Ocean Engineering 189 (2019) 106314

Contents lists available at ScienceDirect

Ocean Engineering

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

# Combining random forests and physics-based models to forecast the electricity generated by ocean waves: A case study of the Mutriku wave farm

Paula Serras [a,*], Gabriel Ibarra-Berastegi [a,b], Jon Sáenz [b,c], Alain Ulazia [d]

[a] NE & Fluid Mechanics Dept, University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU), Bilbao, Spain
[b] Joint Research Unit, Spanish Oceanographic Institute- University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU), Plentziako Itsas Estazioa, Plentzia, Spain
[c] Dept. of Applied Physics II, University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU), Leioa, Spain
[d] NE & Fluid Mechanics Dept, University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU), Eibar, Spain

ARTICLE INFO

*Keywords:*
Mutriku wave farm
Electric power forecasting
Random forest
Machine learning
Fluid mechanics

ABSTRACT

This paper combines random forests with physics-based models to forecast the electricity output of the Mutriku wave farm on the Bay of Biscay. The period analysed was 2014–2016, and the forecast horizon was 24 h in 4-h steps. The Random Forest (RF) machine-learning technique was used, with three sets of inputs: i) the electricity generated at Mutriku, ii) the wave energy flux (WEF) prediction made by the ECMWF wave model at Mutriku's nearest gridpoint, and iii) ocean and atmospheric data for the Bay of Biscay. For this last input, extended empirical orthogonal functions (EOFs) were calculated to reduce the dimensionality of these data, while retaining most of the information. The forecasts are evaluated using the R-Squared, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The model easily outperforms a persistence forecast at 8–10 h and beyond. The most accurate forecasts are achieved by using all three of these inputs. This approach may help to effectively integrate wave farms into the electricity market.

## 1. Introduction

### 1.1. Wave energy development

The sea and oceans are a potential source of renewable energy. Marine energy, which includes both offshore wind and waves among others, could enable the European Union (EU) to reach its environmental goals, as well as record major economic growth. In fact, the Commission's Blue Growth Strategy recently singled out the ocean energy sector as one of the emerging areas in the 'blue economy' (Blue Energy, 2014), and major investments and efforts are currently in place, especially throughout Europe. The ocean energy sector includes different technologies, such as wave energy, tidal energy, salinity gradient, and ocean thermal energy conversion (OTEC), whose main aim is to transform the power from seas and oceans into low-polluting energy (O'Hagan et al., 2016). According to "The 2018 annual economic report on the EU blue economy" (Blue Economy, 2018), the EU is the principal developer of ocean energy technology, aiming to install 100 GW of ocean energy capacity by 2050, thus creating more than 400,000 jobs.

Thus far, waves and tidal energy are the most mature ocean energy technologies, and they are expected to evolve to commercial status in the medium term. By the end of 2013, ocean energy capacity was about 530 MW (World Energy Outlook, 2015), with most of this coming from tidal power. The majority of the tidal power projects currently in operation are located in Korea, Canada and France, reaching capacities of up to 254 MW in the case of Korea. Smaller pre-commercial wave and tidal power projects (in the order of 11 MW) have gained increasing relevance, especially in United Kingdom (Santos et al., 2017). The EU is one of the main developers of ocean energy technology. In terms of the companies involved in the development of converters, devices and technology in general, 50% focus on the field of wave energy and 45% on tidal. Moreover, most of the ocean energy test and implementation sites are located on European coasts (Magagna and Uihlein, 2015).

A higher percentage of the world's electricity could be generated from ocean-based energy sources, so there is a clear need to provide an economic incentive, promoting the use of this type of renewable energy

* Corresponding author.
*E-mail address:* paula.serras@ehu.eus (P. Serras).

https://doi.org/10.1016/j.oceaneng.2019.106314
Received 4 February 2019; Received in revised form 22 July 2019; Accepted 9 August 2019
Available online 26 August 2019

(Esteban and Leary, 2012).

The characterisation and mapping of ocean energy resources, including the identification of areas with high wave energy and its quantification, constitute a necessary first step for the development of ocean energy. Additionally, it is important to accurately estimate available wave energy sources with a high spatial and temporal resolution. The simulation and forecasting of parameters such as wave energy flux or electric power is therefore crucial for optimising the design of wave energy converters (WECs), as well as predicting the electricity available to the grid (Reikard et al., 2011, 2017; Pinson et al., 2012; Parkinson et al., 2015; Widén et al., 2015).

To date, the Mutriku wave farm on the Bay of Biscay (Fig. 1) is the only one in the world continuously supplying electricity to the grid (O'Hagan et al., 2016; Ibarra-Berastegi et al., 2018). The energy production data used in this work have been provided by the Basque Energy Agency EVE (*Ente Vasco de la Energía* - www.eve.eus), making them available for research, and thus solving one of the main limitations when trying to obtain more realistic wave energy forecasting results (Reikard, 2017). Mutriku has 14 oscillating water column (OWC) operating turbines, and supplies energy to the grid 74.4% of the time. One of the more significant parameters recorded at the wave farm is the average power generated by each turbine over the preceding 5 min, leading to an energy output of 246 469 kWh per year in the 2014–2016 period under analysis. The aim of this paper is therefore to predict the power output at Mutriku with a 24-h forecast lead time.

### 1.2. Forecasting electricity production

The participation of renewable energy sources in electricity market mechanisms may incur economic sanctions if any imbalances are detected between generation and supply (Kaur et al., 2016; Mazzi and Pinson, 2017). This issue has received increasing attention in the past, especially in the field of solar and wind power generation, where some authors (Esteban et al., 2010, 2012) propose the use of energy storage devices (e.g., batteries or hydrogen) to solve the intermittency problems associated with renewable energies. Surplus electricity would therefore be stored during high-production periods, and then be made available for subsequent peak demand periods.

In order to reduce the high levels of uncertainty around renewable energy production, information needs to be predicted for better decision-making processes in electricity markets. Thus, a reliable forecasting method will lead to a reduction in integration cost, lower average annual operating costs, and fewer penalties for mismatches between the power contracted and its effective delivery, all of which will impact directly on the economy (Botterud, 2017; Notton et al., 2018).

Short-term forecasting (ranging from a few hours to a few days) is normally used for planning and controlling economic load supplies, as well as for competitive negotiation in electricity markets. Long-term forecasts (ranging from a few days to a week) are generally used for operations that need forward information from renewable sources in the long-term, such as energy storage approaches or maintenance programming of both renewable and conventional power plants (Taşcıkaraoğlu, 2017). All forecasting strategies, whether short or long term, contribute to a more accurate, economic and reliable modus operandi. The accuracy of the power forecast is a crucial factor for the system's optimum performance because both under- and over-forecasting incur additional costs for offsetting the differences between estimation and reality.

In the particular case of wave energy, a particular WEC's electricity production at a given moment and location is driven mainly by local waves. Hence, building cost-effective strategies and models for the short-term prediction of electricity production at a wave farm can largely benefit from the experience, methods and strategies developed for forecasting wave energy (Ibarra-Berastegi et al., 2015a,2016).

Statistical techniques and physics-based models are commonly used for wave forecasting. The WAM wave model, included in the Integrated Forecast System (IFS) of the European Centre for Medium-range Weather Forecasts (ECMWF), is an example of a wave forecasting system based on physical models. Statistical techniques include Random Forest (RF), regression-based, neural networks, and genetic programming. In general, statistical models are more accurate for short time horizons of 3–18 h, (e.g., for utilities operating wave farms), while physics-based models are better predictors for longer horizons of 18–20 h and beyond (e.g., ocean conditions for sailing). (Reikard and Rogers, 2011; Uihlein and Magagna, 2016). The combination of both techniques, statistical and physics-based models, is also analysed by Reikard et al. (2011), concluding that this way they yield more accurate predictions than when used individually.

Spain does not have any specific price regulations on the electricity generated at wave farms. However, there is a more general regulation applicable to all the electricity generated at facilities based on emerging renewable technologies, as in the case of Mutriku. The prices of the electricity generated at these facilities are not currently informed by the unregulated interaction of the players in a free market, as they are fixed by law for a given period of time, after which the applicable regulations may change (BOE, 2017).

The experience gained by the authors through building short-term forecasting models for precipitation (Fernández-Ferrero et al., 2009, 2010) or the wave energy flux in the Bay of Biscay (Ibarra-Berastegi et al., 2015a, 2016) has now been applied to the short-term prediction of the electricity output at Mutriku. The approach described in this paper is a further development of our previous studies in this field (i.e., the use of the RF technique, among others, for predicting a target variable). Additionally, based on the data analysed at Mutriku during the 2014–2016 period, we now go one step further by using Mutriku's output as the predicted variable. Log records of sea states in the Bay of Biscay and future electricity production are mathematically related using the machine-learning RF technique. Due to the highly non-linear relationships to be expected in operations, RF is a most appropriate option for developing the aforementioned forecasting models. The objective of this paper is therefore to build a set of RF-based models to forecast the electricity generated at the Mutriku with forecast lead times of up to 24 h.

The models shown here are fed with current and past observations corresponding to the power generated by the turbines in Mutriku, as well as with wide-area oceanic and atmospheric information from

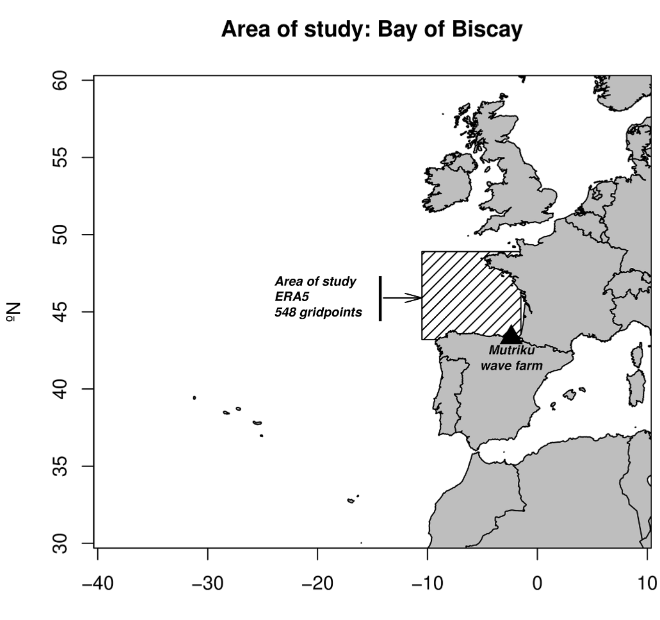


**Fig. 1.** Area of study and location of the Mutriku wave farm.

ECMWF reanalyses. However, in a first step, a reduction in the dimensionality will be applied using extended EOFs. This is undertaken with two goals in mind. The first one is to reduce the spurious spatial correlations common in spatial meteorological fields. As the information at all the grid points is not effectively independent, it is common to use EOFs involving information that is closer to the effective number of degrees of freedom in model output fields (Bretherton et al., 1999; Wang and Shen, 1999). The second goal is related to the effective use of combined EOFs from several variables to handle all the information involved in the forecast. Additionally, the RF model's performance will be evaluated and compared to the most obvious and simplest model: Persistence of electric power. Persistence assumes that what we are currently observing will also be replicated exactly over the fothcoming hours, and is considered here as a benchmark that can be used to compare the forecasting ability of alternative models, as it is the simplest forecasting tool available.

The originality of this work lies in the combination of RF as a statistical technique and the use of predictions made by physics-based models as inputs for RF models. This method has been tested for Mutriku, and has proven to predict electric power effectively.

In the following sections, the data and methodology used will first be explained. The results obtained with both the RF model and Persistence will be compared and discussed, and finally, the more salient conclusions will be presented. The main lines of future research will also be discussed.

## 2. Data and methodology

### 2.1. Data

Data from two different sources have been used:

1. Data from the European Centre for Medium-range Weather Forecast (ECMWF) corresponding to the ERA5 reanalysis (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5); hourly data for the 2014–2016 period and the Bay of Biscay [48.9° N, 349.5 °E 43.2°N 358.5°E] have been used (Fig. 1). ERA5 has a resolution of $0.3 \times 0.3°$, so 548 sea gridpoints were initially used for this study. However, a preliminary screening of these gridpoints led to the conclusion that using only one out of every nine available (61) was enough due to the high spatial correlation of the information contained in them, which led to a low number of independent degrees of freedom.

These are the same variables and geographical environment used in our previous studies, although now, the more recent ERA5 reanalysis has been used as the input source instead of ERA-Interim. As is the case in the field of wind energy (Olauson, 2018) ERA5's higher time and space resolution will probably make it the benchmark for wave energy studies over the coming years. Furthermore, ERA-Interim is soon to be discontinued, so state-of-the-art forecasting systems need to be built using the new ERA5 reanalysis.

The following variables of interest have been chosen: significant wave height (Hs), mean wave period (Tm), mean wave direction (mwd), mean sea-level pressure (MSLP), and finally, zonal and meridional components of wind speed at 10 m (U10, V10). The wave energy flux (WEF [kW/m]) at all the gridpoints has been calculated under the hypothesis of deep waters (Multon, 2012; Lavidas and Venugopal, 2018) according to eq. (2).

$$\mathrm{WEF} = \rho g^2 \mathrm{Hs}^2 \mathrm{Tm}/64\Pi \sim 0.489 \mathrm{Hs}^2 \mathrm{Tm} \quad (2)$$

Where $\rho \sim 1025\,\mathrm{kg/m^3}$ is the density of sea water (variable), and $g \sim 9.8\,\mathrm{m/s^2}$ is gravitational acceleration, whose theoretical value varies according to latitude. The combination of the WEF values and the mean wave direction (mwd) at each gridpoint allows calculating the zonal (WEFu) and meridional (WEFv) components of this WEF.

All these variables have been used for the whole area in analysis mode, although ERA5 also provides forecasts at hourly steps for the same variables at each gridpoint up to a horizon of 18 h ahead at 06:00 and 18:00 GMT.

As this paper's objective is to forecast the electric power at Mutriku up to 24 h ahead, and as electricity production is driven by ocean waves, ERA5 forecasts of Hs and Tm have also been used at the nearest gridpoint (21 km) to Mutriku (Fig. 2) at 4-8-12-16 and 18 h ahead. As a first step, Hs and Tm have been combined using eq. (2) to calculate the WEF forecast at 4-8-12-16 and 18 h ahead. The idea is that the ERA5 forecasts of forthcoming sea states may also be used to feed the models and provide more accurate predictions of the electricity generated.

2. Hourly data from the Mutriku SCADA computer data-recording system corresponding to the same period (2014–2016). This is the initial period of systematic data gathering and validation, and most available cases have been recorded only every 4 h. Predictions have therefore been calculated at 4-h steps from 4 to 24 h ahead.

From the total of 14 Wells-type turbines using OWC technology installed, not all of them are continually generating electricity. On average, and due to maintenance work or experimental purposes, only around 10 turbines are normally running at any given moment. The electricity produced may also vary from one turbine to another due to the irregular seabed and the spatial configuration of Mutriku's breakwater. More details on the general operational characteristics of Mutriku can be found elsewhere (Ibarra-Berastegi et al., 2018).

Although the instantaneous power generated by each turbine is regularly recorded, its value may be due to the action of a particular wave that is not necessarily representative of the current sea state. A more reasonable indicator of electric power has therefore been chosen for current and previous studies, namely, the average power generated by each turbine over the previous 5 min. As the yield of the operating turbines may differ, the average power generated across the active turbines over the previous 5 min has been used as a reference of overall electric power. This will be the target variable to be predicted up to 24 h ahead, and it will be referred to here simply as the electricity generated at Mutriku.

After a strict data screening and pre-processing stage, a total of 2855 hourly cases were available for building the models. The first 1427 were

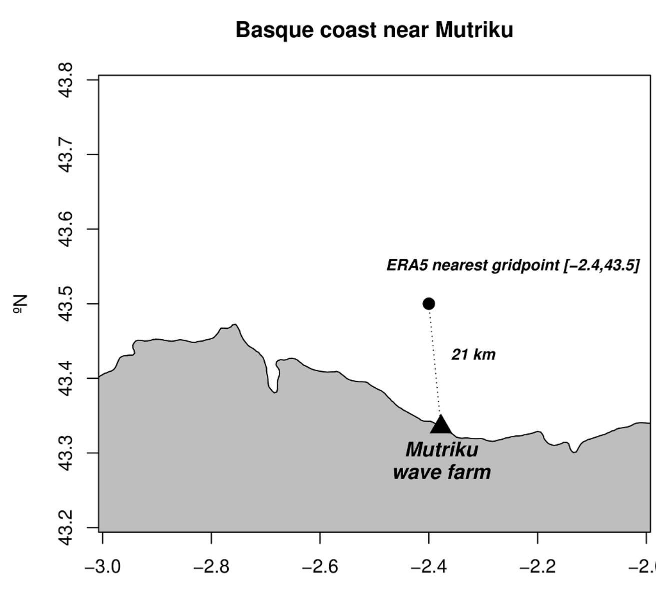


**Fig. 2.** Nearest ERA5 gridpoint to Mutriku.

used to train the models, and the last 1428 to test them. All the results on model performance shown in this paper have been obtained exclusively from the 1428 cases corresponding to the test database.

### 2.2. Methodology

The general aim here was to use RF-based statistical models to forecast the electricity generated by waves, using the data recorded at Mutriku over the 2014–2016 period. The results obtained were compared to the ones provided by Persistence. The maximum forecasting lead time has been set at 24 h. Within this forecast lead time, the RF predictions have been made up to 24 h ahead at 4-h time steps.

#### 2.2.1. Building the models

An RF model has been built for each forecast lead time. In each model, the output or target variable at a given time t is the average power generated through the active turbines at Mutriku over the previous 5 min t+h (h = 4,8, 12, 16, 20, 24). This means a total of six models have been built because, as mentioned above, in most cases data from Mutriku's SCADA are available only every 4 h.

The inputs for all the models have been as follows:

i) The observed electric power generated at time t
ii) The ERA5 wave energy flux forecasts at t+k (k = 4, 8, 12, 16 and 18) at the nearest gridpoint to Mutriku (Fig. 2).
iii) The ERA5 data for the Bay of Biscay using observations (analyses) not only at t but also at t-6, t-12, t-18 and t-24 to capture the daily evolution of the sea states in the area. Twenty inputs have been used after a phase of dimensionality reduction via extended EOF (see below).

#### 2.2.2. Extended empirical orthogonal functions

When it was used as a tool for statistical weather forecasting at the end of the 1940s (North, 1984; Monahan et al., 2009), the principal component analysis technique was referred to as Empirical Orthogonal Functions (EOFs). The method seeks the optimal decomposition of a field that is a function of time and space into a set of patterns that are only a function of space (EOFs) and associated time indices (Principal Components, PCs). EOFs, followed by a sensible truncation of the number of PCs retained, allow optimally reducing the dimensionality of the system to a smaller number of PCs that permit depicting the system's dynamics in a space with fewer dimensions. The use of EOFs for the analysis of spatiotemporal modes is frequently challenged (Dommenget and Latif, 2002; Monahan et al., 2009). Conversely, their use as a dimensionality reduction technique is commonly accepted (Hannachi et al., 2007). The Extended EOF (ExtEOF) technique is oriented toward the identification of temporally varying patterns in multivariate datasets (Weare and Nasstrom, 1982), designed to deal with both the spatial and temporal correlations observed in the data. This method is normally used with variables with a strong temporal autocorrelation (Hannachi et al., 2007; Ibarra-Berastegi et al., 2011; Fukutomi and Yasunari, 2013).

After variable standardisation, 20 ExtEOFs were used in this study, with at least 92% of the total variance being retained. Accordingly, instead of working with 61 highly correlated gridpoints x seven variables (Hs, Tm, MSL,U10,V10, WEFu, WEFv) x five time steps (t, t-6, t-12, t-18, t-24) = 2 135 variables, we only use 20 inputs retaining most of the relevant information, with the additional advantage that the variables introduced into the next step of the statistical model are uncorrelated by construction.

#### 2.2.3. Random forests

The need to deal with big data in many fields of science has led to the development of a whole family of algorithms commonly known as machine learning. Among them, one of the techniques being increasingly used in the field of geophysical fluids involves Random Forests (RFs). A RF has the double advantage of its rather simple implementation and that its outputs tend to be easy to interpret (Ahmad et al., 2017).

RFs can be used either for classification or for regression purposes, and can also yield results in probabilistic mode. In regression mode, RFs perform well at modelling highly nonlinear relationships between a set of inputs and an output. Mathematically speaking, RFs can be considered a further development of the classification and regression trees (CART) that have the additional advantage over other algorithms, such as the neural networks intended to deal with nonlinearities, of always being free from overfitting.

A RF can be understood as the result of multiple random perturbations applied to a regression tree (see Fig. 3), whereby a high number of perturbated trees are obtained. This group of slightly different trees constitutes a forest. In regression mode, the final output given by an RF is simply the average of the output that all the individual regression trees would yield.

The trees are perturbated by randomly selecting both the input cases and the number of $m$ variables at the different split nodes of the regression tree (Jiang et al., 2009). In fact, $m$ is the most important parameter to be selected in an RF, with the second one being the number of perturbated trees -usually more than 1000- that form the forest. Several recommendations can be found in the literature regarding the selection of $m$ (Breiman, 2001; Liaw and Wiener, 2002; Gromping, 2009; Siroky, 2009).

More specific aspects of the practical implementation of the RF technique applied to the prediction of WEF have largely been explained by our research group in journals and conference papers (Ibarra-Berastegi et al., 2015).

The six RF models that forecast the electricity generated at t+h use 26 inputs (current electric power at t, ERA5 WEF forecasts at t+k (k = 4, 8, 12, 16, 18) + 20 ExtEOF). Therefore, following the general recommendations for RF implementation in all cases (Breiman, 2001; Liaw and Wiener, 2002; Gromping, 2009; Siroky, 2009), each forest contained 1000 tress, and the $m$ parameter was set at $26/3 \approx 9$.

Additionally, the models' performance has been compared to Persistence. The Persistence of electricity generated is the simplest and cheapest forecast that can be made, as it is always possible to predict a value that matches the present value. In the particular case of the production at Mutriku recorded by the plant's electricity meter, the slowly decaying autocorrelation observed with values of 0.6 at 24-h lags (Ibarra-Berastegi et al., 2018) suggests that Persistence seems to perform fairly well in the short term. In order to confirm that the RF prediction model is a good forecasting tool, the indicators used for the analysis of the differences between real and predicted data must be better for the prediction model than for Persistence. This latter is not a prediction model per se, but only a bottom line or reference to be taken into account, as any modelling effort will only be justified if it can outperform Persistence.

### 2.3. Evaluation of models

After fitting the models with data from the training dataset for the different forecasting horizons, performance has been assessed here by comparing the forecasts provided by Persistence and the RF models with observations belonging exclusively to the test dataset. The three statistical indicators selected are those commonly used in wave energy studies.

i) R-Squared, because it provides a general idea of the variance fraction explained by each model.
ii) Mean Absolute Error (MAE), because the target variable is kw, and an absolute estimation of the error is illustrative both in engineering terms and for developing future wave farm management strategies.
iii) Nevertheless, percentage errors are also interesting, as they are independent of the scale, and hence different datasets can be easily compared. In further developments of this study, model

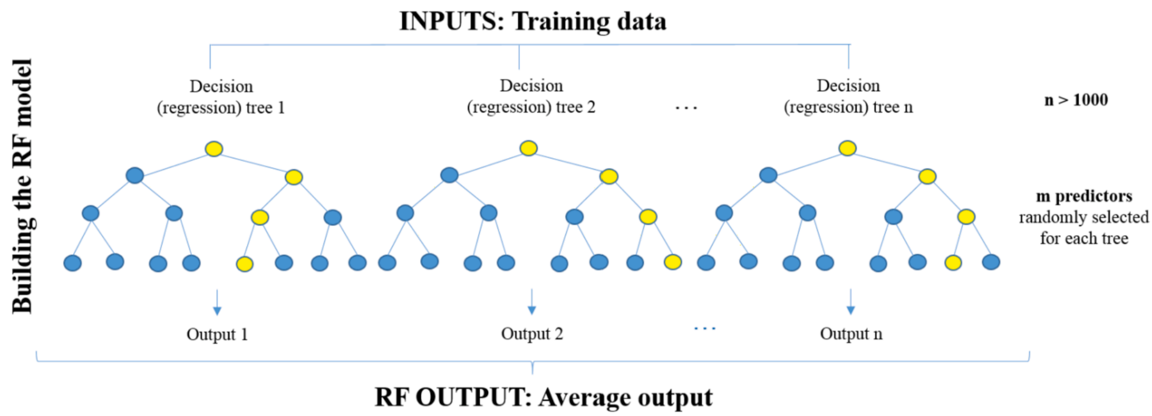


**Fig. 3.** Random Forest flow chart.

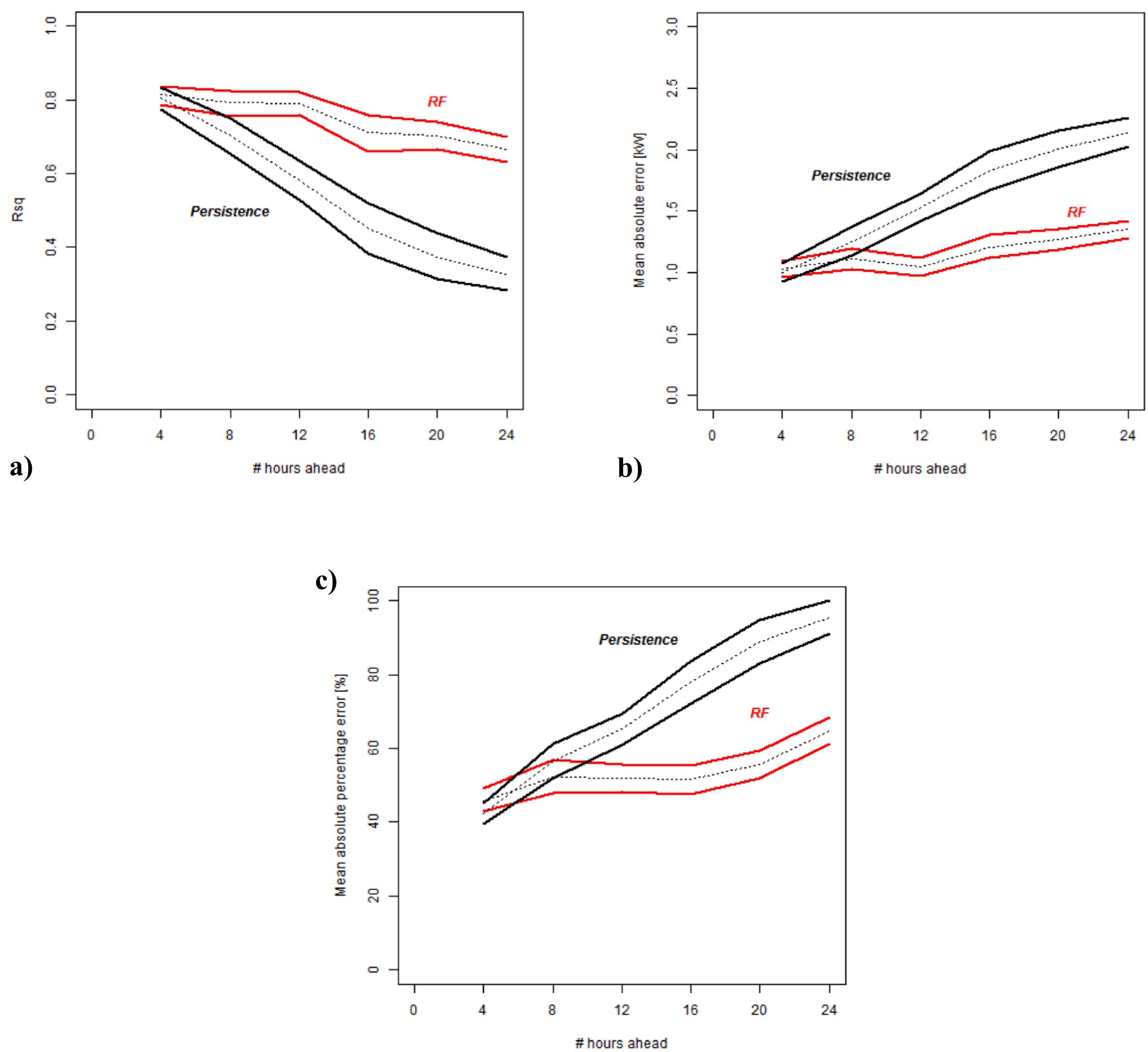


**Fig. 4.** a) R-Squared (RSQ), b) Mean Absolute Error (MAE), and c) Mean Absolute Percentage Error (MAPE) for Mutriku's electric power forecast (95% confidence boundaries). Persistence vs. Random Forests.

performance as described by such a scale-independent indicator may also be helpful for electricity auctions and, ultimately, for a better integration of wave energy into the electricity grid. The most commonly used measure is the Mean Absolute Percentage Error (MAPE). It is defined as the average value of [Abs(log of forecast-log of actual)] × 100. This indicator is a derivation of the MAE, but instead of raw values of observations and predictions, natural logarithmic values are used, thus providing an estimation of the error in percentage terms. It is therefore possible to rescale the differences in the data ranges, and obtain a homogeneous comparison based on the proportional error (%) for each model.

To correctly evaluate RF models and Persistence performance for the different forecasting scenarios, the confidence ranges of these indicators have been calculated here by bootstrap resampling (5000 resamples), whereby model performance has been evaluated at a 95% confidence level, according to these three indicators.

R language (R Core Team, 2018) was used to evaluate the data here, and the principal packages used for the calculations were as follows:

i) 'randomForest' (Liaw and Wiener, 2002)
ii) 'FactoMineR' for the calculation of EOF (Le et al., 2008)
iii) 'sp', 'maps' and 'rgdal' for creating and managing maps (Pebesma and Bivand, 2005; Bivand et al., 2013, 2018; Becker et al.,2018).

## 3. Results and discussion

### 3.1. Results

Fig. 4a–c shows the R-squared, Mean Absolute Error, and Mean Absolute Percentage Error results, respectively, obtained for both RF and Persistence models, at a 95% confidence level. The results from Fig. 4c show that errors for the RF model fall within the range of 46–65%. All the results indicate that the RF-based models outperform Persistence beyond the 8-10-h horizon, confirming the good performance of the RF model used to predict the electric power at Mutriku. Fig. 4a–c shows it is important to stress that when compared to Persistence, the deterioration in the RF model's performance with the forecasting horizon follows a slow downward trend. The MAPE moves from values of around 50% at t+10–60% at t+24. RF errors correspond to the overprediction of future values for all forecasting horizons. The positive bias detected in our model's error can be related –in general terms- to the fact that the most influential inputs also tend to have a positive bias (results not shown). However, an in-depth study of the error requires many more cases and a systematic analysis depending on the season, electric power, observed WEF, and frequencies.

An interesting aspect of RF is that it allows identifying the most influential input variables by calculating the percentage increase in the mean square error (incMSE) due to the case-shuffling of these variables at the model fitting stage (Gromping, 2009). As the incMSE number decreases, the importance of the input variable analysed also decreases.

An analysis of the input variables (see left-hand column in Table 1) indicates that the most influential input for the RF model at t+4 was the current electric power at time t, followed by the WEF forecasts (referred to as WEF_FCST_variable number in the table) and the extEOFs (referred to as extEOF_variable number in the table), which provide an estimation of the regional atmospheric and oceanic conditions. For t+8 and t+12, the WEF forecasts were followed with almost the same weight by the current electric power, as the main inputs, while the extEOFs played a moderate role. This means that at those forecast ranges, the results from a numerical model of the ocean's state provide substantial information for the short-range forecast. Thus, as the system's atmosphere and ocean state at the time the forecast is issued forgoes any connection with the state of the system at longer lead times, numerical forecasts provide information about the system that corrects these forecasts. From t ≥ 16, the leading factors were the WEF numerical forecasts followed by the general sea state in the Bay of Biscay, as described by the extEOFs, while the relevance of the observed power at time t significantly decreased. This reveals that pure Persistence cannot provide valid information for such long forecast lead times.

The system's memory is playing a significant role in providing the models with predictive capabilities for forecasts up to 12 h ahead. This is consistent with the high autocorrelation shown by the overall electricity production at Mutriku (Ibarra-Berastegi et al., 2018), and stresses the need to compare the RF models' results with Persistence.

### 3.2. Discussion

As mentioned above, this is an initial study to explore whether electricity output can be predicted at a wave farm, and the first one ever applied to an operational facility, namely, Mutriku on the Bay of Biscay.

It should be stressed that the data gathered by the SCADA substantially improve the models' results, particularly at the initial lead times. Thus, any improvement in this feature will lead to better generation forecasts. On the other hand, real-time inputs from the ECMWF can also be obtained for real-time applications for commercial uses, albeit depending on the number of fields, their resolution, and the number of forecast steps that need to be used during the real-time application of the forecasting scheme presented here. A rough estimate of the cost of the input data required ranges from 30 kEUR to 50 kEUR per year, although detailed estimates of this cost must be negotiated in each case with the appropriate national meteorological services, or the ECMWF itself. They must also be refined depending on the spatial and temporal resolution requested from the ECMWF. The results obtained here highlight the possibility of predicting the average electricity generated by the farm's working turbines through the use of RF models. These models outperform Persistence beyond 8–10 h. Thus, if this forecasting method is finally implemented, the actual number of turbines in operation at

**Table 1**
Percentage increase in the mean square error (incMSE) of the inputs at different forecasting horizons (t+4, +8, +12, +16, +20, +24).

| incMSE of inputs | t+4 | t+8 | t+12 | t+16 | t+20 | t+24 |
|---|---|---|---|---|---|---|
| Input 1 | Current electric power at t | WEF_FCST_08 | WEF_FCST_12 | WEF_FCST_16 | WEF_FCST_24 | WEF_FCST_24 |
| incMSE (%) | 38.13 | 30.65 | 31.32 | 24.38 | 27.98 | 34.45 |
| Input 2 | WEF_FCST_04 | Current electric power at t | ExtEOF_02 | WEF_FCST_12 | WEF_FCST_20 | WEF_FCST_20 |
| incMSE (%) | 32.86 | 26.69 | 25.65 | 20.93 | 27.17 | 33.85 |
| Input 3 | WEF_FCST_08 | WEF_FCST_12 | Current electric power at t | WEF_FCST_20 | WEF_FCST_16 | ExtEOF_02 |
| incMSE (%) | 23.35 | 23.4 | 23.22 | 17.59 | 15.68 | 14.68 |
| Input 4 | ExtEOF_02 | WEF_FCST_04 | WEF_FCST_08 | WEF_FCST_24 | ExtEOF_01 | WEF_FCST_04 |
| incMSE (%) | 20.34 | 23.22 | 24.85 | 20.35 | 17.08 | 27.58 |
| Input 5 | WEF_FCST_12 | ExtEOF_02 | WEF_FCST_16 | Current electric power at t | ExtEOF_08 | WEF_FCST_16 |
| incMSE (%) | 15.94 | 17.59 | 21.49 | 16.08 | 16.9 | 24.28 |
| Input 6 | WEF_FCST_20 | WEF_FCST_16 | WEF_FCST_04 | WEF_FCST_04 | ExtEOF_04 | ExtEOF_08 |
| incMSE (%) | 13.2 | 15.68 | 18.48 | 14.36 | 16.12 | 23.04 |
| Input 7 | ExtEOF_08 | ExtEOF_01 | ExtEOF_01 | ExtEOF_01 | WEF_FCST_12 | ExtEOF_05 |
| incMSE (%) | 12.95 | 14.68 | 15.8 | 13.22 | 15.79 | 22.36 |

Mutriku will be known and, assuming a constant value of the electricity predicted during an hour, the total electricity (kWh) generated at Mutriku could also be predicted for the same forecasting horizons as used here.

Electricity production at Mutriku is driven by waves, so it depends on the sea state, and follows a markedly seasonal cycle. The results of this research therefore correspond to an average performance over the period analysed. However, the forecasting results may vary according to the season or sea state, so an analysis of a longer time series for the site is needed for a more detailed characterisation of the prediction of the electricity generated at Mutriku. Mutriku is currently the only wave farm in the world continuously delivering electricity to the grid, and has been doing so now for more than seven years. The models described here are the initial design of a cost-effective forecasting strategy applied to a real-world wave farm, and not only prototypes, simulated wave farms, or standalone devices operating for short periods. Although the results shown here might be highly site-dependent and associated with Mutriku's particular conditions and technology (OWC), they provide a sound methodological framework on how to tackle the issue of the intermittency of wave-generated electricity, and may be applied in other environments, as shown by Reikard et al. in a recent publication (Reikard et al., 2015). It is important to stress that the approach described in this study is the authors' latest effort to predict wave energy and characterise the Mutriku wave farm. This study would not have been possible without our prior research.

Concerning renewable energies, one of the EU's main goals for 2030 is that at least 27% of the energy consumed will come from renewable sources, and one of the strategies involves the delivery of renewable energy to the power grid. The importance of the development of smart grid technologies has therefore been stressed (Gouardères et al., 2018).

Waves, wind and sun share a common drawback, namely, the intermittent availability of the resource. This means that the electricity supplied to the grid by an intermittent resource is also intermittent, and this poses serious grid management challenges. Accordingly, the forecasting of both the resource and the electricity generated by renewables below certain error thresholds is a key aspect for including such varying energy sources into the electricity market. In the case of more developed renewable energies, such as wind or solar power, this means different forecasting strategies are used.

Wind is the renewable energy that is spreading more rapidly, and it is expected to reach 2000 GW by 2030, accounting for 17–19% of overall electricity generation. To support this growth, higher quality forecasting methods are required (Costa et al., 2008; Wang et al., 2011), so two approaches are being pursued: the direct forecasting of power values, and the use of power curves to convert the predicted wind speed into forecasted power. In the case of wind farms, which include turbines with different power ratings, the latter option may be more expedient. Concerning forecasting models, two major types are normally used: physical and statistical. In general, physical models provide better results for longer-term horizons, while statistical ones are normally used for relatively short terms. As regards statistical models, machine-learning-based methods are considered to be the most effective due to their ability to model nonlinear relationships. However, there is no perfect forecasting model, and so it is important to use the one that gives the most accurate results depending on different factors, such as the characteristics and size of the data, and the prediction horizon. In 2014, Taşcıkaraoğlu and Uzunoglu (2014) presented a table comparing the more widely used wind forecasting models and their different approaches. There are also recent studies that use hybrid methods combining statistical and physics-based model forecasting, and which provide very accurate results for specific wind farms (Liu et al., 2017). In terms of the most widely studied sites in Europe, the performance of single wind farm forecasting in power production is between 8 and 10% 24 h ahead, according to MAE. By contrast, Chinese forecasting techniques for wind farms present R-Squared values of almost 20% for the same forecasting period. Concerning wind speed prediction, the Prediktor forecasting system (Watson et al., 2001) records a MAE of 2.4 m/s in one day; that is, a deviation of 40% for a typical site with an average wind speed of 6 m/s, with subsequent errors for the prediction of wind power output. If we compare the results obtained with the advanced forecasting models used in wind energy to the ones obtained here for Mutriku (forecasting errors from 40% to 60% for a 24-h horizon), the latter can be considered satisfactory results in view of the recent emergence of the wave energy field.

Regarding solar power forecasts, and as with wind energy, two different methods are normally used: the prediction of power values using power measurements, and the use of power conversion models for converting the predicted solar irradiance into power forecasts. In the latter case, other variables such as wind speed, humidity and temperature are also used (besides irradiance) to obtain more accurate forecasts.

Some recent studies propose a hypothetical 100% renewable energy scenario (Esteban et al., 2012, 2018). The authors therefore simulate future hourly energy generation based on wind, solar and ocean sources, and propose the use of storage devices, such as batteries or hydrogen, to balance the system. They also highlight the importance of levelling the generation peaks through the use of interconnection lines between different regions or countries, especially in Europe. Other studies have focused on the demand for electricity by pointing to the need for better energy policies related to more efficient energy consumption (Roldan et al., 2016). In this study, focused on Spain, a seven-year period is analysed (2008–2014), and a series of scenarios are proposed to analyse the effects of renewable energy production and the demand-side management of the electricity market.

Iberian Electricity Market Operator, OMIE, manages the wholesale electricity market on the Iberian Peninsula. Electricity prices are set every day at 12 a.m. in the day-ahead market, and are valid for the following 24 h. According to the EU marginal pricing model, the intersection between the demand and supply curves establishes the volume of energy and each hourly price. After the day-ahead market, electricity can be bought and sold on the intraday market, which is structured into six different trading sessions and allows agents to readjust their bids up to 4 h before real time (OMIE). Thus, both electricity market agents and system operators are gaining experience in the field of renewable energy forecasting, but more effort is required to develop prediction tools that can be used within the energy market's decision-making framework.

The models presented here, based on RF fed with ERA5 data in both analysis and forecasting mode, may constitute the general approach that will enable wave energy to take a major step towards its full integration in the electricity grid. In the particular case of Mutriku, and within the context of the Spanish electricity market, the models shown here provide reasonable results from 8 to 10 h ahead onwards, and can be an interesting tool for auctions in the day-ahead market. They can also provide useful insights for the intraday market. However, beyond these initial comments, further studies with more data are needed to better establish the specific role that RF-based models can play in the day-ahead and intraday markets. Further studies with more data will also help elucidate how forecasting models with errors such as those reported in this study can be effectively implemented and incorporated into successful strategies for electricity market auctions.

Thus, the analysis and construction of reliable models based on real data, like the ones presented here, are needed to obtain accurate forecasting that would make it easier to integrate renewable energies into the electricity grid.

## 4. Conclusions and future outlook

### 4.1. Conclusions

The electricity generated at the Mutriku wave farm on the Bay of Biscay was predicted for a 24-h forecast lead time through the use of RF-based models. For short lead times (t+4), the most influential input variable was the current electric power at time t, thus indicating the important role that system memory has in the prediction. As the prediction horizons increased (t+8 and t+12), the ERA5 wave energy flux (WEF) forecast gained in importance, while the current electric power at time t became less important for longer horizons (t+16 and beyond). When comparing the R-Squared, MAE and MAPE (at a 95% confidence level), the RF model proved to be a better prediction tool than Persistence beyond 8–10 h. RF models tend to overpredict electric power with forecasting errors ranging from 40% to 60% for a 24-h horizon.

This is the first time the output power of a fully operational wave farm has been forecast up to a 24-h ahead horizon through the use of logged electric power data. The role of ERA5 at providing observational data, and especially wave energy flux forecasts with an unprecedented time and space resolution, has also been crucial for this study.

In general terms, and before becoming a reliable source, wave energy needs major improvements in many aspects. Among them, an important point is to maximise the accuracy of the forecasting methods to solve the intermittency and variability problems associated with renewable energies. In this sense, and taking into account day-ahead and intraday market performance in the electricity market, more accurate prediction tools would enable this energy to be better accommodated within the grid, and so government and private sector engagement and funding are crucial for a renewable energy-based future. The study presented here makes an initial contribution to this objective, and the general availability of ERA5, the RF-based methodology –although so far developed only for Mutriku- can also be applied to future wave energy facilities, which while continuously supplying electricity to the grid operate with systematic protocols of operational data recording and management.

### *4.2. Future outlook*

As noted in the Introduction, 246 469 kWh of energy was produced each year at the Mutriku wave farm during the period 2014–2016. According to the emission factors proposed by the Spanish Government for 2017 (*ca.* 0.3 kg of $CO_2$/kWh) (Ministry of Agriculture and Fisheries, Food and the Environment, 2018) this means that the atmospheric discharge of 75 tonnes of $CO_2$ was avoided each year, indicating that higher percentages of renewable energy are needed to reduce greenhouse gas emissions. However, some improvements are needed for the optimum use and management of renewable energies. For instance, despite the advances in forecasting models, there are still several targets to be achieved. One of them is the development of probabilistic spatiotemporal forecasting methods to improve the operation of power systems. Probabilistic methods provide information on the probabilities of predicted values, as well as uncertainty values, which is very important for power systems, especially for the renewable energy sector. In addition, analysing the high volume of data available would currently require the use of different "big-data" techniques to select the most relevant information (Mazzi and Pinson, 2017; Taşcıkaraoğlu, 2017).

In our case, the use of a larger volume of data from the Mutriku wave farm and their corresponding screening would permit a more accurate analysis of the sea state, and the establishment of better and more precise forecasting limits for the RF model. The results shown in this paper provide a general or average model performance characterisation. However, the electricity production at Mutriku is very variable, so using a higher number of cases will allow a more detailed characterisation of the models' performance by sea state or season of the year. As a result, a deeper understanding may be gained of the specific role these models play in the day-ahead and intraday electricity markets.

When more cases are available, the RF technique also allows obtaining probabilistic forecasting outputs that can be more useful for electricity market auctions. A further advantage of future research with more data is that RF models will be developed for hourly instead of 4-h time steps, thus increasing the general accuracy in electric power predictions and allowing a better characterisation of model errors. The authors are currently conducting further research along these lines.

The authors have also developed a calibration method for studying and analysing the wave energy resource's evolution over the last century in the Bay of Biscay (Ulazia et al., 2017) and Ireland (Penalba et al., 2018). They have implemented these variations in different WECs to show the subsequent energy production trends in terms of capacity width, and not electricity production. Strong positive decadal increments of wave energy have been found in general, which cannot be absorbed properly by the WECs. This study therefore allows obtaining a transfer function from the resource to the electrical power in each sea state in the time series. Another future research line is to analyse the hypothetical annual energy production at the Mutriku OWC plant in each year of the last century, as well as other important aspects about seasonality and historical trends in electricity production.

## Acknowledgements

This work has been financially supported by the Spanish Government through the MINECO project CGL2016-76561-R, (MINECO/ERDF, UE) and the University of the Basque Country (UPV/EHU, GIU 17/002). ERA5 hindcast data have been downloaded at no cost from the MARS server of the ECMWF. All the calculations have been carried out within the framework of R (R Core Team, 2018). Special thanks to the Basque Energy Agency (EVE www.eve.eus) for kindly providing the data from the Mutriku wave farm used in this study.

## References

Ahmad, M.W., Mourshed, M., Rezgui, Y., 2017. Tree vs Neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy Build. 147, 77–89. https://doi.org/10.1016/j.enbuild.2017.04.038.

Blue Economy, 2018. The 2018 annual economic report on the EU blue economy. Eur. Comm. https://publications.europa.eu/en/publication-detail/-/publication/79299d10-8a35-11e8-ac6a-01aa75ed71a1.

Blue energy, 2014. Action needed to deliver on the potential of ocean energy in European seas and oceans by 2020 and beyond. In: Communication from the Commission to the European Parliament, the Council, vol. 20. The European Economic and Social Committee and the Committee of the Regions, Brussels, 1. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1396419828231&uri=CELEX:52014DC0008.

Becker, R.A., Wilks, A.R., (Original S code). R version by Ray Brownrigg Enhancements by, Minka, T.P., Deckmyn, A., 2018. Maps: Draw Geographical Maps. R package version 3.3.0. https://CRAN.R-project.org/package=maps.

Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. Applied Spatial Data Analysis with R, second ed. Springer, NY http://www.asdar-book.org/.

Bivand, R., Keitt, T., Rowlingson, B., 2018. Rgdal: bindings for the 'Geospatial' data abstraction library. R package version 1.3-4. https://CRAN.R-project.org/package=rgdal.

BOE: Boletín Oficial del Estado, 2017. Spain's official state gazette. BOE-A-2017-1793. https://www.boe.es/buscar/doc.php?id=BOE-A-2017-1793.

Botterud, A., 2017. Forecasting renewable energy for grid operations. In: Second Edition, Renewable Energy Integration. Practical Management of Variability, Uncertainty, and Flexibility in Power Grids, second ed. Academic Press – Elsevier, pp. 133–143 https://www.sciencedirect.com/science/article/pii/B978012809592800010X.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Bretherton, C.S., Widmann, M., Dymnikov, V.P., Wallace, J.M., Bladé, I., 1999. The effective number of spatial degrees of freedom of a time-varying field. J. Clim. 12, 1990–2009. <1990:TENOSD>2.0.CO;2. https://doi.org/10.1175/1520-0442(1999)012.

Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., Feitosa, E., 2008. A review on the young history of the wind power short-term prediction. Renew. Sustain. Energy Rev. 12 (6), 1725–1744. https://doi.org/10.1016/j.rser.2007.01.015.

Dommenget, D., Latif, M., 2002. A cautionary note on the interpretation of EOFs. J. Clim. 15, 216–225. <0216:ACNOTI>2.0.CO;2. https://doi.org/10.1175/1520-0442(2002)015.

Esteban, M., Zhang, Q., Utama, A., 2010. Methodology to estimate the output of a dual solar-wind renewable energy system in Japan. Energy Policy 38 (12), 7793–7802. https://doi.org/10.1016/j.enpol.2010.08.039.

Esteban, M., Leary, D., 2012. Current developments and future prospects of offshore wind and ocean energy. Appl. Energy 90, 128–136. https://doi.org/10.1016/j.apenergy.2011.06.011.

Esteban, M., Zhang, Q., Utama, A., 2012. Estimation of the energy storage requirement of a future 100% renewable energy system in Japan. Energy Policy 47, 22–31. https://doi.org/10.1016/j.enpol.2012.03.078.

Esteban, M., Portugal-Pereira, J., Mclellan, B.C., Bricker, J., Farzaneh, H., Djalilova, N., Ishihara, K.N., Takagi, H., Roeber, V., 2018. 100% renewable energy system in Japan: smoothening and ancillary services. Appl. Energy 224, 698–707. https://doi.org/10.1016/j.apenergy.2018.04.067.

Fernández-Ferrero, A., Sáenz, J., Ibarra-Berastegi, G., Fernández, J., 2009. Evaluation of statistical downscaling in short range precipitation forecasting. Atmos. Res. 94, 448–461. https://doi.org/10.1016/j.atmosres.2009.07.007.
Fernández-Ferrero, A., Sáenz, J., Ibarra-Berastegi, G., 2010. Comparison of the performance of different analogue-based Bayesian probabilistic precipitation forecasts over Bilbao, Spain. Mon. Weather Rev. 138, 3107–3119. https://doi.org/10.1175/2010MWR3284.1.
Fukutomi, Y., Yasunari, T., 2013. Structure and characteristics of submonthly-scale waves along the Indian Ocean ITCZ. Clim. Dyn. 40 (7–8), 1819–1839. https://doi.org/10.1007/s00382-012-1417-x.
Gouardères, F., McWatt, V., Fleuret, L., 2018. European parliament. Fact sheets on the European Union. Renew. Energy. http://www.europarl.europa.eu/factsheets/en/sheet/70/la-energia-renovablele.
Gromping, U., 2009. Variable importance assessment in regression: linear regression versus random forest. Am. Statistician 63, 308–319. https://doi.org/10.1198/tast.2009.08199.
Hannachi, A., Jolliffe, I.T., Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: a review. Int. J. Climatol. 27, 1119–1152. https://doi.org/10.1002/joc.1499.
Ibarra-Berastegi, G., Saenz, J., Ezcurra, A., Ezcurra, A., Elias, A., Diaz Argandoña, J., Errasti, I., 2011. Downscaling of surface moisture flux and precipitation in the Ebro Valley (Spain) using analogues and analogues followed by random forests and multiple linear regression. Hydrol. Earth Syst. Sci. 15 (6), 1895–1907. https://doi.org/10.5194/hess-15-1895-2011.
Ibarra-Berastegi, G., Saenz, J., Esnaola, G., Ezcurra, A., Ulazia, A., 2015. Short-term Forecasting of Zonal and Meridional Wave Energy Flux in the Bay of Biscay Using Random Forests. OCEANS, Genova, pp. 1–6. https://doi.org/10.1109/OCEANS-Genova.2015.7271404.
Ibarra-Berastegi, G., Saenz, J., Esnaola, G., Ezcurra, A., Ulazia, A., 2015. Short-term forecasting of the wave energy flux: analogues, random forests, and physics-based models. Ocean. Eng. 104, 530–539. https://doi.org/10.1016/j.oceaneng.2015.05.038.
Ibarra-Berastegi, G., Saenz, J., Esnaola, G., Ezcurra, A., Ulazia, A., 2016. Wave energy forecasting at three coastal buoys in the Bay of Biscay. J. Ocean. Eng. 41, 923–929. https://doi.org/10.1109/JOE.2016.2529400.
Ibarra-Berastegi, G., Saenz, J., Ulazia, A., Serras, P., Esnaola, G., García-Soto, C., 2018. Electricity production, capacity factor, and plant efficiency index at the Mutriku wave farm (2014-2016). Ocean. Eng. 147, 20–29. https://doi.org/10.1016/j.oceaneng.2017.10.018.
Jiang, R., Tang, W., Wu, X., Fu, W., 2009. A random forest approach to the detection of epistatic interactions in case-control studies. BMC Bioform 10 (1), 1. https://doi.org/10.1186/1471-2105-10-S1-S65.
Kaur, A., Nonnenmacher, L., Pedro, H.T.C., Coimbra, C.F.M., 2016. Benefits of solar forecasting for energy imbalance markets. Renew. Energy 86, 819–830. https://doi.org/10.1016/j.renene.2015.09.011.
Lavidas, G., Venugopal, V., 2018. Application of numerical wave models at European coastlines: a review. Renew. Sustain. Energy Rev. 92, 489–500. https://doi.org/10.1016/j.rser.2018.04.112.
Le, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25 (1), 1–18. https://doi.org/10.18637/jss.v025.i01.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. RNews2 (3), 18–22.
Liu, J., Wang, X., Lu, Y., 2017. A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system. Renew. Energy 103, 620–629. https://doi.org/10.1016/j.renene.2016.10.074.
Magagna, D., Uihlein, A., 2015. Ocean energy development in Europe: current status and future perspectives. Int. J. Mar. Energy 11, 84–104. https://doi.org/10.1016/j.ijome.2015.05.001.
Mazzi, N., Pinson, P., 2017. Wind power in electricity markets and the value of forecasting. In: Kariniotakis, George (Ed.), Woodhead Publishing Series in Energy. Renewable Energy Forecasting: From Models to Applications, pp. 259–278. https://www.sciencedirect.com/book/9780081005040/renewable-energy-forecasting.
Ministry of Agriculture and Fisheries, 2018. Food and the Environment. Government of Spain. OECD. Factores de emisión. Registro de huella de carbono, compensación y proyectos de absorción de dióxido de carbono. https://www.miteco.gob.es/es/cambio-climatico/temas/mitigacion-politicas-y-medidas/factores_emision_tcm30-446710.pdf.
Monahan, A.H., Fyfe, J.C., Ambaum, M.H.P., Stephenson, D.B., North, G.R., 2009. Empirical orthogonal functions, the medium is the message. J. Clim. 22, 6501–6514. https://doi.org/10.1175/2009JCLI3062.1.
Multon, B., 2012. In: Marine Renewable Energy Handbook. Wiley. https://doi.org/10.1002/9781118603185.
North, G.R., 1984. Empirical orthogonal functions and normal modes. J. Atmos. Sci. 41, 879–887.
Notton, G., Nivet, M.-L., Voyant, C., Paoli, C., Darras, C., Motte, F., Fouilloy, A., 2018. Intermittent and stochastic character of renewable energy sources: consequences, cost of intermittence and benefit of forecasting. Renew. Sustain. Energy Rev. 87, 96–105. https://doi.org/10.1016/j.rser.2018.02.007.
O'Hagan, A.M., Huertas, C., O'Callaghan, J., Greaves, D., 2016. Wave energy in Europe: views on experiences and progress to date. Int. J. Mar. Energy 14, 180–197. https://doi.org/10.1016/j.ijome.2015.09.001.
Olauson, J., 2018. ERA5: the new champion of wind power modelling? Renew. Energy 126, 322–331. ISSN 0960-1481. https://doi.org/10.1016/j.renene.2018.03.056.
OMIE. Iberian electricity market operator. http://www.omie.es/en/inicio.
Parkinson, S.C., Dragoon, K., Reikard, G., García-Medina, G., Özkan-Haller, H.T., Brekken, T.K.A., 2015. Integrating ocean wave energy at large-scales: a study of the US Pacific Northwest. Renew. Energy 76, 551–559. https://doi.org/10.1016/j.renene.2014.11.038.
Pebesma, E.J., Bivand, R.S., 2005. Classes and methods for spatial data in R. R. News 5 (2). https://cran.r-project.org/doc/Rnews/.
Penalba, M., Ulazia, A., Ibarra-Berastegui, G., Ringwood, J., Sáenz, J., 2018. Wave energy resource variation off the west coast of Ireland and its impact on realistic wave energy converters' power absorption. Appl. Energy 224, 205–219. https://doi.org/10.1016/j.apenergy.2018.04.121.
Pinson, P., Reikard, G., Bidlot, J.-R., 2012. Probabilistic forecasting of the wave energy flux. Appl. Energy 93, 364–370. https://doi.org/10.1016/j.apenergy.2011.12.040.
R Core Team, 2018. The R Project for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/.
Reikard, G., Pinson, P., Bidlot, J.-R., 2011. Forecasting ocean wave energy: the ECMWF wave model and time series methods. Ocean. Eng. 38, 1089–1099. https://doi.org/10.1016/j.oceaneng.2011.04.009.
Reikard, G., Rogers, W.E., 2011. Forecasting ocean waves: comparing a physics-based model with statistical models. Coast. Eng. 58, 409–416. https://doi.org/10.1016/j.coastaleng.2010.12.001.
Reikard, G., Robertson, B., Buckham, B., Bidlot, J.-R., Hiles, C., 2015. Simulating and forecasting ocean wave energy in western Canada. Ocean. Eng. 103, 223–236. https://doi.org/10.1016/j.oceaneng.2015.04.081.
Reikard, G., 2017. Wave energy forecasting. In: Kariniotakis, George (Ed.), Renewable Energy Forecasting. From Models to Applications. Woodhead Publishing Series in Energy, pp. 199–217. https://www.sciencedirect.com/science/article/pii/B978008100504000007X.
Reikard, G., Robertson, B., Bidlot, J.-R., 2017. Wave energy worldwide: simulating wave farms, forecasting, and calculating reserves. Int. J. Mar. Energy 17, 156–185. https://doi.org/10.1016/j.ijome.2017.01.004.
Roldan, J.M., Burgos, M., Riquelme, J.M., Trigo, A.L., 2016. Renewable generation versus demand-side management. A comparison for the Spanish market. Energy Policy 96, 458–470. https://doi.org/10.1016/j.enpol.2016.06.014.
Santos, S.F., Fitiwi, D.Z., Shafie-khah, M., Bizuayehu, A.W., Catalão, J.P.S., 2017. Introduction to renewable energy systems. In: Erdinç, Ozan (Ed.), Optimization in Renewable Energy Systems. Recent Perspectives. Butterworth-Heinemann. Elsevier, pp. 1–26. https://www.sciencedirect.com/science/article/pii/B9780081010419000016.
Siroky, D.S., 2009. Navigating Random Forests and related advances in algorithmic model. Stat. Surv. 3, 147–163. https://doi.org/10.1214/07-SS033.
Taşcıkaraoğlu, A., Uzunoglu, M., 2014. A review of combined approaches for prediction of short-term wind speed and power. Renew. Sustain. Energy Rev. 34, 243–254. https://doi.org/10.1016/j.rser.2014.03.033.
Taşcıkaraoğlu, A., 2017. Impacts of accurate renewable power forecasting. In: Erdinç, Ozan (Ed.), Optimization in Renewable Energy Systems. Recent Perspectives. Butterworth-Heinemann. Elsevier, pp. 159–175. https://www.sciencedirect.com/science/article/pii/B9780081010419000053.
Uihlein, A., Magagna, D., 2016. Wave and tidal current energy – a review of the current state of research beyond technology. Renew. Sustain. Energy Rev. 58, 1070–1081. https://doi.org/10.1016/j.rser.2015.12.284.
Ulazia, A., Penalba, M., Ibarra-Berastegui, G., Ringwood, J., Sáenz, J., 2017. Wave energy trends over the Bay of Biscay and the consequences for wave energy converters. Energy 141, 624–634. https://doi.org/10.1016/j.energy.2017.09.099.
Wang, X., Shen, S.S., 1999. Estimation of spatial degrees of freedom of a climate field. J. Clim. 12, 1280–1291, 1280:EOSDOF>2.0.CO;2. https://doi.org/10.1175/1520-0442(1999)012.
Wang, X., Guo, P., Huang, X., 2011. A review of wind power forecasting models. Energy Procedia 12, 770–778. https://doi.org/10.1016/j.egypro.2011.10.103.
Watson, R., Landberg, L., Costello, R., McCoy, D., O'Donnell, P., 2001. Evaluation of the Prediktor wind power forecasting system in Ireland. In: Proceedings of European Wind Energy Conference. Copenhagen.
Weare, B.C., Nasstrom, J.C., 1982. Examples of extended empirical orthogonal function analyses. Mon. Weather Rev. 110, 481–485.
Widén, J., Carpman, N., Castellucci, V., Lingfors, D., Olauson, J., Remouit, F., Bergkvist, M., Grabbe, M., Waters, R., 2015. Variability assessment and forecasting of renewables: a review of solar, wind, wave and tidal resources. Renew. Sustain. Energy Rev. 44, 356–375. https://doi.org/10.1016/j.rser.2014.12.019.
report, World Energy Outlook Special Report, 2015. International Energy Agency, Paris, France.