in a three-phase four-wire distribution grid was evaluated. The objective was to find out if PV inverters were able to maintain voltages within ANSI limits without any other regulation device on the grid while minimizing the amount of active power curtailed and the use of reactive power. To evaluate the performance of the optimization algorithm, six scenarios with different load and irradiance conditions were analyzed, and the results were compared to other standard local controls.

The OptGC outperformed every other alternative in terms of apparent energy dedicated to control, averaging 0.48 kVAh per hour of good voltage values. In terms of voltage quality, only the VV local control (VV) performs slightly better (with voltage violations 0.07% of the time versus 0.14% for the OptGC), but with a less efficient use of the energy based on the EIs.

The viability of using PV inverters as the only source of voltage regulation has been demonstrated, at least for the IEEE123 test feeder. The main limitations encountered by the strategy are related to the formulation of the optimization problem, which does not take into account variations in load conditions, which lead to some undervoltages.

The suitability of using EDA as approximators in power system related optimization problems with a high number of variables was demonstrated. EDA provides quality solutions in a reasonable time.

REFERENCES
1. Stetz, T., et al.: Technical and economical assessment of voltage control strategies in distribution grids: Assessment of voltage control strategies in distribution grids. Prog. Photovolt.: Res. Appl. 21(6), 1292–1307 (2013) https://doi.org/10.1002/pip.2331.
2. Molina-Garcia, A., et al.: Reactive power flow control for PV inverters voltage support in LV distribution networks. IEEE Trans. Smart Grid 8(1), 447–456 (2017) https://doi.org/10.1109/TSG.2016.2625314.
3. Barbieri, F., et al.: Intrahour cloud tracking based on probability hypothesis density filtering. IEEE Trans. Sustain. Energy 9(1), 340–349 (2018), https://doi.org/10.1109/TSTE.2017.2733258.
4. Tang, J., et al.: Optimal configuration of battery energy storage systems using for rooftop residential photovoltaic to improve voltage profile of distributed network. J. Eng. 2019(16), 728–732 (2019) https://doi.org/10.1049/joe.2018.8386.
5. Braun, M., et al.: Is the distribution grid ready to accept large-scale photovoltaic deployment? State of the art, progress, and future prospects. Prog. Photovolt. Res. Appl. 20(6), 681–697 (2012) https://doi.org/10.1002/pip.1204.
6. Alizadeh, S.M., et al.: The impact of X/R ratio on voltage stability in a distribution network penetrated by wind farms. In 2016 Australasian Universities Power Engineering Conference (AUPEC), Brisbane, Australia, Sep. 2016, pp. 1–6, https://doi.org/10.1109/AUPEC.2016.7749289
7. Ochi, T., et al.: The development and the application of fast decoupled load flow method for distribution systems with high R/X ratios lines. In 2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT), Feb. 2013, pp. 1–6, https://doi.org/10.1109/ISGT.2013.6497842
8. Demirok, E., et al.: Local reactive power control methods for overvoltage prevention of distributed solar inverters in low-voltage grids. IEEE J. Photovolt. 1(2), 174–182 (2011) https://doi.org/10.1109/JPHOTOV.2011.2174821.
9. O'Connell, A., Keane, A.: Volt–var curves for photovoltaic inverters in distribution systems. IET Gener., Transmiss. Distrib. 11(3), 730–739 (2017) https://doi.org/10.1049/iet-grd.2016.0409
10. Weckx, S., et al.: Combined Central and Local Active and Reactive Power Control of PV Inverters. IEEE Trans. Sustain. Energy 5(3), 776–784 (2014) https://doi.org/10.1109/TSTE.2014.2300934.
11. Martí, P., et al.: Distributed reactive power control methods to avoid voltage rise in grid-connected photovoltaic power generation systems. In 2013 IEEE International Symposium on Industrial Electronics, May 2013, pp. 1–6, https://doi.org/10.1109/ISIE.2013.6563803
12. Ma, R., et al.: Smart grid communication: Its challenges and opportunities. IEEE Trans. Smart Grid 4(1), 36–46 (2013) https://doi.org/10.1109/TSG.2012.2225851.
13. Masoum, M.A.S., et al.: Optimal placement of hybrid PV-wind systems using genetic algorithm. In 2010 Innovative Smart Grid Technologies (ISGT), Jan. 2010, pp. 1–5, https://doi.org/10.1109/ISGT.2010.5434746
14. Liu, L., et al.: Research on short-term optimization for integrated hydro-PV power system based on genetic algorithm. Energy Procedia 152, 1097–1102 (2018) https://doi.org/10.1016/j.egypro.2018.09.132.
15. L. jing Hu, et al.: Capacity optimization of wind /PV/storage power system based on simulated annealing-particle swarm optimization. In 2018 37th Chinese Control Conference (CCC), Jul. 2018, pp. 2222–2227, doi: 10.23919/ChiCC.2018.8482706
16. Mohammed, O.H., et al.: Particle swarm optimization of a hybrid wind/tidal/PV/battery energy system. Application to a remote area in Bretagne, France. Energy Procedia 162, 87–96 (2019) https://doi.org/10.1016/j.egypro.2019.04.010.
17. Gómez-Lorente, D., et al.: Evolutionary algorithms for the design of grid-connected PV-systems. Expert Syst. Appl. 39(9), 8086–8094 (2012) https://doi.org/10.1016/j.eswa.2012.01.159.
18. In: P. Larrañaga, J. A. Lozano (eds. Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation. Springer, New York, (2002)
19. Su, W., Chow, M.-Y.: Performance evaluation of an EDA-based large-scale plug-in hybrid electric vehicle charging algorithm. IEEE Trans. Smart Grid 3(1), 308–315 (2012) https://doi.org/10.1109/TSG.2011.2151888.
20. Qiu, Y., et al.: Network optimization based on genetic algorithm and estimation of distribution algorithm. In 2008 International Conference on Computer Science and Software Engineering, Dec. 2008, pp. 1058–1061, https://doi.org/10.1109/CSSE.2008.1511
21. Pelikan, M.: Analysis of estimation of distribution algorithms and genetic algorithms on NK landscapes. In Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation - GECCO '08, Atlanta, GA, USA, 2008, p. 1033, https://doi.org/10.1145/1389095.1389287
22. Li, H., et al.: EDA-based speed control of a networked DC motor system with time delays and packet losses. IEEE Trans. Ind. Electron. 56(5), 1727–1735 (2009) https://doi.org/10.1109/TIE.2009.2013749.
23. Wang, L., et al.: Solving system-level synthesis problem by a multi-objective estimation of distribution algorithm. Expert Syst. Appl. 41(5), 2496–2513 (2014) https://doi.org/10.1016/j.eswa.2013.09.049.

170

24. Zheng, H., Wang, L.: Reduction of carbon emissions and project makespan by a Pareto-based estimation of distribution algorithm. Int. J. Prod. Econ. 164, 421–432 (2015) https://doi.org/10.1016/j.ijpe.2014.12.010.

25. Wang, Z., Gong, M.: Dynamic deployment optimization of near space communication system using a novel estimation of distribution algorithm. Appl. Soft Comput. 78, 569–582 (2019) https://doi.org/10.1016/j.asoc.2019.02.045.

26. Wu, C., Wang, L.: A multi-model estimation of distribution algorithm for energy efficient scheduling under cloud computing system. J. Parallel Distrib. Comput. 117, 63–72 (2018) https://doi.org/10.1016/j.jpdc.2018.02.009.

27. Rafati Fard, M., Shariat Mohaymany, A.: A copula-based estimation of distribution algorithm for calibration of microscopic traffic models. Transp. Res. Part C: Emerg. Technol. 98, 449–470 (2019) https://doi.org/10.1016/j.trc.2018.12.008.

28. Common Functions for Smart Inverters: 4th ed.. EPRI, Palo Alto, CA, 2016. 3002008217

29. Seuss, J., et al.: Analysis of PV advanced inverter functions and setpoints under time series simulation. SAND2016-4856, 1259558, May 2016. https://doi.org/10.2172/1259558

30. Weckx, S., et al.: Voltage sensitivity analysis of a laboratory distribution grid with incomplete data. IEEE Trans. Smart Grid 6(3), 1271–1280 (2015) https://doi.org/10.1109/TSG.2014.2380642.

31. Wang, Z., et al.: Robust optimization based optimal DG placement in microgrids. IEEE Trans. Smart Grid 5(5), 2173–2182 (2014), https://doi.org/10.1109/TSG.2014.2321748.

32. Tamp, F., Ciufo, P.: A sensitivity analysis toolkit for the simplification of MV distribution network voltage management. IEEE Trans. Smart Grid 5(2), 559–568 (2014) https://doi.org/10.1109/TSG.2014.2300146.

33. Armañanzas, R., et al.: A review of estimation of distribution algorithms in bioinformatics. BioData Min 1, 6 (2008) https://doi.org/10.1186/1756-0381-1-6.

34. Chen, Y., et al.: Solving deceptive problems using a genetic algorithm with reserve selection. In 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence), Jun. 2008, pp. 884–889, https://doi.org/10.1109/CEC.2008.4630900

35. Schneider, K.P., et al.: Analytic considerations and design basis for the IEEE distribution test feeders. IEEE Trans. Power Syst. 33(3), 3181–3188 (2018) https://doi.org/10.1109/TPWRS.2017.2760011.

36. Dall'Anese, E., et al.: Optimal dispatch of photovoltaic inverters in residential distribution systems. IEEE Trans. Sustain. Energy 5(2), 487–497 (2014) https://doi.org/10.1109/TSTE.2013.2292828.

37. Larranaga, P., et al.: Learning Bayesian network structures by searching for the best ordering with genetic algorithms. IEEE Trans. Syst., Man, Cybern. Part A: Syst. Hum. 26(4), 487–493 (1996) https://doi.org/10.1109/3468.508827.

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

172

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
              UNIVERSIDAD DE LA LAGUNA

              María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
              UNIVERSIDAD DE LA LAGUNA

Universidad de La Laguna
Oficina de Sede Electrónica
Entrada
Nº registro: 2021/44847
Nº reg. oficina: OF002/2021/44186
Fecha: 02/06/2021 14:57:08

173

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
Su autenticidad puede ser contrastada en la siguiente dirección https://sede.ull.es/validacion/

Identificador del documento: 3479426   Código de verificación: 0B+OCdAL

Firmado por: David Cañadillas Ramallo   Fecha: 02/06/2021 14:51:11
             UNIVERSIDAD DE LA LAGUNA

             María de las Maravillas Aguiar Aguilar   07/06/2021 16:06:22
             UNIVERSIDAD DE LA LAGUNA