POLITECNICO DI TORINO
Repository ISTITUZIONALE

A comparative analysis of Machine Learning Techniques for short-term grid power forecasting and uncertainty analysis of Wave Energy Converters

Original
A comparative analysis of Machine Learning Techniques for short-term grid power forecasting and uncertainty analysis of Wave Energy Converters / FONTANA CRESPO, R.N., Aliberti, A., Bottaccioli, L., Pasta, E., Sirigu, S.A., Macii, E., Mattiazzo, G., Patti, E.. - In: ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE. - ISSN 0952-1976. - 138, Part A:(2024). [10.1016/j.engappai.2024.109352]

Availability:
This version is available at: 11583/2992828 since: 2024-09-26T16:31:29Z

Publisher:
Elsevier

Published
DOI:10.1016/j.engappai.2024.109352

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

25 November 2026

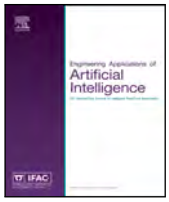
Engineering Applications of Artificial Intelligence 138 (2024) 109352



Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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



Research paper



A comparative analysis of Machine Learning Techniques for short-term grid power forecasting and uncertainty analysis of Wave Energy Converters

Rafael Natalio Fontana Crespo [a,*], Alessandro Aliberti [a], Lorenzo Bottaccioli [b], Edoardo Pasta [c], Sergej Antonello Sirigu [c], Enrico Macii [a], Giuliana Mattiazzo [c], Edoardo Patti [a,b]

[a] Department of Control and Computer Engineering, Politecnico di Torino, 10129, Turin, Italy
[b] Energy Center Lab, Politecnico di Torino, 10129, Turin, Italy
[c] Department of Mechanical and Aerospace Engineering, Politecnico di Torino, 10129, Turin, Italy

ARTICLE INFO

Keywords:
Renewable energy
Wave energy converter
Energy forecast
Neural networks
Non-parametric kernel density estimation
Uncertainty analysis
Prediction intervals

ABSTRACT

Wave Energy is one of the renewable sources with greatest potential. Since power coming from waves fluctuates, the grid integration of wave energy involves several power conditioning stages to comply with grid quality requirements. However, to ensure full integration of wave energy in a smart grid scenario and unlock advanced monitoring and control techniques (e.g. Demand/Response), it is crucial to forecast the output power. This work proposes a methodology to forecast in short-term horizons (i.e. 15 min to 240 min) the power delivered to the grid of the Inertial Sea Wave Energy Converter (ISWEC), a device that harnesses wave power through the inertial effect of a gyroscope. Therefore, we designed, optimized and compared the performance of five known machine learning techniques for time series point forecasting: Random Forest, Support Vector Regression, Long Short-Term Memory Neural Network, Transformer Neural Network and 1 Dimensional Convolutional Neural Network. Additionally, we studied the efficacy of downsampling technique aggregating original dataset sampled every 0.1 s in time steps of 1min, 3min, 5min and 15min to compare the performance behaviour of the different machine learning models for these datasets. Furthermore, we implemented Prediction Intervals (PIs) to calculate the inherent uncertainties associated with the previously mentioned machine learning techniques. These PIs were built based on the Non-Parametric Kernel Density Estimator technique. The point forecasting and the PIs results showed that models' performance improved as the downsampling increased. Moreover, the Random Forest model was the worst-performing in all cases. Finally, none of the other models can be considered the best overall.

1. Introduction

Nowadays, Renewable Energy Sources (RES) are a very hot topic, attracting more and more attention. Implementing environmentally friendly energy alternatives to either replace or drastically reduce the usage of fossil fuels has become an essential objective. According to the International Energy Agency, the capacity of RES is expected to increase by almost 75% between 2022 and 2027 (International Energy Agency, 2023). In this context, wave energy will become a key player in the renewables field as a result of its vast energy potential: wave energy's annual potential that can be fully harnessed is estimated to be 18 000 TWh per year, almost 70% of the global electricity consumption in 2022 (Eni S.p.A, 2023; Enerdata, 2023). Moreover, wave power has greater predictability and is significantly more reliable than wind and solar power: it can generate electricity for as much as 90% of the time, compared to 20% and 30% for wind and solar energy (Drew et al., 2009).

The process of harnessing wave energy and converting it into electrical power involves the utilization of Wave Energy Converters (WECs). These devices first convert the wave's motion into mechanical energy in the Power Take-Off (PTO) and later into electrical energy. WECs are categorized based on their installation location (i.e. offshore, near-shore and onshore), operating principle (i.e. overtopping, floating structures, pressure differential, and impact devices), and size (i.e. terminator, attenuator and point absorber) (Bracco et al., 2016). Although all WECs have the same purpose, the design approach varies among the solutions. Point absorbers, relatively small to the wavelength, move within the waves. On the contrary, attenuators have a dimension aligned with the wavelength and use the wave to create relative motion between its

* Corresponding author.
E-mail address: rafael.fontana@polito.it (R.N. Fontana Crespo).

https://doi.org/10.1016/j.engappai.2024.109352
Received 22 April 2024; Received in revised form 13 August 2024; Accepted 15 September 2024
Available online 26 September 2024
0952-1976/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

1D-CNN	1 Dimensional Convolutional Neural Network
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
CRPS	Continuous Rank Probability Score
ISWEC	Inertial Sea Wave Energy Converter
LSTM	Long Short-Term Memory
LSTM E-D	LSTM Encoder-Decoder
LSTM VEC	LSTM Vector
LUBE	Lower Upper Bound Estimate
MAD	Mean Absolute Difference
MAE	Mean Absolute Error
MALE	Mean Absolute Log Error
MAPE	Mean Absolute Percentage Error
MCS	Model Confidence Set
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NPKDE	Non-Parameteric Kernel Density Estimation
PF	Point Forecasting
PCA	Principal Component Analysis
PI	Prediction Interval
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalized Average Width
PTO	Power Take-Off
R	Pearson's Correlation Coefficient
R^2	Coefficient of determination
RES	Renewable Energy Sources
RF	Random Forest
RMSD	Root Mean Squared Deviation
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SVR	Support Vector Regression
TNN	Transformer Neural Network
VAF	Variance Accounted For
WEC	Wave Energy Converter

parts. A terminator aims to maximize power extraction by effectively terminating wave motion. Another category of WEC is the inertial one, where the PTO reactions are generated by inertial factors. A particular WEC of this category is the Inertial Sea Wave Energy Converter (ISWEC). The ISWEC is an offshore floating-point absorber that uses a buoy and the reactive inertial effects of a gyroscope to harness wave energy (Bracco et al., 2016).

In recent years, the increasing utilization of RES led to the advance towards more advanced power grids known as Smart Grids (Siano, 2014). A traditional power grid implements a unidirectional power flow from the power generator to the consumers. On the contrary, the Smart Grid introduces a distributed approach where consumers can provide valuable feedback on how electricity is consumed, enabling the grid to optimize its performance and enhance its overall reliability. Nonetheless, a significant challenge arises for almost every RES due to their non-dispatchable and intermittent characteristics. The output power of a Wave Energy Converter varies due to the stochastic nature of waves, causing a misalignment between power generation and load (Reikard, 2013).

Recent research indicates that smart system operational techniques, such as Demand-Response (Siano, 2014), State Estimation (Ahmad et al., 2018) and Real-time forecasting (Sugihara et al., 2017), can mitigate fluctuating power issues. Demand-Response encompasses alterations in the end-users electricity consumption habits in response to the variations of RES power production and grid requirements, together with financial considerations such as shifts in electricity pricing. On the other hand, State Estimation is the procedure of assessing the grid condition by comparing the system measurements and predicted ones at a specific time *t*. This procedure facilitates the identification of potential system anomalies and measurement inaccuracies.

To implement the forenamed procedures, it is fundamental to forecast the output power delivered from WECs. Depending on the forecasting horizon, Wave Power forecasts can be categorized into nowcasting (from a few seconds to 1 min), very-short-term (from 1 min to 15 min), short-term (from 15 min to 4 h), mid-term (from 4 h to 1 day) and long-term forecasts (from days to even years). State estimation is one of the applications of nowcasting. On the other hand, very short-term forecasts can be employed for WEC control. Short-term forecasts can be used for efficient dispatching of the power grid and for ensuring power quality supply. On the contrary, Mid-long-term forecasts can be exploited for feasibility studies of the WEC's farm design and location together with forecasts of the annual power generation (Soman et al., 2010).

Since the 1960s, predictions in the wave energy field have been performed employing physical models. Nonetheless, the computational cost of these models has significantly increased over time in an attempt to minimize errors (Nalamati, 2021b). In recent years, the Machine Learning (ML) field introduced significant progress in advanced and efficient learning algorithms (Janiesch et al., 2021). Therefore, many researchers consider these algorithms a robust and efficient methodology to predict a wide range of phenomena, such as fault estimation (Wang et al., 2023), solar radiation (Voyant et al., 2017), and wave power (Shadmani et al., 2023). Artificial Neural Networks (ANN) are one of the ML algorithms that has shown remarkable advancements in learning capabilities, leading to deep neural network architectures, known as Deep Learning (DL) (LeCun et al., 2015). The development of these deep complex architectures has led to a great interest in resource efficiency within the research community. Song et al. (2023) proposed a bipartite synchronization method in reaction–diffusion neural networks, which employed a dural event-triggered control algorithm for resource optimization. Moreover, Wan et al. (2023) introduced an adaptive self-triggered mechanism for discrete-time Markov jump systems to minimize resource consumption.

In the wave energy field, ML and DL algorithms have been employed to perform predictions in the area, showing significant advantages in terms of execution time and cost (Mousavi et al., 2021). These comprise the application of Random Forest (RF), Support Vector Regression (SVR), Artificial Neural Network (ANN), Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN). The Long Short-Term Memory (LSTM) algorithm, derived from RNN, has shown a remarkable ability to handle a range of sequential learning-related issues (Srivastava and Lessmann, 2018). The LSTM algorithm has been employed by several authors in the literature (Ni et al., 2019; Mousavi et al., 2021; Ni, 2021; Neshat et al., 2022) for predicting the power generated by WECs. Furthermore, in Mousavi et al. (2021) and Ni (2021), the original dataset was aggregated, i.e. downsampled (Hatalis et al., 2014), in time intervals (e.g. 15 min). Downsampling is frequently employed to make data smoother, reducing the peaks of signals and fluctuations that data may have. Nonetheless, the existing models proposed in literature only address the forecasting challenge of the instantaneous output power of the WEC PTO, ignoring the grid integration of the device. The grid integration of a WEC, an essential phase in the wave-to-grid power generation process, involves many challenges (e.g. power conditioning) and has to be considered. Fig. 1 presents a comparison between the power curves of the instantaneous output power of the WEC PTO and the power delivered to the grid of the same WEC. The power has been normalized from 0 to 1 using *Min-Max* normalization, where 0 and the

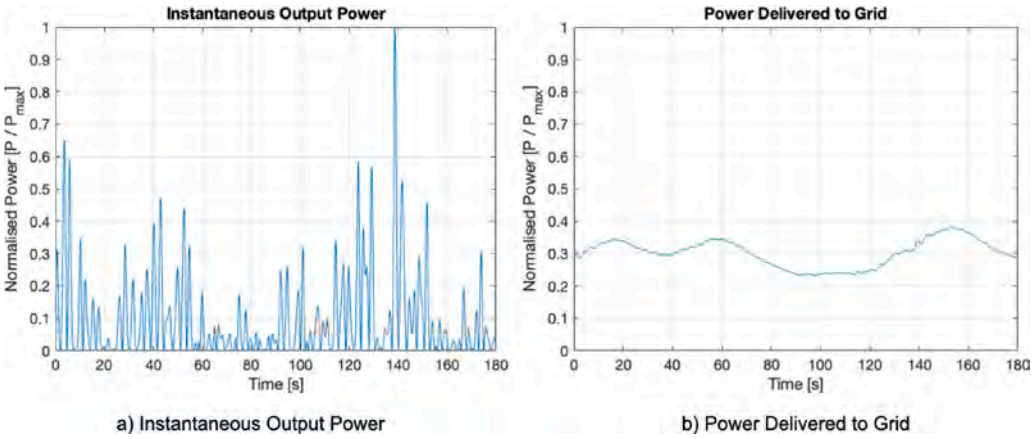


Fig. 1. Comparison instantaneous electric power and power delivered to the grid.

maximum value found in the dataset were used as the *Min* and *Max* values, respectively. We adopted this approach because the machine specifications (e.g. minimum, nominal and maximum power delivered to the grid) were not available. For a real-time implementation, where the complete dataset will not be available, the minimum and maximum values can be determined based on the device specifications. As can be seen, the power delivered to the grid is significantly more stable than the instantaneous one. Furthermore, this stability will allow obtaining better predictions as well as forecasting with longer time horizons. Nevertheless, real-world data accounting for this phase is rare or does not exist.

Moreover, the abovementioned machine learning algorithms perform point forecasting of the instantaneous output power. This means the algorithm's output single values for each point of interest in the forecasting horizon. For predicting the output power of WECs, it is fundamental to perform an accurate analysis of the uncertainties associated with the forecasting. Prediction Intervals (PIs) are a way for computing uncertainty. PIs have been employed in the energy forecasting area, particularly for solar (Chu and Coimbra, 2017; Gu et al., 2021a; Liu et al., 2018) and wind power (Banik et al., 2020; Gu et al., 2021b; Niu et al., 2022; Quan et al., 2013; Shi et al., 2017; Xie et al., 2023). Nonetheless, to the best of our knowledge, there is no reference to the implementation of Prediction Intervals for the forecasting of WEC's power delivered to the grid. There are several ways to calculate the Prediction Intervals for determining the power forecasting uncertainty, including parametric and non-parametric techniques. Non-Parameteric Kernel Density Estimation (NPKDE) is one of the non-parametric techniques that has emerged and gained significant attention. It has been widely used in the literature for determining Prediction Intervals for solar and wind power forecasting (Gu et al., 2021a; Liu et al., 2018; Gu et al., 2021b; Niu et al., 2022; Xie et al., 2023). The main advantage of the NPKDE, in contrast to previous approaches, is that it can predict the Prediction Intervals based on the forecasting errors and does not require any prior assumption about the data distribution.

Wave power fluctuating issues caused by waves' intermittent and stochastic nature are still key challenges that need to be addressed. Addressing these challenges is essential for enabling the full integration of wave energy in a smart grid scenario (Reikard, 2013). These fluctuations can be mitigated by implementing advanced smart grid operational techniques, like State Estimation and Demand-Response. However, the successful implementation of these smart grid techniques relies on accurate wave power forecasting. In this context, ML algorithms and uncertainty analysis methods represent a robust approach for obtaining precise forecasts of the power delivered from WECs and determining their associated uncertainties.

In this work, we propose a novel methodology to forecast in short-term horizons (i.e. 15 min to 240 min) the power delivered to the grid of the ISWEC device by exploiting ML algorithms and uncertainty analysis methods. The main goal is to obtain accurate predictions and be able to determine their associated uncertainties, making our forecasts suitable for smart grid advanced monitoring and control techniques, such as State Estimation and Demand-Response. In detail, we design, optimize and compare the performance of five Machine Learning techniques for point forecasting (i.e. LSTM, RF, SVR, 1D-CNN and TNN), which have shown promising results for wave power and time series forecasting. We exploit a realistic dataset consisting of 1.5 years of the electric power delivered to the grid of the ISWEC device, collected with the "realistic" *ISWEC Wave-to-Grid model* (Fontana Crespo et al., 2023). Additionally, to study the effectiveness of the downsampling technique, we aggregate the original dataset sampled every 0.1 s in time windows of 1 min, 3 min, 5 min and 15 min and compare the prediction performance of the different ML point forecasting algorithms. Finally, we build Prediction Intervals using the Non-Parametric Kernel Density Estimator technique to evaluate and quantify the uncertainties associated with the different point forecasting algorithms.

The rest of the article is organized as follows. Section 2 reviews the literature solutions for Wave Energy Forecasting as well as the literature found for uncertainty analysis for energy forecasting. Section 3 presents the proposed methodology for power-to-grid forecasting of the ISWEC device and introduces the proposed technique for computing the associated point uncertainties predictions. Section 4 discusses the experimental results for point forecasting and their associated uncertainties. Finally, Section 5 discusses the concluding remarks.

2. Related work

To efficiently integrate wave energy into the grid, the output power delivered from WECs must be forecasted. As indicated previously, the computational cost of physical models has increased notably when trying to minimize forecasting errors. On the other hand, time-series forecasting predicts WEC output power using past values and other exogenous variables, such as significant wave height. In this context, Machine Learning and Deep Learning algorithms have become one of the most popular approaches for performing these predictions due to their significant advantages in terms of time and cost (Mousavi et al., 2021). Consequently, these algorithms have been extensively used in the literature, as presented in Table 1 and Table 2.

Numerous works (Ni et al., 2019; Nalamati, 2021b; Ni, 2021; Deberneh and Kim, 2018b; Pereira et al., 2018; Burramukku, 2020; Nalamati, 2021a) forecasted the output power utilizing parameters derived from the Wave Energy Converter as input features. Nalamati (2021b), Burramukku (2020) and Nalamati (2021a) predicted the output power of an array of WECs from 4 distinct wave farms. The inputs employed were the device's power output and position (latitude and longitude). Nalamati (2021b) and Burramukku (2020) compared the prediction of the output power of each farm independently using several

Table 1
Literature on wave power forecasting.

Author and year	Fore. Hor.	Fore. Resol.	Variables	Methodology	Output	Uncertainty analysis	Metrics	Results
Hatalis et al. (2014)	3, 6, 12, 24 h	10 min	Weather-related exogenous variables	Conversion of input using power matrix. NAR NN for forecast	Output power	No	MAPE, MAD, R	Better results with smooth data. R results of up to 0.99
Ibarra-Berastegi et al. (2015)	0–24 h	1 h	Weather-related exogenous variables and wave energy flux	Compared WAM with persistence, analogues, RF, Analogues+RF	Wave energy flux	No	MAPE	RF showed the best results.
Deberneh and Kim (2018b)	15 min and 3 days	1 s to 6 s, 1 min to 10 min	Input parameters from WEC	PCA+ML (RF, SVM, ANN). Compared with PCA+Classifier+ML	Output power	No	MAE, MSE, R^2	PCA+Classifier+ML models presented better performance
Pereira et al. (2018)	–	10 min	Input parameters from WEC	NNARX to predict the wave induced excitation torque	Wave induced excitation torque	No	MSE, VAF	VAF results shown that the prediction using the NNARX were accurate
Ni et al. (2019)	–	–	Input parameters from WEC	PCA+LSTM (seq2seq)	Output power and WEC parameters	No	R^2, MAE, RMSE	PCA-LSTM outperform other traditional ML algorithms such as RT, ET, GPR, SVM
Burramukku (2020)	–	–	Input parameters from WEC	Sequence to one NN multi layer perceptron	Farm power output	No	MAE, RMSE	Compared the farm power output in different locations
Ni (2021)	0–72 h	15 min	Input parameters from WEC	PCA+LSTM (seq2one)	Ouput Power	No	MAPE, MAE, RMSE	PCA-LSTM outperform other ML algorithms such as LSTM, SVM, and Neural networks
Nalamati (2021b)	–	–	Input parameters from WEC	Vanilla neural network (seq2one)	Farm power output	No	MSE, MAE, MAPE, RMSE	NN with one hidden layer showed better results
Nalamati (2021a)	–	–	Input parameters from WEC	Compared different LSTM models, including Vanilla LSTM, Bi-LSTM, Stacked LSTM, and CNN-LSTM	Farm power output	No	MSE, MAE, MAPE, R^2	Bi-LSTM model showed better performance

configurations of a simple Multi-Layer perceptron model. Instead, Nalamati (2021a) studied the performance of different LSTM models, including Vanilla LSTM, Bi-Directional LSTM (Bi-LSTM), Stacked LSTM, and CNN-LSTM, with the Bi-LSTM model showing the best results.

Many authors try to lower the complexity of the forecasting task by employing Principal Component Analysis (PCA) to investigate possible correlations among the input data. Ni et al. (2019) implemented PCA to reduce the input dataset to features representing 90% of the original information. These features were later applied to an LSTM sequence-to-sequence neural network. The performance of this PCA-LSTM model was then compared with traditional ML algorithms, including Regression Tree (RT), Ensembled Trees (ET), Gaussian Process Regression (GPR) and Support Vector Machine (SVM). According to the results, the proposed PCA-LSTM model showed the best accuracy. Likewise, Ni (2021) employed a similar PCA-LSTM model but adopted a sequence-to-one architecture. Various models were studied, including raw LSTM, SVM and Neural Networks, with the proposed PCA-LSTM model demonstrating the highest accuracy. On the other hand, Deberneh and Kim (2018b) used as input a dataset comprising the angular velocities and accelerations along the three axes of the WEC together with its magnitudes. Afterwards, the data was downsampled into different window sizes. Later, PCA was applied to reduce the dimension of the input dataset. For the predictions, two methods were investigated: one implementing solely the predictor and another which combined the predictor with a classifier (Deberneh and Kim, 2018a). In addition, several predictors were tested: Random Forest (RF), SVM, and Artificial Neural Networks. The findings indicated that models incorporating a classifier combined with a regressor demonstrated better behaviour than the other models, with the model accuracy varying based on the adopted window size.

Pereira et al. (2018) employed a Neural Network Autoregressive with Exogenous Input (NNARX) to forecast the wave-induced torque of the Inertial Sea Wave Energy Converter. The inputs for this prediction model included the angular speed of the PTO and the hull, and the PTO torque.

Other authors, motivated by the possible correlation between weather variables (e.g. wind speed and wave height) and wave power, used weather-related data as exogenous inputs for their models to forecast the output power. Mousavi et al. (2021) employed an LSTM neural network to predict the output power of WEC, including wave height and wind speed as exogenous inputs. Similarly, Casas and Lekube (2023) used the wave height as exogenous input and compared the forecasting performance of different models: Random Walk (RW), Functional Coefficients Linear Model (FCLM), Functional Coefficients Autoregressive-Exogenous model (FCARX), Time-Varying Linear model (TVLM), Time-Varying Autoregressive-Exogenous model (TVARX), Random Forest (RF), Bayesian Regularization Neural Network (BRNN), Multi-Layer Perceptron (MLP), linear Support Vector Machine (SVM), and radial SVM. The results showed that the FCLM offered the best results. Likewise, Ni and Peng (2023) used the wave height and wave period as exogenous inputs and the WEC output power as endogenous input. These time-series signals were transformed into frequency domain signals by employing Empirical Wavelet Transform (EWT) and later converted into two dimensional signals. Then, the converted signals were input to a Convolutional Neural Network (CNN) model. The obtained predictions using the described EWT-CNN model were compared against a Wave Power Matrix model (WPM) together with EWT-SVM, EWT-LSTM and EWT-NN. The results showed that the proposed EWT-CNN achieved the best results among the compared models for different wave conditions. Hatalis et al. (2014) used the

Table 2
Literature on wave power forecasting part 2.

Author and year	Fore. Hor.	Fore. Resol.	Variables	Methodology	Output	Uncertainty analysis	Metrics	Results
Mousavi et al. (2021)	–	28.5 min	Weather-related exogenous variables and Output power	Compared forecast of output power using numerical solutions and LSTM (seq2one)	Output power	No	RMSE	LSTM model outperform numerical solutions
Bento et al. (2021)	1, 2, 3, 6, and 12 h	1 h	Weather-related exogenous variables and wave energy flux	FFNN optimized using MFO for hyperparameters tuning. Compared forecast of wave energy flux directly and indirectly	Wave energy flux	No	R, MALE, RMSE	Better results when forecasting wave energy flux indirectly (with forecast of wave height and period).
Neshat et al. (2022)	6 h	6 h	Weather-related exogenous variables and wave energy flux	Compared different ML models (e.g. PNN, FFNN, ANFIS, Vanilla LSTM, Bi-LSTM and Stacked LSTM).	Wave energy flux	No	MSE, RMSE, MAE, R	Bi-LSTM showed the best performance
Casas and Lekube (2023)	13–36 h	1 h	Weather-related exogenous variable and nominal power	Compared forecasting with RW, FCLM, FCARX, TVLM, TVARX, RF, BRNN, MLP, SVM, and Radial SVM	Output power	No	RMSE, MAE, MCS	FCLM model outperformed the other models
Ni and Peng (2023)	–	–	Weather-related exogenous variable and output power	Compared EWT-CNN with WPM, EWT-SVM, EWT-LSTM, EWT-NN	Output power	No	R^2, MAE, RMSE	EWT-CNN showed the best results
Fontana Crespo et al. (2023)	1 min	0.1, 1, 2.5, 5, 10, 30 s	Output power delivered to the grid	LSTM for forecast and compared different downsampling of the input dataset	Output power delivered to the grid	No	MAE, RMSE, R^2	The biggest downsampling (30 s) significantly outperformed the other models
This article	240 min	1, 3, 5 and 15 min	Output power delivered to the grid	Compares different ML algorithms (RF, SVR, 1D-CNN, LSTM E-D, LSTM VEC, TNN). PI for model uncertainties	Output power delivered to the grid	NPKDE to compute prediction intervals	MAE, RMSE, and R^2; PICP, PINAW and CRPS	Models performance improves as the downsampling increases. Not a best model for all downsamplings. Not always the best point forecasting model showed the best results for PI

significant wave height as input to the proposed model, to be later converted into output power using the power matrix of the employed WEC. The output power was input to a Non-Linear Autoregressive Recurrent Neural Network to perform the forecasting.

Besides, other works employed weather-related exogenous variables to predict the wave energy flux. Neshat et al. (2022) used the wave height, wave period, and wave direction as exogenous input to forecast the wave energy flux and compared several ML models for the forecasting: Polynomial Neural Network (PNN), Feed-Forward Neural Network (FFNN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Vanilla LSTM, Bi-LSTM and Stacked LSTM. The Bi-LSTM model showed the best performance. Additionally, a variety of input combinations were evaluated to determine the most prevailing wave parameters for forecasting wave energy flux, with wave height and wave direction being the most dominant parameters. Similarly, Ibarra-Berastegi et al. (2015) forecasted the wave energy flux employing as exogenous inputs the mean sea level pressure, surface wind vector, significant height of combined wind waves and swell, mean wave period and mean wave direction. The study compared the forecasting performance of a physical model (Wave Model – WAM) and different statistical models: Persistence, Analogues, Random Forest and Analogues combined with Random Forest. The RF model outperformed the others until a forecasting range between 3 and 7 h for open sea buoys, and 16-19 h for near-shore sites. Likewise, Bento et al. (2021) employed the Wave Energy Period, Significant Wave Height, wind speed, and wind direction as exogenous inputs to the proposed forecasting model: a Feed Forward Neural Network (FFNN) in which the hyperparameters were tuned using Moth Flame Optimization (MFO). The predictions included the forecast of the wave energy flux directly, and indirectly using the forecast of significant wave height and wave energy period. The findings suggested a better performance when forecasting the wave energy flux indirectly (using the forecasts of wave height and period) than when it was forecasted directly.

However, the previously highlighted works deal with point predictions. When forecasting the power delivered to the grid of WECs, it is crucial to calculate the uncertainties associated with the forecasting. Prediction Intervals (PIs) are one of the main ways to compute the uncertainties quantitatively. They have been widely employed for RES power forecasting, particularly for wind and solar. The construction of Prediction Intervals for solar and/or wind power forecasting has been done mainly using parametric and non-parametric methods. The parametric technique often assumes that the error distribution follows a particular distribution or a combination of multiple distributions, e.g., mixed Gaussian (Prusty and Jena, 2017), mixed skewness (De Giorgi et al., 2015), and beta (Von Loeper et al., 2020; Yuan et al., 2019; Liu and Xie, 2014). Nonetheless, the prediction output will be unsatisfactory if the assumed distribution is wrong (Liu et al., 2018; Li et al., 2021).

On the other hand, the non-parametric technique does not assume any distribution form of the prediction error; instead, it computes the distribution of the forecasting error directly from the prediction error value. One of the most common non-parametric methods used to calculate Prediction Intervals is Quantile Regression (QR) (Koenker and Bassett, 1978; Huang et al., 2017; Wang et al., 2022). Another

popular non-parametric method is the Lower Upper Bound Estimate (LUBE) (Khosravi et al., 2010) method (Khosravi and Nahavandi, 2013; Quan et al., 2013; Shi et al., 2017; Banik et al., 2020). The Non-Parameteric Kernel Density Estimation (NPKDE) is a well-diffused non-parametric technique. It has the benefits of simplicity and ease of implementation and has proven its capability to outperform other alternative methods (Liu et al., 2018). The NPKDE has been widely employed for building PIs for wind (Gu et al., 2021b; Niu et al., 2022; Xie et al., 2023; Li et al., 2021) and solar (Gu et al., 2021a; Liu et al., 2018) power forecasting.

Nonetheless, to the best of our knowledge, there is no reference in the literature to the implementation of Prediction Intervals for forecasting the power delivered to the grid of WECs. Previous studies in the wave energy field that have addressed the point forecasting problem of the output power of WECs have overlooked the grid integration of the devices. Grid integration is a vital stage in the wave-to-grid power conversion process and cannot be ignored. However, real-world data for this case is scarce or non-existent. In our previous work (Fontana Crespo et al., 2023), we presented a methodology based on a Long Short-Term Memory neural network to compute the point forecasting in very-short-term (i.e. 1 min) of the power delivered to the grid of the ISWEC device. The methodology comprised a realistic simulation model of the entire wave-to-grid power conversion process of the ISWEC, able to generate a realistic dataset of the power delivered to the grid. In addition, we studied the effectiveness of the downsampling technique by comparing the performance of the forecasts using different time-step aggregations.

In this work, we build upon the method we presented in Fontana Crespo et al. (2023) presenting a methodology for forecasting the power delivered to the grid of the ISWEC device in short-term horizons (i.e. from 15 min to 240 min). To this end, we design, tune and compare the performance of five Machine Learning models for time series point forecasting which have shown promising results for both wave power and time series forecasting: RF (Ni et al., 2019; Ibarra-Berastegi et al., 2015; Deberneh and Kim, 2018b; Casas and Lekube, 2023), SVR (Ni, 2021; Deberneh and Kim, 2018b; Casas and Lekube, 2023; Ni and Peng, 2023), LSTM (Ni et al., 2019; Mousavi et al., 2021; Ni, 2021; Neshat et al., 2022; Fontana Crespo et al., 2023; Nalamati, 2021a; Ni and Peng, 2023), 1D-CNN (Ni and Peng, 2023; Castangia et al., 2021; Kiranyaz et al., 2021), and TNN (Wen et al., 2022). The methodology includes the complete model of the wave-to-grid power conversion process of the ISWEC device, capable of producing a realistic dataset of the power transmitted to the grid. Furthermore, to study the effectiveness of downsampling, we downsample the original dataset sampled every 0.1 s in windows of 1 min, 3 min, 5 min and 15 min and compare the prediction performance of the different Machine Learning models. Lastly, to accurately quantify and compute the uncertainties of the forecasts, we build PIs for assessing the forecasting uncertainties employing the Non-Parameteric Kernel Density Estimation technique.

The novelties and contributions of this work are summarized as follows:

- Implementation of the complete wave-to-grid power conversion process of the ISWEC with the "realistic" *ISWEC Wave-to-Grid Simulator* (Fontana Crespo et al., 2023). This model allows the collection of a realistic dataset of the power transmitted to the grid. To the best of our knowledge, real-world data accounting for the grid integration of a WEC is scarce or non-existent.
- Design, optimize, and compare the performance of five Machine Learning algorithms for point forecasting: Random Forest, Support Vector Regression, Long Short-Term Memory Neural Network, 1 Dimensional Convolutional Neural Network and Transformer Neural Network. In the literature, these algorithms have demonstrated promising results for time-series and wave power forecasting. Our goal is to design and optimize highly specialized models for predicting in short-term horizons (i.e. from 15 min to 240 min) the power delivered to the grid of the ISWEC device. This approach allows obtaining accurate forecasts which are suitable for smart grid advanced monitoring and control techniques, like State Estimation and Demand-Response, unlocking the full integration of wave energy in a smart grid framework. To the best of our extent, our study is the first to forecast the power delivered to the grid of a WEC. Previous studies in the wave energy field that addressed the point forecasting problem of the output power of WECs, predicted the instantaneous output power of the WEC PTO, neglecting the grid integration of the device. This stage is vital in the wave-to-grid power conversion and must be considered.
- Compared to our previous work (Fontana Crespo et al., 2023), in which we presented a methodology for forecasting the power delivered to the grid of the ISWEC in the next minute, we employ a significantly more robust dataset of 130 days (versus a 30-h dataset that we employed in Fontana Crespo et al. (2023)). This "new" dataset better captures variability in wave power generation. Furthermore, the forecasting horizon is expanded from 1 to 240 min, enabling the application of the forecasts for advanced grid management techniques, like Demand/Response procedures.
- Compared to traditional physical approaches, employing the proposed Machine Learning algorithms for wave power prediction offers significant advantages in terms of execution time and cost (Mousavi et al., 2021).
- Analysis and quantification of forecasting uncertainties employing Prediction Intervals. These Prediction Intervals are built utilizing the Non-Parametric Kernel Density Estimation technique. To the best of our knowledge, we employ Prediction Intervals for computing forecasting uncertainties in the wave power prediction field for the first time.

3. Methodology

This section outlines our methodology, illustrated in Fig. 2. We propose a novel approach to forecast in short-term horizons (i.e. 15 min to 240 min) the power delivered to the grid of the ISWEC device. We used short-term forecasts to align with our main goal: obtain accurate forecasts suitable for smart grid advanced monitoring and control techniques (e.g. Demand-Response). For the predictions, we employed five Point Forecasting Machine Learning algorithms (*PF ML algorithms*) due to their widespread adoption and promising results in the literature for wave power and time series forecasting: RF (Ni et al., 2019; Ibarra-Berastegi et al., 2015; Deberneh and Kim, 2018b; Casas and Lekube, 2023), SVR (Ni, 2021; Deberneh and Kim, 2018b; Casas and Lekube, 2023; Ni and Peng, 2023), LSTM (Ni et al., 2019; Mousavi et al., 2021; Ni, 2021; Neshat et al., 2022; Fontana Crespo et al., 2023; Nalamati, 2021a; Ni and Peng, 2023), 1D-CNN (Ni and Peng, 2023; Castangia et al., 2021; Kiranyaz et al., 2021), and TNN (Wen et al., 2022). Nonetheless, to the best of our knowledge, real-world data of this case is rare or non-existent. Therefore, we modelled the ISWEC behaviour to obtain a realistic dataset through the *ISWEC Wave-to-Grid Simulator*. This module integrates the *ISWEC Wave-to-PTO model* (Bracco et al., 2016, 2011; Battezzato et al., 2015; Vissio et al., 2017) and the *PTO-to-Grid Electric system model* (Vissio, 2017). The output of the *ISWEC Wave-to-Grid Simulator* is the electric power delivered to the grid. This power served as input dataset for our forecasting algorithms. The *dataset* was partitioned into *training*, *test* and *PI Evaluation set*.

In the *point forecasting* module, the *training set* was employed to train the different *PF ML algorithms*. To analyse the efficacy of the downsampling technique, we aggregated the original dataset sampled every 0.1 s in different time steps (i.e. 1 min, 3 min, 5 min, and 15 min) before inputting it into the different *PF ML algorithms*. During the *test phase*, the *test set* was supplied to the *trained PF ML algorithms*. The output of the *trained PF ML algorithms* were the *point predictions*. The

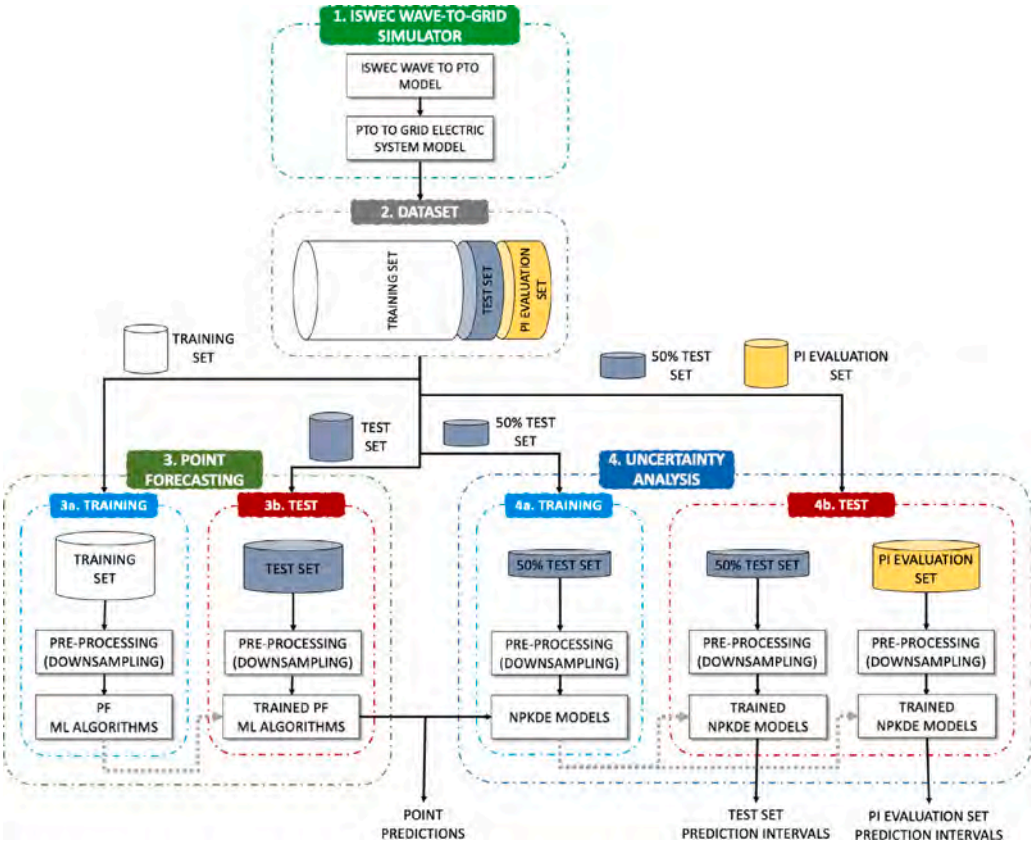


Fig. 2. Schema of the proposed methodology.

different *trained PF ML algorithms* were evaluated and compared based on their predictive performance on the *point predictions* of the *test set*.

Furthermore, in the *uncertainty analysis* module, we employed the *NPKDE models* to compute and quantify the uncertainties associated with the *point predictions*. Henceforth, in the *training phase*, we employed 50% of the *test set* and the corresponding *point predictions* of the different algorithms and calculated the associated errors. These errors were used to train the *NPKDE models* for computing and building the Prediction Intervals. Later, in the *test phase*, the *trained NPKDE models* were fed with the other 50% of the *test set* and evaluated based on their *prediction intervals* performance. Moreover, to further assess the *trained NPKDE models*, we compared the PIs performance on the *PI Evaluation set*.

3.1. ISWEC wave-to-grid simulator

This section will provide a detailed explanation of the *ISWEC Wave-to-grid Simulator* module depicted in Fig. 2 of the proposed methodology. ISWEC is a pitching-resonant floating WEC, that is able, by means of a gyroscopic system, to extract energy from the pitching motion generated by the action that the waves exert on the hull (Bracco et al., 2011; Battezzato et al., 2015). The main components besides the gyroscope and the floater are the gearbox, the eccentric mass, and the Power Take-Off (PTO). The PTO and the gyroscopic systems are responsible of the conversion process, and are enclosed in the hull, making them protected against the corrosive action of the water, and guaranteeing a higher durability. More in particular, the precession axis of rotation of the gyroscopic system is connected directly to the shaft of the PTO itself, which is integral with the hull structure, by means of a gearbox. In this way, the oscillating motion of the gyroscope (given by the precession effect generated by the flywheel characterizing the gyroscopic system) generated by the pitching motion of the floater is directly converted into electrical energy by means of the PTO generator. A schematic representation of this kind of devices is presented in Fig. 3.

Assuming a properly designed mooring system, ISWEC can be considered to be self-orienting with respect to the incoming wave. In this way, its main floater motions are pitch, surge and heave. Considering also an even number of gyroscopes (with opposite rotation speeds), we can also consider negligible effects discharged on the hull on the roll degree of freedom by the gyroscopes themselves. Because of that, being sway, roll and yaw negligible, they are not considered in the formulation of the *ISWEC Wave-to-grid Simulator*. With these assumptions, the vector of the coordinates that describe the floater motion is formulated in Eq. (1),

$$X_f(t) = \left[x(t), z(t), \delta(t)\right]^{\mathsf{T}} \tag{1}$$

where $x(t)$ and $z(t)$ are the floater surge and heave, while $\delta(t)$ is the rotation of the floater with respect to the inertial axis y. Following the principles of linear wave theory, and assuming small floater oscillations (as usually done Falnes, 2002), the floater motion follows the Cummins' equation (Cummins, 1962)

$$\left(M + A_\infty\right)\ddot{X}_f(t) + \int_0^t K_{rad}\left(t-\tau\right)\dot{X}_f(t)d\tau + B_v|\dot{X}_f(t)|\dot{X}_f(t) + K_h X_f(t) = F_{ext}(t) \tag{2}$$

where M is the inertia matrix that takes into account of both the hull and the internal components of ISWEC (e.g. gyroscope, structure, PTO, gearbox), A_∞ is the added mass at infinity frequency, K_{rad} is the radiation response functions matrix, K_h models the hydrostatic stiffness matrix, and $F_{ext}(t)$ represents the external forces (coming from the waves, the reaction of internal gyroscope systems and the ISWEC mooring) acting on the different degrees of freedom of the floater. The coefficients B_v that models the nonlinear viscous forces has been computed by means of fully-viscous *Computational Fluid Dynamics* (CFD)

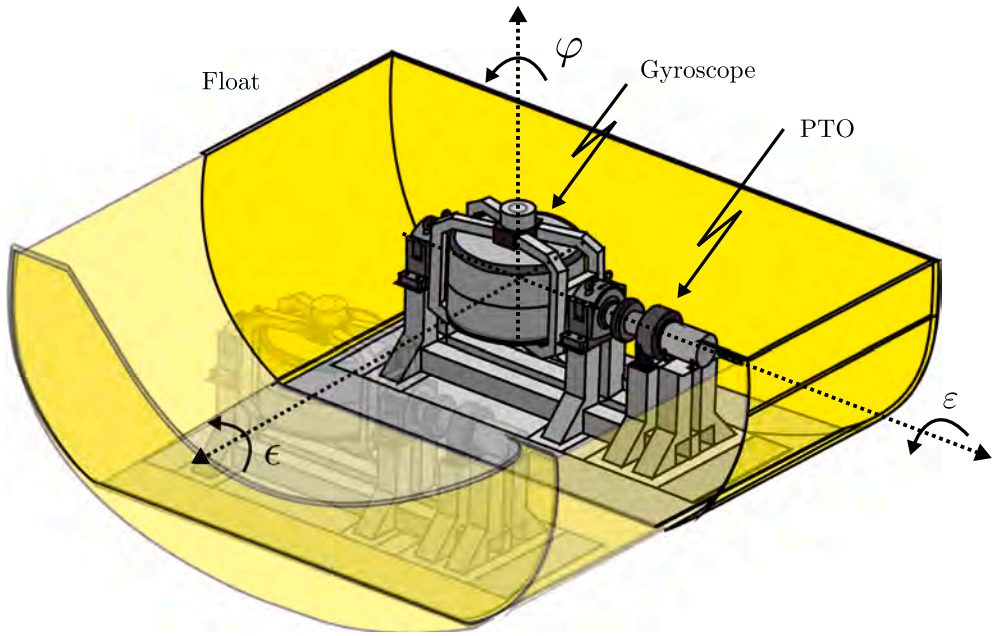


Fig. 3. Schematic representation of the ISWEC device and its main components.

simulations of the ISWEC hull. To make the model more computational efficient, the convolutional term is approximated through a state-space representation:

$$\int_0^t K_{rad}(t-\tau)\dot{X}_f(t)d\tau = F_{rad}(t) \simeq \begin{cases} \dot{\zeta}_{rad}(t) & = A_{rad}\zeta_{rad}(t) + B_{rad}\dot{X}_f(t) \\ F_{rad}(t) & = C_{rad}\zeta_{rad}(t) + D_{rad}\dot{X}_f(t), \end{cases} \tag{3}$$

where A_{rad}, B_{rad}, C_{rad}, and D_{rad} are the state-space matrices employed in the state-space approximation of radiation convolution term, $\zeta_r(t)$ is an array of additional states used to represent the convolution term F_r of the radiation forces acting on ISWEC hull. As mentioned before, the external forces $F_{ext}(t)$ are constituted by the contributions discharged on the hull by the waves, the reactions of the gyroscope system and the mooring. The external forces $F_{ext}(t)$ include contributions from the wave F_{wave}, the mooring F_{moor}, and the reactions F_{pend} that the pendulum and PTO system discharge onto the hull axes:

$$F_{ext}(t) = F_{wave}(t) + F_{moor}(t) + F_{pend}(t). \tag{4}$$

In Eq. (4), wave forces $F_{wave}(t)$ have been computed considering them as the superposition of several sinusoidal components that follows a *JONSWAP* (Joint North Sea Wave Observation Project) wave spectrum. To make the 1.5-year simulation reliable, the parameters describing the sea conditions (and, in this way, the JONSWAP spectrum) are modified hourly, as done in Pasta et al. (2023), following the records from a deployment site in the Mediterranean Sea available in ERA5 dataset (Hersbach et al., 2020). More in particular, considering a spectrum $S_\eta(\omega)$ of N components and component amplitudes that follows the *Random Amplitude scheme* presented in Merigaud and Ringwood (2018), the wave elevation can be written as:

$$\eta(t) = \sum_{n=1}^{N} A_n \sin(\omega_n t + \theta_n), \tag{5}$$

where the nth component has frequency ω_n, amplitude A_n follows a Rayleigh distribution with variance $2S_\eta(\omega_n)\Delta\omega$, and random phase θ_n defined as a sample from a uniform distribution between 0 and 2π. In this way, the wave force acting on the ith axis of the hull $F_{wave,i}(t)$ is given by:

$$F_{wave_i}(t) = \sum_{n=1}^{N} |f_{e_{i,n}}| A_n \sin\left(\omega_n t + \theta_n + \angle f_{e_{i,n}}\right). \tag{6}$$

In Eq. (6), $f_{e_{i,n}}$ is the excitation force nth coefficient (which depends upon the geometry of ISWEC submerged volume), at frequency ω_n and computed with respect to the ith axis, by means of Nemoh (Babarit and Delhommeau, 2015), an open-source tool for *Boundary Element Method* (BEM) computations. The mooring system of the simulator has been modelled following the *Quasi-static* (QS) approach presented in Paduano et al. (2020).

Considering a single gyroscope, its internal motion follows the equation[1]:

$$I\ddot{\varepsilon} + (I-J)\dot{\delta}^2 \sin(\varepsilon)\cos(\varepsilon) + T_{ctrl.} = J\dot{\phi}\dot{\delta}\cos(\varepsilon), \tag{7}$$

where I and J are respectively the moment of inertia of the gyroscope system along the PTO axis and the moment of inertia of the flywheel, while ε and ϕ are the angular displacement around the gyroscope and flywheel axes respectively. $T_{ctrl.}$ is the control torque that the PTO applies after the gearbox. The torque effectively applied by the PTO, before the gearbox is regulated by:

$$T_{PTO} = \frac{T_{ctrl}}{\tau_{gear}}, \tag{8}$$

where T_{PTO} is the PTO torque (coupled by the PTO angular velocity $\dot{\varepsilon}_{PTO}$), while τ_{gear} is the gearbox ratio. The instantaneous mechanical power absorbed by the PTO is, in this way:

$$P_{PTO,abs} = T_{PTO}\dot{\varepsilon}. \tag{9}$$

This mechanical power is then converted into electrical power by the PTO system, and then injected inside the grid. The PTO and the link between generators and the grid is modelled as in Vissio (2017). To reduce the computational effort in the 1.5-year simulation, the subsystems chain that transforms $P_{PTO,abs}$ into the power at the grid P_{grid} is substituted by a linear model obtained through a process of system identification (Ljung, 1999). Different tests have been performed in 50 operational wave conditions with the original grid model, and a transfer function fitting the resulting average frequency response is obtained from them and applied inside the ISWEC simulator (McKelvey et al., 1996).

[1] From now on, the dependence on t is dropped when clear from the context.

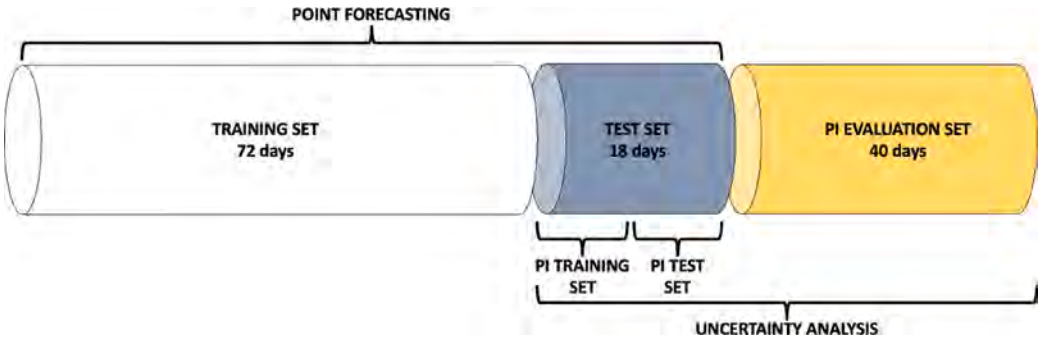


Fig. 4. Schema of the dataset partition.

The model employed within the ISWEC simulator has been extensively validated against experimental tests performed with different scaled versions of the device, as presented in Bracco et al. (2016) and Paduano et al. (2024). The average percentage errors observed within these tests are 8.8% on the pitch velocity $\dot{\delta}$ (which constitutes the main floater degree of freedom), 8.5% on the gyroscope velocity $\dot{\epsilon}$, and 11.2% on the delivered power (Bracco et al., 2016; Paduano et al., 2024).

3.2. Dataset

This section will introduce in depth the *Dataset* block shown in Fig. 2 of the proposed solution. We employed a dataset sampled every 0.1 s (10 Hz) over 1.5 years of the power delivered to the grid, acquired using the *ISWEC Wave-to-Grid Simulator*. The original dataset encompasses the measurements of the ISWEC sway, surge, heave, pitch, roll and yaw position, angular velocities and accelerations (summing 18 variables) and the power transmitted to the grid. For our specific case, we focused solely on the delivered power to the grid. From the 1.5-year power measurements, we chose the 130 days that account for more variability and were the most representative. Moreover, we employed the sliding non-overlapping moving average window technique to downsample the data into larger time steps. The downsampling technique aims to reduce the number of data points while preserving the essential characteristics of the original signal. We employed downsampling as a data preprocessing technique, working as a data smoother, decreasing signal peaks and fluctuations. However, this technique is strictly related to the use case for which the predictions are intended. For example, if the objective of the forecasts is State Estimation, where nowcasting is usually employed, the downsampling is restricted to prevent losing valuable information. Furthermore, downsampling is also "restricted" by the forecasting horizon. For instance, if we aim to perform forecasts 1 min ahead, the largest data aggregation feasible is 1 min. Contrariwise, if we downsample the data in steps of 3 min, the "smaller" step ahead we can predict is 3 min. Nonetheless, for the case under study, there is no consensus on a universally preferred size of the time steps (Banos et al., 2014). In the proposed methodology, we aim to compute forecasts in short-term horizons (i.e. 15 min to 240 min). Consequently, we performed downsampling at time steps of 1 min, 3 min, 5 min and 15 min as done in literature (Ni, 2021; Deberneh and Kim, 2018b), referring to these datasets as *D-1min*, *D-3min*, *D-5min*, and *D-15min*.

Fig. 4 illustrates the partition of the *dataset* for *point forecasting* and *uncertainty analysis*. We partitioned the 130-day dataset into 90 days for *point forecasting* and 40 days for validating the Prediction Intervals (*PI Evaluation set*). To study and compare the prediction performance of the *PF ML algorithms*, we divided the 90 days dataset for *point forecasting* into *training* (80% - 72 days) and *test set* (20% - 18 days). Additionally, the last 10% of the *training set* was used for validation during the *training phase*. Considering that the data is time series, where adjacent observations are often correlated, we employed the simple holdout method (Cerqueira et al., 2020) to split the dataset into *training* and *test set*. On the other hand, during *uncertainty analysis*, we employed 50% of the *test set* (hereinafter referred to as *PI training set*) for training the *NPKDE models*, and the remaining 50% of the *test set* (hereinafter referred to as *PI test set*) for testing purposes. Once again, the holdout method was used to divide the dataset into *PI training set* and *PI test set*. To further evaluate the performance of the *Trained NPKDE models*, we compared the Prediction Intervals performance on the *PI Evaluation set*. Table 3 presents the details of the data points of the *training*, *test*, and *PI Evaluation set*.

Table 3
Training, test and PI evaluation set size (number of samples) of the datasets.

	Dataset			
	D-1min	D-3min	D-5min	D-15min
Training set	103 680	34 560	20 736	6912
Test set	25 920	8640	5184	1728
PI evaluation set	56 160	18 720	11 232	3744

Before training the *PF ML algorithms* algorithms, it is essential to implement feature scaling to improve and accelerate the training process (Castangia et al., 2021). Therefore, we employed *Min-Max* normalization to scale the input dataset from 0 to 1:

$$x_{scaled} = \frac{x - min(x)}{max(x) - min(x)} \tag{10}$$

where x represents the vector of values to be scaled, and its minimum and maximum values are denoted by $min(x)$ and $max(x)$. In the proposed solution, the *Min* and *Max* values used for normalization were 0 and the maximum value found in the dataset, respectively. This is because the machine specifications (e.g. minimum, nominal, and maximum power delivered to the grid) were not available. In the case of a real-time implementation, since the entire dataset will not be available, the *Min* and *Max* values can be set following the device specifications.

3.3. Point forecasting

This section will provide a detailed explanation of the *Point Forecasting* module of the proposed methodology depicted in Fig. 2. For predicting the ISWEC output power delivered to the grid, we designed and fine-tuned five distinct state-of-the-art *PF ML algorithms*: Random Forest (RF), Support Vector Regression (SVR), Long Short-Term Memory (LSTM) neural network, 1 Dimensional Convolutional Neural Network (1D-CNN), and Transformer Neural Network (TNN). As shown in Fig. 2, during the *training phase*, the *PF ML algorithms* were trained with the corresponding *training set* after *pre-processing (downsampling)*. To assess the prediction horizon over which the models can still achieve acceptable forecasting errors, we implemented the multiple output approach, denoting a singular network with n outputs, where n corresponds to the number of steps ahead to predict (Kline,

Table 4
Selected models random forest.

Dataset	Time Lag	Max Features	Min Samples	Training time [s]
D-1min	240	0.33	0.01	283.54
D-3min	160	0.33	0.01	9.57
D-5min	84	0.33	0.01	1.49
D-15min	28	0.33	0.01	0.25

Table 5
Selected models support vector regression.

Dataset	Time Lag	Kernel	C	ϵ	Training time [s]
D-1min	180	rbf	100	0.1	3744.34
D-3min	60	rbf	100	0.001	669.06
D-5min	36	rbf	100	0.001	206.76
D-15min	12	rbf	100	0.001	12.43

2004). Therefore, we conducted a multi-step ahead prediction with a prediction horizon equal to 240 min, i.e. 4 h. The maximum prediction horizon was selected arbitrarily, contemplating the challenges of predicting numerous time steps ahead while maintaining an acceptable forecasting error. We identified the optimal architecture for each model via a trial-and-error approach with a grid search algorithm, where we evaluated several possible configurations. We employed the grid search because it is one of the most diffused methodologies in the literature for hyperparameters optimization (Bergstra et al., 2011) due to its simplicity and exhaustiveness compared to more intricate methodologies (Khalid and Javaid, 2020; Shekar and Dagnew, 2019). During the *test phase*, the *Trained PF ML algorithms* were fed with the downsampled *test set* accordingly to obtain the *point predictions*, as shown in Fig. 2.

3.3.1. Random forest

Random Forest (RF) is a regression Machine Learning algorithm derived from decision trees. RF regression combines the forecasts of many decision trees to compute a more precise prediction than those obtained with any singular model. This combination, known as model ensembling, harnesses the collective strength of multiple models to improve the overall performance. In the RF algorithm, each base tree is built independently through a deterministic algorithm that selects a random subset of input data from the training set. For a detailed description of the RF structure, refer to Breiman (2001).

We implemented Random Forest in Python using the *sklearn* library (Pedregosa et al., 2011). The RF algorithm has four main parameters to tune: (i) *number of estimators*, which represent the number of trees in the forest; (ii) *max features* that limit the number of features a tree can use; (iii) *min samples leaf* which determine the smallest size of the final nodes of the trees; and (iv) *number of regressors*. We tuned the hyperparameters through a trial-and-error approach using a grid search. For the number of regressors (i.e. time lag), we tested values ranging from 2 to 600. The number of estimators was set to be equal to the number of regressors since different values for the number of estimators did not show substantial improvement. The minimum number of samples in a leaf is a key parameter employed as a stopping criterion to prevent overfitting. We searched the min samples between 1% and 100% of the total samples. For max features, we tested using 33% and 100% of the input features.

Table 4 summarizes the chosen Random Forest models hyperparameters for each of the evaluated datasets. As presented, all models employed the same *Max Features* and *Min Samples* hyperparameters. Moreover, it is important to highlight the significant reduction of the *training time* as the downsampling increased. Furthermore, the number of data points required for one forecast (i.e. *Time Lag*) also decreased for bigger downsamplings: the *D-1min* model required 240 points; the *D-3min* 160 points; the *D-5min* required 84 points; and the *D-15min* model 28 points. This number of data points refers to the data after being downsampled, which means that, for example, the 240 points for the *D-1min* equal 240 min of past information. If we consider the time period to which these past observations correspond, it can be seen that the *D-3min* model employed the most past information (160 points equals 480 min). On the contrary, the *D-1min* considered only the most recent information (240 min). The *D-5min* and *D-15min* models used past data spanning the previous 420 min.

3.3.2. Support vector regression

Support Vector Regression (SVR) is a non-linear regression algorithm that aims to find a hyperplane in a high-dimensional space that best represents the relationship between the inputs and outputs. To this end, it maps the input data x into a high-dimensional feature space through non-linear mapping and computes a linear regression within this feature space (Mukherjee et al., 1997). More details of the structure of the SVR employed in this study can be found in Vapnik et al. (1996).

The SVR was implemented in Python using the open-source library *MSVR* (Bao et al., 2014). The *MSVR* is a generalization of the standard SVR implementation that enables multi-step ahead time series prediction. The *MSVR* has 4 parameters to be optimized: (i) *kernel*, which is the function that maps the points into the higher-dimensional feature space; (ii) C is a penalty parameter for misclassified data points during the training process which regulates the trade-off between a smooth decision boundary and correct classification of training points; (iii) ϵ is a parameter that defines a margin of tolerance for not penalizing certain training errors; and (iv) *number of regressors*.

As before, we adopted a grid search approach to find the best hyperparameters for the SVR. For the kernel, we tested the Radial Basis Function (RBF), Polynomial, and Sigmoid kernel. Furthermore, we searched the number of regressors ranging from 2 to 600. For the penalty parameter C, we investigated values ranging from 0.001 to 1000. Additionally, we tested values for ϵ in the interval between 0.001 and 10. Table 5 reports the selected parameters for the different SVR models. All models showed the best results with the *rbf* kernel, C equal to 100 and ϵ equal to 0.001 (except for the 1 min dataset with ϵ 0.1). As for the RF models, the training time significantly decreased with bigger downsamplings. Additionally, the *Time Lag* (i.e. number of regressors) also diminished as the downsampling increased. However, considering the timespan of these past observations, the four models considered 180 min of past information.

3.3.3. Long short-term memory neural network

The Long Short-Term Memory Neural Network (LSTM), an improvement of traditional recurrent neural networks, was designed to address the "vanishing gradient" issue (Hochreiter and Schmidhuber, 1997). This issue emerges during the training phase of neural networks employing backpropagation techniques. The LSTM architecture is especially well-suited for time-series prediction due to its internal structure, enabling connections across more than 1000 time steps and mitigating the issue of backpropagating errors between layers (Guo et al., 2016). A detailed description of the LSTM structure is provided in Hochreiter and Schmidhuber (1997).

There are two main LSTM models for multi-step forecasting: Vector Output model (LSTM VEC) and Encoder-Decoder model (LSTM E-D). In the first model, as the name indicates, the LSTM network outputs a vector directly. On the other hand, the Encoder-Decoder model has been designed for sequence-to-sequence (seq2seq) problems (e.g. text translation from one language to another) and consists of two components: the first component reads the input sequence, encoding it into a fixed-length vector, while the second model decodes the fixed-length vector, generating the predicted sequence (Brownlee, 2018).

In our work, we employed the *Keras* library with Tensorflow backend (Chollet et al., 2015) for implementing the LSTM neural network in Python. We opted for the hyperbolic tangent (*tanh*) activation function as the hidden layer activation function due to its widespread adoption

Table 6
Selected models LSTM vector.

Dataset	Hidden layers	Batch size	Time Lag	Neurons layer 1	Neurons layer 2	Training time [s]
D-1min	2	128	180	128	64	527.82
D-3min	1	128	280	16	0	97.71
D-5min	1	32	72	32	32	231.45
D-15min	1	16	20	48	0	81.27

Table 7
Selected models LSTM encoder-decoder.

Dataset	Hidden layers	Batch size	Time Lag	Neurons layer 1	Neurons layer 2	Training time [s]
D-1min	2	128	240	96	48	543.75
D-3min	1	128	320	256	0	149.50
D-5min	1	128	36	32	0	72.61
D-15min	1	128	40	256	0	21.77

and proven efficacy (Kline, 2004). On the contrary, we chose a linear activation function for the output layer. We explored various configurations for the overall architecture, implementing both the Vector Output model and the Encoder-Decoder one, utilizing in each case one or two hidden layers. For tuning the hyperparameters, i.e. number of regressors, batch size and number of LSTM units, we exploited the same grid search approach previously highlighted. We searched the number of regressors in the range from 2 to 600. Additionally, we investigated the batch size in the range from 16 to 128. For the LSTM units, we studied values ranging from 8 to 256. For the models employing two hidden layers, we tested two distinct configurations for the second hidden layer: one with the identical number of units of the first hidden layer, and the other with half the units. We employed the Adaptive Moment Estimation (Adam Optimizer) (Kingma and Ba, 2014) as the optimization algorithm. Furthermore, we implemented the early-stopping technique (Prechelt, 2002) to mitigate overfitting. This technique interrupts the training phase if there is no improvement in the validation set after a specific number of steps. Besides mitigating overfitting, the Early-stopping technique substantially reduces the training time.

Tables 6 and 7 report both hyperparameters for the LSTM Vector (LSTM VEC) and LSTM Encoder-Decoder (LSTM E-D) models, respectively. The only model implementing 2 hidden layers in both cases was the one for the *D-1min* dataset. This can be due to the complexity of this model in terms of inputs and outputs. However, the *Time Lag* of the *D-1min* model was smaller than for the *D-3min* models. This increasing behaviour in the *Time Lag* for bigger downsampling also occurred for the *D-15min* model compared to the *D-5min* model of LSTM E-D case. Moreover, if we take into account the time period corresponding to the past observations, the *D-3min* model used the most extensive historical data in both the LSTM VEC and LSTM E-D cases, employing 840 min and 960 min of past data, respectively. On the other hand, the *D-1min* model was the model which focused on the most recent information (180 min) for the LSTM VEC case, whereas the *D-5min* model (180 min) for the LSTM E-D case. For the LSTM Vector case, the batch size was reduced as the downsampling increased. Instead, in the Encoder-Decoder case, all models employed the same batch size (128).

3.3.4. 1 dimensional convolutional neural network

The 1 Dimensional Convolutional Neural Network (1D-CNN) is an evolution of the CNN developed for modelling 1-dimensional inputs. It has demonstrated exceptional performance in various signal processing applications due to its remarkable ability to extract meaningful features from sequential data (Kiranyaz et al., 2021). The filter is the central component of the feature selection process. It is a feature detector trained to identify specific patterns within the input data. As the filter slides across the input sequence, it activates the corresponding neuron in the feature map upon detecting a match for the specific pattern. This operation is known as convolution and involves a series of multiplications between the input values and the filter. Typically, convolutional layers consist of numerous filters operating across multiple input channels, allowing the neural network to identify diverse patterns within the input data. These convolutional layers can be stacked sequentially to create a deep architecture, with each layer building upon the features identified by the preceding layer. Usually, a pooling layer follows the convolutional layers. The pooling layer is in charge of diminishing the volume of information transferred from the previous layer. Lastly, one or more fully connected layers are appended at the end of the neural network to interpret the extracted features. More details of the 1D-CNN structure employed in this work can be found in Kiranyaz et al. (2021).

Table 8
Selected models 1D-CNN.

Dataset	Batch size	Time Lag	Pool	Kernel	Filt. Lay. 1	Filt. Lay. 2	Dense	Training time [s]
D-1min	128	180	2	3	16	8	128	43.23
D-3min	128	40	2	3	16	8	32	23.54
D-5min	64	36	2	5	16	8	32	31.82
D-15min	128	12	2	3	16	16	16	14.80

We implemented the 1D-CNN models using the *Keras* library with Tensorflow backend (Chollet et al., 2015). We used the hyperbolic tangent function as an activation function for each convolutional layer. Once again, we implemented a grid search approach to find the best hyperparameter configuration. For the batch size, we searched in the range 16 to 128. Moreover, we tested the models using from 4 to 600 past observations. We investigated models using one or two 1-dimensional convolutional layers. For each convolutional layer, we studied kernel sizes of 3 and 5. Additionally, we searched the number of filters ranging between 16 and 128. For the models employing two convolutional layers, we tested two configurations in the second layer as explained in Section 3.3.3: one with an equal number of filters as in the first layer, and the other with half the filters. After the convolutional layers, a pooling layer was added. We investigated values of the pooling window in the range of 2 to 8. Finally, we append a fully connected layer at the end of the model. For this layer, we searched for units ranging from 16 to 512. Like in the LSTM network, we employed Adam as the optimization algorithm, and the early stopping technique to prevent overfitting.

Table 8 presents the hyperparameters for the 1D-CNN models. All models implemented the same *Pool* (2) and *kernel size*(3) parameter, as well as for the number of filters in the first hidden layer (16). Moreover, there can be seen a decreasing trend in the number of units in the fully connected layer *Dense*. Furthermore, the number of data points required for one forecast (i.e. *Time Lag*) also decreased for bigger downsamplings: the *D-1min* model required 180 points, the *D-3min* 40 points, *D-5min* 36 points and the *D-15min* 12 points. Nonetheless, analysing the equivalent time period of these past values, the *D-3min* model considered the least past information (120 min), whereas the other three models employed 180 min of past information, as occurs in the SVR case.

3.3.5. Transformer neural network

The Transformer Neural Network (TNN), initially proposed by Vaswani et al. (2017), emerged as a revolutionary neural network architecture designed to address machine translation problems. The Transformer has been widespread in various application domains with the need to process large input sequences, like classification algorithms and time-series forecasting (Li et al., 2019). The core of the Transformer's success relies on the self-attention mechanism and the multiple-head attention, which enable the model to capture intricate patterns and dependencies within the input sequence. The self-attention mechanism in the Transformer allows the model to focus on different

Table 9
Selected models transformer.

Dataset	Num. Transf. Blocks	Batch size	Time Lag	Num. Heads	Head size	Filter dim	MLP units	Training time [s]
D-1min	2	128	180	8	256	32	256	1490.85
D-3min	2	128	40	8	256	16	128	83.34
D-5min	4	64	36	8	128	32	256	213.15
D-15min	2	64	12	8	128	16	128	33.96

time steps in the input sequence when making predictions for a specific time step. This mechanism enables the model to capture dependencies between distant time points effectively, addressing a crucial aspect of time-series forecasting where long-range relationships play a significant role. Additionally, the Transformer utilizes multiple-head attention to enhance its ability to capture diverse temporal patterns. This involves running the self-attention mechanism in parallel across multiple attention heads. Each head learns different aspects and dependencies within the time series simultaneously, contributing to a more comprehensive understanding of complex temporal relationships. This parallel learning is particularly beneficial in capturing various patterns and trends in time-series data. For more details on the TNN structure, refer to Li et al. (2019).

In our work, the transformer was implemented in Python using *Keras* with Tensorflow backend (Chollet et al., 2015). Like in the previous models, we tuned the hyperparameters with a grid search approach. We searched the batch size in the range 16–64. For the number of lagged inputs, we investigated values in the range 2–600. Moreover, we tested models using 2 or 4 transformer blocks. For the multi-head attention module, we used 8 attention heads. Instead, for the size of each attention head, we searched in the range from 128 to 512. Additionally, we searched values in the range 16 to 128 for the Feed Forward part. We trained the transformer using the Adam Optimizer and a learning rate of 0.000001. As in the other models, we used the early stopping mechanism to prevent overfitting. Table 9 reports the hyperparameters for the Transformer models. As it can be seen, the batch size and the size of the attention head were reduced for the *D-5min* and *D-15min* models compared to the *D-1min* and *D-3min* models. Moreover, it is important to highlight the significant reduction in training time if we compared the *D-1min* model with the others. Similarly, the *Time Lag* (i.e. number of regressors) also decreased significantly comparing the *D-1min* model with the others. Nonetheless, if we evaluate in terms of the time interval that these past observations correspond to, the *D-3min* model focused on the shortest past data interval (120 min). Instead, the other three models covered 180 min of previous information. Last but not least, the only model implementing 4 transformer blocks was the *D-5min* model, which can be one of the reasons for the increasing training time if compared with the *D-3min* model. Regarding the *D-5min* case, we chose the model with four transformer blocks since it was an optimum trade-off between computational cost and accuracy: despite being more expensive than the models with two transformer blocks, the improvement in accuracy justifies this choice. For this case, to determine if it was necessary to use more than 4 transformer blocks, we decided to extend the grid search for models employing 6 transformer blocks. Nonetheless, the performance of the models with 4 transformer blocks was better.

3.4. Uncertainty analysis

In this section, we will provide a detailed explanation of the *Uncertainty Analysis* module of the proposed solution illustrated in Fig. 2. The computation of the *Prediction Intervals* of the ISWEC delivered power to the grid means determining the upper and lower limits of the *point predictions* of this power and computing the *Prediction Intervals* for a particular confidence level. The calculation of precise *PIs* provides decision-makers with highly valuable forecast information and mitigates the operational hazards of power systems (Niu et al., 2022). Hence, we analysed and calculated the point prediction errors using the Non-Parameteric Kernel Density Estimation (NPKDE) method. Subsequently, we combined the *point predictions* and the computed interval estimations (using *point predictions* errors) to obtain the lower and upper bounds of the *Prediction Intervals* of the ISWEC delivered power to the grid for various confidence levels.

Non-Parameteric Kernel Density Estimation (NPKDE) is a non-parametric data-driven method for calculating probability density functions. It is known for its efficiency and simplicity (Khorramdel et al., 2018) and is particularly well-suited for computing the probability density when the underlying data distribution is unknown (Zhao et al., 2018). The choice of the kernel function is a crucial stage for an effective implementation of the NPKDE. We adopted the Gaussian kernel since it provides high accuracy and speed (Qin et al., 2011). Eq. (11) presents the expression of the Gaussian kernel.

$$g(x) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{11}$$

where, $g(x)$ is the Gaussian kernel function at any given point x, μ is the mean, and σ is the standard deviation.

Eq. (12) reports the probability density distribution calculated using NPKDE.

$$f(x) = \frac{1}{N}\sum g\left(\frac{x-x_i}{h}\right) \tag{12}$$

where $f(x)$ is the probability density function at any given point x, N refers to the number of samples, h is the bandwidth coefficient, and x_i is the ith sample.

The ISWEC point forecasting error was calculated as the difference between the *point prediction* value $P_{forecast}$ and the actual ISWEC power value P_{real} at a specific time point, as presented in Eq. (13).

$$e = P_{forecast} - P_{real} \tag{13}$$

The confidence level of ISWEC point prediction error was computed as shown in Eq. (14).

$$P(e_{low} < e < e_{up}) = 1 - \sigma \tag{14}$$

where the range $[e_{low}, e_{up}]$ is the Prediction Interval; e_{low} denotes the lower boundary of the PI, and e_{up} is the upper boundary. $P(e_{low} < e < e_{up})$ denotes the probability that the forecasting error is in the range $[e_{low}, e_{up}]$. Therefore, the Prediction Interval of ISWEC power forecasting is $[P_{forecast} - |e_{up}|,\ P_{forecast} + |e_{low}|]$.

4. Results and discussion

In this section, we present the experimental results obtained by employing the proposed methodology. Firstly, we introduce the statistical metrics employed for comparing the different models. Subsequently, we show the results obtained for the point forecasting. Later, we describe and analyse the results of the associated forecasting uncertainties of the different models.

4.1. Performance metrics

To assess the different *point forecasting* model's performance, we relied on three metrics frequently employed in literature to compute the similarities between forecasted and observed time series for regression and descriptive analysis (Gueymard, 2014): (i) the *Mean Absolute Difference* (MAD); (ii) the *Root Mean Squared Deviation* (RMSD) and (iii) the *Coefficient of determination* (R^2). MAD determines the absolute

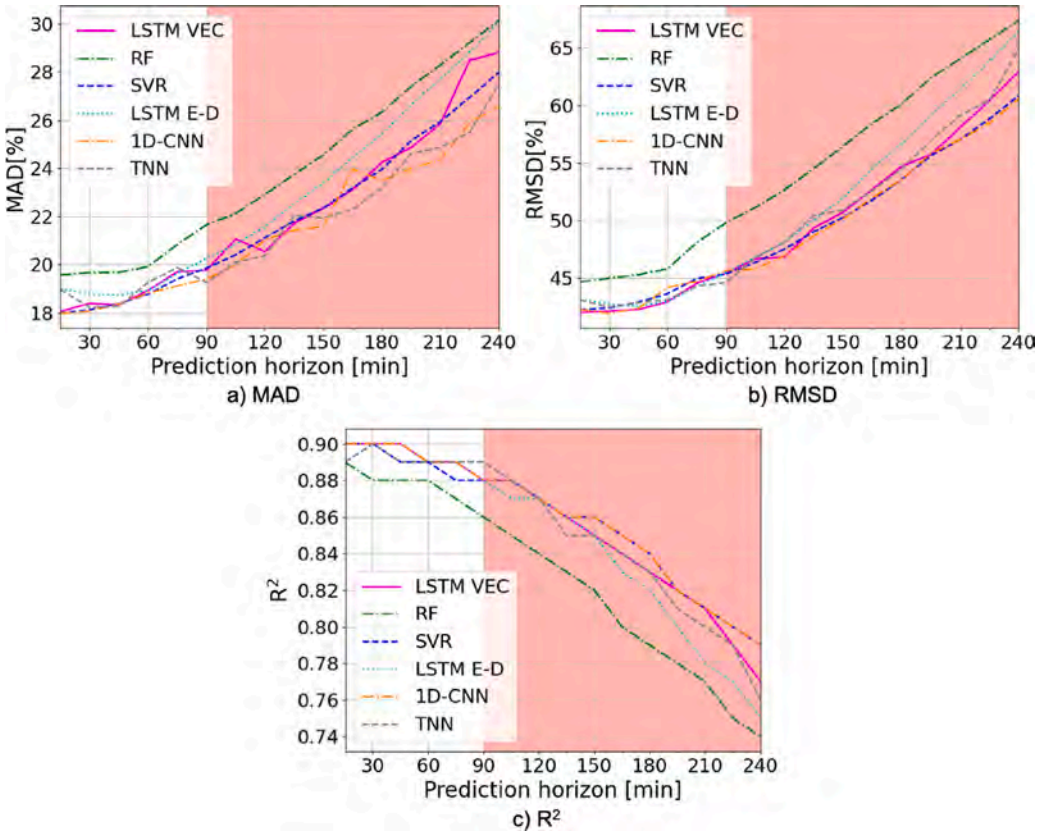


Fig. 5. Comparison of models point forecasting performance for the D-1 min dataset.

difference between forecasted and observed values. RMSD computes the standard deviation of the difference between the observed and the forecasted values. R^2 calculates the percentage of the observed value variance that is determined by the forecasted value. The three metrics are described by the following equations:

$$MAD = \frac{100}{\overline{y}_{test}} \frac{\sum_{i=1}^{n} |y_{fore,i} - y_{test,i}|}{n} \tag{15}$$

$$RMSD = \frac{100}{\overline{y}_{test}} \sqrt{\frac{\sum_{i=1}^{n} (y_{fore,i} - y_{test,i})^2}{n}} \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{test,i} - y_{fore,i})^2}{\sum_{i=1}^{n} (y_{test,i} - \overline{y}_{test})^2} \tag{17}$$

where n is the number of predictions, $\overline{y}_{test}$ represents the mean value, y_{fore} and y_{test} are the forecasted and observed values, respectively. The MAD and RMSD are expressed in percentages. Lower values for MAD and RMSD imply a lower error and, therefore, indicate better performance. Instead, R^2 computes the correlation between observed and predicted values, where 1 indicates a complete correlation, while smaller values denote a weaker correlation.

To evaluate and quantify the model's Prediction Intervals performance, alternative metrics from the ones presented before must be employed. The effectiveness of probabilistic models is commonly assessed based on their reliability and sharpness (Van der Meer et al., 2018). The *Prediction Interval Coverage Probability* (PICP) is employed to compute the model's reliability. It measures the rate at which the observed wave power values lie within the Prediction Interval with a particular confidence level as shown in Eq. (18). Higher PICP means that more values fall into the Prediction Interval. Ideally, the PICP value should be higher than the confidence level, but roughly equivalent. Instead, the *Prediction Interval Normalized Average Width* (PINAW), a complementary metric, indicates the width of the Prediction Interval by computing the difference between the lower and upper bounds under a particular confidence level, as presented in Eq. (19). In contrast to the PINAW and the PICP, the *Continuous Rank Probability Score (CRPS)* measures both the sharpness and the reliability of a probabilistic model. The CRPS computes the difference between the forecasted and the observed cumulative distribution (see Eq. (20)).

$$PICP = \frac{100}{T} \sum_{t=1}^{T} \epsilon_t \longrightarrow \epsilon_t \begin{cases} 1 \text{ if } y_{test,t} \in [L_t, U_t] \\ 0 \text{ if } y_{test,t} \notin [L_t, U_t] \end{cases} \tag{18}$$

$$PINAW = \frac{100}{TR} \sum_{t=1}^{T} (U_t - L_t) \tag{19}$$

$$CRPS = \frac{100}{T} \sum_{t=1}^{T} \int_{-\infty}^{\infty} (F_{fore,t}(x) - F_{test,t}(x))^2 dx \tag{20}$$

where $y_{test,t}$ is the measured value at time t, T is the length of the time-series dataset, R is the difference between the maximum and the minimum measured value, U_t and L_t are the upper and lower bound of the Prediction Interval at time-step t, $F_{fore,t}(x)$ is the Cumulative Distribution Function (CDF) of the predicted variable of interest x (i.e. power delivered to the grid) and $F_{test,t}(x)$ is the CDF of the observed variable of interest x at time-step t. $F_{test,t}(x)$ is a cumulative-probability step function that jumps from 0 to 1 when the predicted value is equal to the observed one (Lauret et al., 2019). The average of the squared difference between the two CDFs is computed over the T pairs of forecasts and observations. The CRPS score rewards the concentration of probability around the step function positioned at the real value. It is important to highlight that the CRPS is a negatively oriented metric, which means smaller values indicate better performance. Additionally, it shares the dimension of the forecasted variable.

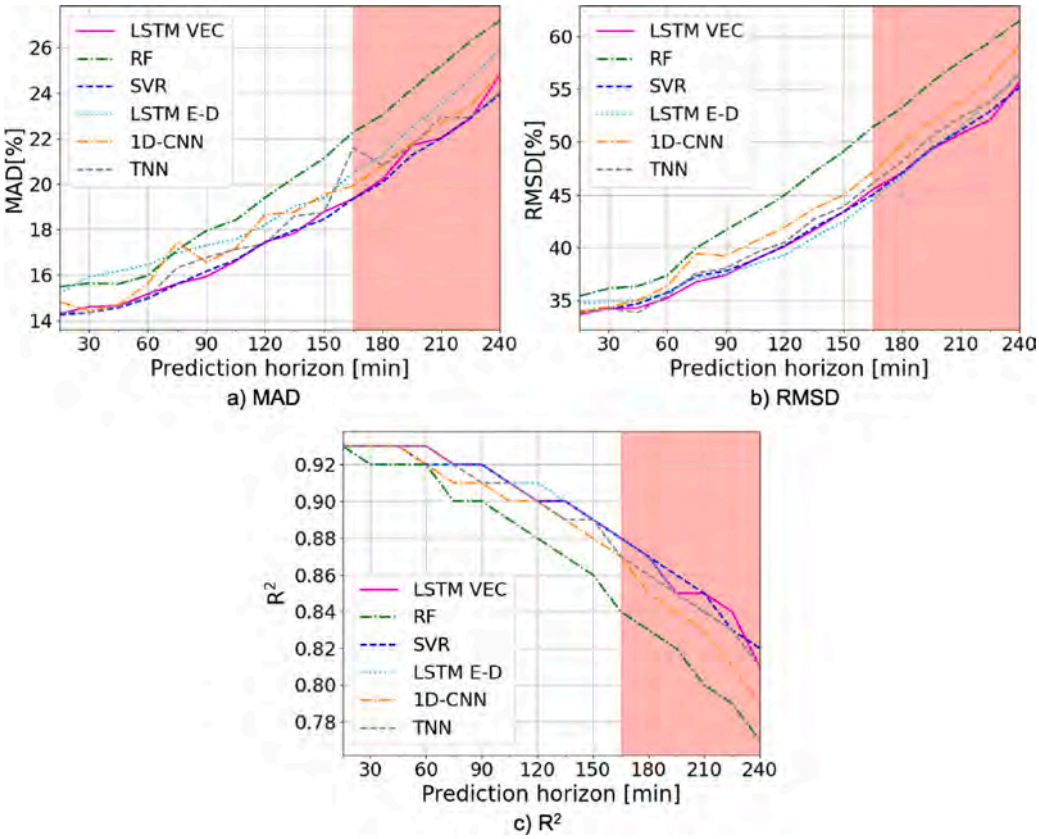


Fig. 6. Comparison of models point forecasting performance for the D-3 min dataset.

4.2. Model evaluation

4.2.1. Point forecasting models

For each of the six ML models for the different datasets (i.e. *D-1min*, *D-3min*, *D-5min* and *D-15min*), we selected the hyperparameter configuration that offered the best performance as described in Section 3.3. Afterwards, we compared the models based on their prediction performance in the *test set*.

Fig. 5 provides a visual comparison of the prediction performance of the different models in terms of MAD, RMSD and R^2 for the *D-1min* dataset. The plots indicate a general uptrend in the forecasting errors as the prediction horizon increases. We arbitrarily decided to set a 20% MAD threshold to define an upper limit for acceptable prediction performance. Above this threshold, we consider an excessive degradation in the model's prediction performance. Based on this criterion, we can affirm that the present methodology for the *D-1min* dataset showed acceptable results up to a forecasting horizon of 90 min (1 h and 30 min), with the TNN and the 1D-CNN models having a maximum error of 19.26% and 19.40% in terms of MAD, respectively. The red area in Fig. 5 highlights the forecasting horizon in which the model performance is worse than the established limit. Therefore, we will focus our analysis only on the forecasting horizon with acceptable results, i.e. from 15 min up to 90 min. The numerical results of the forecasting errors illustrated in Fig. 5 are reported in Table 10. Overall, the 1D-CNN model performed well, especially in terms of MAD and R^2. For the forecasting horizon of interest (i.e. 15 min to 90 min), the prediction errors for this model ranged from 17.99% to 19.40% in terms of MAD, from 42.23% to 45.63% in terms of RMSD, and from 0.90 to 0.88 in terms of R^2. On the contrary, the plots of the three metrics in Fig. 5 clearly depict that the RF model was the worst-performing model. The LSTM VEC demonstrated comparable performance to the 1D-CNN model for a prediction horizon of up to 60 min, even outperforming it at some points. Nonetheless, as reported in Table 10, the training time of the 1D-CNN (43.23 s) was significantly less than for the LSTM VEC model (527.82 s), making the first a more suitable option. The SVR model also showed comparable performance to the 1D-CNN model in terms of MAD for the same horizon. However, the training time of the SVR (3744.34 s) was considerably larger than that of the 1D-CNN model (43.23 s), making the last a better alternative. Interestingly, the TNN presented comparable performance to the 1D-CNN in terms of R^2 and RMSD for a prediction horizon larger than 45 min, even outperforming it in some sporadic cases. Moreover, the TNN outperformed the 1D-CNN model in terms of the three metrics for a forecasting horizon of 90 min. Nonetheless, the 1D-CNN demonstrated better performance than the TNN in terms of MAD for the other points of the same horizon, and its training time (43.23 s) was significantly shorter than for the TNN (1490.85 s), making the 1D-CNN a better choice. Similarly, the LSTM E-D model also displayed comparable performance to the 1D-CNN in terms of R^2 and RMSD for the same prediction horizon. Once again, the 1D-CNN seems a better alternative since it outperformed the LSTM E-D in terms of MAD for the same horizon, and the training time (43.23 s) was lower than that of the LSTM E-D (543.75 s).

The prediction performance of the different models for the *D-3min* dataset is depicted in Fig. 6. Like in the *D-1min* case, the plots reveal an evident degradation of the models' forecasting accuracy as the prediction horizon increases. According to the 20% limit in terms of MAD, the models of the proposed methodology for the *D-3min* dataset achieved acceptable performance up to a prediction horizon of 165 min (2 h and 45 min). This forecasting horizon is 75 min larger than the horizon obtained for the *D-1min* dataset, demonstrating that the models for the *D-3min* case can provide satisfactory forecasts for longer prediction horizons than the *D-1min* case. Table 11 presents the prediction errors shown in Fig. 6. Overall, the LSTM VEC and the SVR models demonstrated promising results regarding the three metrics, particularly in terms of MAD. The prediction errors for the SVR model varied between 14.25% and 19.35% in terms of MAD, from 33.90% to

Table 10
Models point forecasting performance based on MAD [%]. RMSD [%] and R^2 for the D-1min dataset.

Model	Training Time [s]	Metric	Prediction Horizon [min]					
			15	30	45	60	75	90
RF	283.54	MAD	19.58	19.67	19.68	19.92	20.84	21.66
		R^2	0.89	0.88	0.88	0.88	0.87	0.86
		RMSD	44.64	44.96	45.27	45.78	48.06	49.81
SVR	3744.34	MAD	18.00	18.13	18.39	**18.78**	19.41	19.86
		R^2	**0.90**	**0.90**	0.89	**0.89**	0.88	0.88
		RMSD	42.22	42.39	42.89	43.62	44.93	45.33
LSTM VEC	527.82	MAD	18.06	18.40	18.34	18.93	19.71	19.77
		R^2	**0.90**	**0.90**	**0.90**	**0.89**	**0.89**	0.88
		RMSD	**41.97**	42.11	**42.24**	**42.91**	44.53	45.41
LSTM E-D	543.75	MAD	19.00	18.80	18.75	18.93	19.64	20.26
		R^2	0.89	**0.90**	**0.90**	**0.89**	**0.89**	0.88
		RMSD	43.12	42.72	42.53	42.94	**44.26**	45.46
1D-CNN	43.23	MAD	**17.99**	**18.07**	18.38	18.82	**19.12**	19.40
		R^2	**0.90**	**0.90**	**0.90**	**0.89**	**0.89**	0.88
		RMSD	42.23	**41.92**	42.34	44.17	44.69	45.63
TNN	1490.85	MAD	18.98	18.24	**18.27**	19.27	19.88	**19.26**
		R^2	0.89	**0.90**	0.89	**0.89**	**0.89**	**0.89**
		RMSD	43.08	42.46	42.85	43.12	**44.26**	**44.59**

Table 11
Models point forecasting performance based on MAD [%]. RMSD [%] and R^2 for the D-3min dataset.

Model	Training Time [s]	Metric	Prediction Horizon [min]									
			15	30	45	60	75	90	105	120	135	150	165
RF	9.57	MAD	15.48	15.63	15.60	15.97	17.09	17.95	18.42	19.40	20.29	21.14	22.30
		R^2	**0.93**	0.92	0.92	0.92	0.90	0.90	0.89	0.88	0.87	0.86	0.84
		RMSD	35.37	36.09	36.31	37.32	39.94	41.64	43.25	44.98	47.16	49.11	51.43
SVR	669.06	MAD	**14.25**	**14.35**	**14.56**	**14.97**	**15.59**	16.16	16.66	17.44	17.97	**18.45**	**19.35**
		R^2	**0.93**	**0.93**	**0.93**	0.92	**0.92**	**0.92**	**0.91**	0.90	**0.90**	**0.89**	**0.88**
		RMSD	33.90	**34.16**	34.65	35.69	37.30	37.75	38.83	40.17	41.89	43.35	45.00
LSTM VEC	97.71	MAD	14.29	14.61	14.65	15.16	15.61	**15.93**	**16.61**	17.45	**17.83**	18.78	19.36
		R^2	**0.93**	**0.93**	**0.93**	**0.93**	**0.92**	**0.92**	**0.91**	0.90	**0.90**	**0.89**	**0.88**
		RMSD	**33.60**	34.20	34.26	**35.18**	**36.73**	**37.36**	38.89	40.09	41.65	43.37	45.50
LSTM E-D	149.50	MAD	15.18	15.92	16.16	16.45	16.97	17.30	17.58	18.21	19.02	19.38	20.51
		R^2	**0.93**	**0.93**	**0.93**	0.92	**0.92**	**0.92**	**0.91**	**0.91**	**0.90**	**0.89**	**0.88**
		RMSD	34.73	34.83	34.97	35.70	37.09	37.47	**38.42**	**39.26**	**40.97**	**42.45**	**44.53**
1D-CNN	23.54	MAD	14.82	14.45	14.69	15.61	17.45	16.56	17.18	18.65	18.78	19.54	19.93
		R^2	**0.93**	**0.93**	**0.93**	0.92	0.91	0.91	0.90	0.90	0.89	0.88	0.87
		RMSD	33.93	34.35	34.97	36.30	39.41	39.19	40.57	41.91	43.69	44.97	47.12
TNN	83.34	MAD	14.36	**14.35**	14.63	15.02	16.31	16.77	17.13	**17.36**	18.60	18.74	21.60
		R^2	**0.93**	**0.93**	**0.93**	**0.93**	**0.92**	0.91	**0.91**	0.90	0.89	**0.89**	0.87
		RMSD	33.73	34.20	**33.82**	35.40	37.57	37.99	39.47	40.48	42.70	43.86	46.08

45.00% in terms of RMSD, and between 0.93 and 0.88 in terms of R^2, as shown in Table 11. Moreover, the forecasting errors of the LSTM VEC model ranged from 14.29% to 19.36% in terms of MAD, from 33.60% to 45.50% in terms of RMSD and from 0.93 to 0.88 in terms of R^2. However, the significantly shorter training time of the LSTM VEC model (97.71 s) compared to the SVR model (669.06 s) makes the former a better alternative. The LSTM E-D model showed comparable performance to the LSTM VEC in terms of RMSD and R^2, even slightly surpassing it in some sporadic cases. Nonetheless, its results in terms of MAD were consistently much worse than those of the LSTM VEC model throughout the entire prediction horizon. On the contrary, the RF was, once again, the model presenting the worst performance, particularly in terms of RMSD and R^2. Furthermore, the 1D-CNN model performed better than the RF model but presented lower prediction accuracy than the other evaluated models. Finally, the TNN showed comparable performance to the LSTM VEC for a prediction horizon of up to 60 min, making it a valid competitor for forecasts with this horizon.

Fig. 7 visually compares the models' performance for the *D-5min* dataset. In line with the observations for the *D-1min* and *D-3min* datasets, the models' accuracy in terms of the three metrics worsened as the forecasting horizon increased. Based on the 20% MAD threshold, the methodology delivered acceptable results until a prediction horizon of 195 min (3 h and 15 min), 30 min longer than the horizon for the *D-3min* dataset. These results follow the trend of larger prediction horizons for increasing downsamplings. The numerical results of the forecasting errors depicted in Fig. 7 are reported in Table 12. The superiority of the LSTM E-D model compared to the other models is evident until a forecasting horizon of 150 min in terms of MAD and RMSD from the plots in Fig. 7. Table 12 confirms the superior performance of the LSTM E-D model over the other models within the temporal window of interest (i.e. 15 min to 195 min). The performance of the LSTM E-D model ranged from 12.38% to 19.62% in terms of MAD, from 28.80% to 45.99% in terms of RMSD and from 0.95 to 0.87 in terms of R^2. The SVR model also showed outstanding performance,

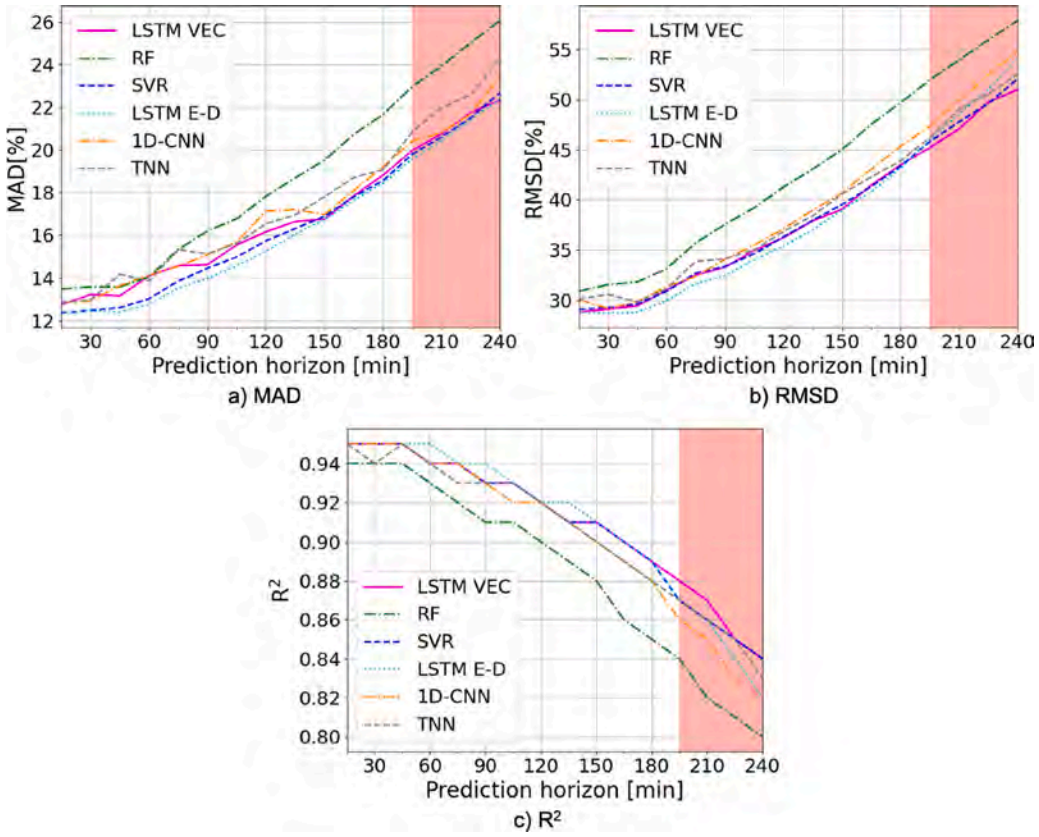


Fig. 7. Comparison of models point forecasting performance for the D-5 min dataset.

even outperforming the LSTM E-D model in terms of MAD for the 15 min prediction horizon. Nonetheless, it remained steadily behind the LSTM E-D for the other horizons. Furthermore, the LSTM VEC model presented comparable performance to the LSTM E-D for prediction horizons starting from 150 min, outperforming it in some sporadic cases. For example, the LSTM VEC model slightly surpassed the LSTM E-D model in terms of RMSD and R^2 for the 195 min prediction horizon, as shown in Table 12. However, this marginal performance improvement comes at a higher training time, with the LSTM VEC model requiring 231.45 s versus 72.61 s for the LSTM E-D, making the latter a more suitable choice even for these prediction horizons. On the other hand, the RF model demonstrated the lowest performance, which is more distinguishable for prediction horizons greater than 75 min. The 1D-CNN and the TNN models outperformed the RF model but presented lower prediction performance than the other evaluated models.

The various models' prediction performance in terms of MAD, RMSD and R^2 for the *D-15min* dataset are visually compared in Fig. 8. Consistent with what occurred for the models of the previous datasets, the performance of the models deteriorates for growing prediction horizons. Following the 20% MAD limit criteria, the models achieved satisfactory results up to a prediction horizon of 225 min (3 h and 45 min). As expected, this horizon is 30 min longer than for the *D-5min* dataset, consistent with the trend of longer prediction horizons with bigger downsampling. For short-term horizons (until 120 min), the LSTM VEC, SVR, LSTM E-D and 1D-CNN prediction errors were similar. Beyond 120 min, the superior performance of the LSTM VEC model became more evident, especially in terms of RMSD. Table 13 reports the numerical results of the prediction errors shown in Fig. 8. The SVR model outperformed the other models for forecasting horizons under two hours (i.e. 120 min), with prediction errors varying between 8.03% and 12.12% in terms of MAD, from 0.98 to 0.95 in terms of R^2, and from 18.17% to 28.09% in terms of RMSD. For forecasting horizons greater or equal to two hours, the LSTM VEC model demonstrated the best performance, with errors ranging from 12.31% to 18.73% in terms of MAD, from 0.95 to 0.89 in terms of R^2 and from 27.85% to 41.96% in terms of RMSD. However, the SVR model still showed promising results for this horizon (i.e. greater than two hours), occasionally outperforming the LSTM VEC model in some sporadic cases. For example, the SVR model surpassed the LSTM VEC model in terms of MAD at forecasting horizons of 120 min, 135 min, and 180 min, as reported in Table 13. Additionally, the SVR model had a slight advantage in training time (12.43 s) compared to the LSTM VEC model (81.27 s), making the SVR model a viable option for the entire prediction horizon of interest (i.e. 15 min to 225 min). Furthermore, the LSTM E-D model demonstrated comparable performance to the SVR model for the whole prediction horizon, even outperforming it for the largest prediction horizons (210 min and 225 min). However, for prediction horizons up to 165 min, the LSTM E-D model remained consistently behind the SVR. Therefore, the LSTM E-D model can be considered a valid competitor for prediction horizons greater than 165 min. The 1D-CNN model also showed very good prediction performance, even surpassing the SVR model for a prediction horizon of 75 min in terms of RMSD and 105 min in terms of the three metrics. Nonetheless, the SVR model outperformed the 1D-CNN model in all the other forecasting horizons. On the contrary, the RF model demonstrated the poorest performance among the models. Moreover, the TNN model performed slightly better than the RF model but still presented unsatisfactory results, especially in terms of RMSD.

For all the cases (i.e. *D-1min*, *D-3min*, *D-5min* and *D-15min*), the models showed a degradation of the prediction accuracy in terms of the three metrics for growing forecasting horizons. Nonetheless, the results highlighted that the models' performance improved for increasing

Table 12
Models point forecasting performance based on MAD [%]. RMSD [%] and R^2 for the D-5min dataset.

Model	Training Time [s]	Metric	Prediction Horizon [min]												
			15	30	45	60	75	90	105	120	135	150	165	180	195
RF	1.49	MAD	13.48	13.58	13.58	14.05	15.34	16.21	16.78	17.84	18.70	19.54	20.75	21.68	22.99
		R^2	0.94	0.94	0.94	0.93	0.92	0.91	0.91	0.90	0.89	0.88	0.86	0.85	0.84
		RMSD	30.87	31.54	31.81	33.12	35.70	37.57	39.23	41.29	43.10	45.03	47.56	49.69	51.96
SVR	206.76	MAD	**12.37**	12.48	12.61	13.03	13.85	14.46	14.98	15.74	16.32	16.90	17.82	18.58	19.82
		R^2	**0.95**	**0.95**	**0.95**	0.94	**0.94**	0.93	**0.93**	**0.92**	0.91	**0.91**	**0.90**	**0.89**	0.87
		RMSD	29.08	29.26	29.62	30.87	32.69	33.36	34.63	36.32	37.98	39.54	41.33	43.38	45.83
LSTM VEC	231.45	MAD	12.75	13.22	13.16	14.10	14.58	14.62	15.55	16.17	16.65	**16.76**	17.86	18.83	19.99
		R^2	**0.95**	**0.95**	**0.95**	0.94	**0.94**	0.93	**0.93**	**0.92**	0.91	**0.91**	**0.90**	**0.89**	**0.88**
		RMSD	28.82	29.09	29.43	30.99	32.49	33.28	34.96	36.23	37.92	39.07	41.48	43.50	**45.22**
LSTM E-D	72.61	MAD	12.38	**12.46**	**12.41**	**12.77**	**13.52**	**13.98**	**14.57**	**15.25**	**16.06**	**16.76**	**17.67**	**18.45**	**19.62**
		R^2	**0.95**	**0.95**	**0.95**	**0.95**	**0.94**	**0.94**	**0.93**	**0.92**	**0.92**	**0.91**	**0.90**	**0.89**	0.87
		RMSD	**28.80**	**28.74**	**28.82**	**29.94**	**31.63**	**32.46**	**34.03**	**35.40**	**37.03**	**38.99**	**40.81**	**43.19**	45.99
1D-CNN	31.82	MAD	12.86	12.93	13.63	14.13	14.55	15.09	15.69	17.14	17.21	16.98	18.12	19.22	20.44
		R^2	**0.95**	**0.95**	**0.95**	0.94	**0.94**	0.93	0.92	**0.92**	0.91	0.90	0.89	0.88	0.86
		RMSD	30.07	29.16	29.85	31.33	32.37	34.07	35.49	37.14	38.97	40.69	43.09	45.34	47.37
TNN	213.15	MAD	12.87	12.97	14.17	13.86	15.33	15.13	15.62	16.55	16.96	17.82	18.72	19.08	20.88
		R^2	**0.95**	0.94	**0.95**	0.94	0.93	0.93	**0.93**	**0.92**	0.91	0.90	0.89	0.88	0.87
		RMSD	30.11	30.56	29.84	31.13	33.86	34.11	35.00	36.87	38.42	40.61	42.31	43.89	46.22

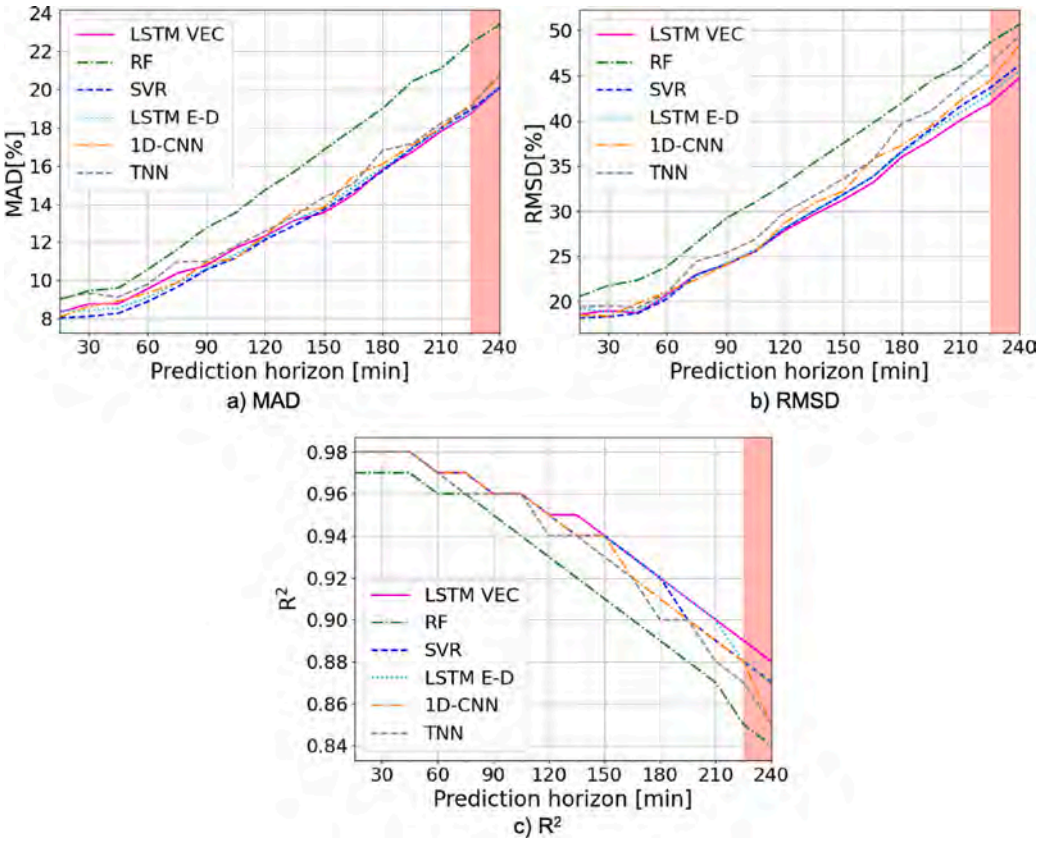


Fig. 8. Comparison of models point forecasting performance for the D-15 min dataset.

downsamplings. As a consequence, the prediction horizon enlarged as the downsampling increased: 90 min for the *D-1min* case; 165 min for *D-3min* case; 195 min for the *D-5min* case; and 225 min for the *D-15min* case.

4.2.2. Uncertainty analysis

We calculated the forecasting errors for the selected models of the various datasets. Afterwards, we computed the PIs using the NPKDE technique as detailed in Section 3.4. We trained a kernel model for every prediction horizon for each point forecasting model. We employed 50% of the *test set* (*PI training set*) to train the NPKDE models (to compute the Prediction Intervals) and the other 50% (*PI test set*) for testing purposes. To further assess the behaviour of the calculated PIs, we compare the performance of the Prediction Intervals on the *PI Evaluation set*. We implemented three confidence levels: 90%, 95% and 97.5%.

Table 13
Models point forecasting performance based on MAD [%]. RMSD [%] and R^2 for the D-15min dataset.

Model	Training Time [s]	Metric	Prediction Horizon [min]														
			15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	0.25	MAD	8.99	9.45	9.59	10.57	11.60	12.75	13.54	14.73	15.70	16.83	17.94	19.00	20.41	21.09	22.42
		R^2	0.97	0.97	0.97	0.96	0.96	0.95	0.94	0.93	0.92	0.91	0.90	0.89	0.88	0.87	0.85
		RMSD	20.52	21.76	22.34	23.85	26.49	29.18	31.00	33.02	35.37	37.53	39.74	41.98	44.48	46.06	48.70
SVR	12.43	MAD	**8.03**	**8.11**	**8.28**	**8.89**	**9.64**	**10.55**	11.17	**12.12**	**12.87**	13.69	14.67	**15.72**	16.92	18.01	18.92
		R^2	**0.98**	**0.98**	**0.98**	**0.97**	**0.97**	**0.96**	**0.96**	**0.95**	0.94	**0.94**	**0.93**	**0.92**	0.90	0.89	0.88
		RMSD	**18.17**	**18.31**	**18.69**	**20.28**	22.97	24.09	25.52	28.09	29.97	31.87	33.79	36.64	39.15	41.64	43.66
LSTM VEC	81.27	MAD	8.33	8.76	8.77	9.55	10.37	10.77	11.69	12.31	13.14	**13.54**	**14.49**	15.82	**16.72**	**17.82**	**18.73**
		R^2	**0.98**	**0.98**	**0.98**	**0.97**	**0.97**	**0.96**	**0.96**	**0.95**	**0.95**	**0.94**	**0.93**	**0.92**	**0.91**	**0.90**	**0.89**
		RMSD	18.58	19.00	18.73	20.68	23.00	**24.01**	25.58	**27.85**	**29.62**	**31.26**	**33.14**	**36.01**	**37.93**	**40.12**	**41.96**
LSTM E-D	21.77	MAD	8.41	8.42	8.55	9.10	9.84	10.61	11.35	12.22	13.13	13.84	14.88	15.89	16.94	17.89	18.74
		R^2	**0.98**	**0.98**	**0.98**	**0.97**	**0.97**	**0.96**	**0.96**	**0.95**	0.94	**0.94**	**0.93**	**0.92**	**0.91**	**0.90**	0.88
		RMSD	19.26	18.83	19.02	20.53	22.83	24.20	25.77	28.14	30.03	31.79	33.84	36.51	38.83	40.92	43.02
1D-CNN	14.80	MAD	8.07	8.63	8.91	9.33	9.87	10.94	**11.12**	12.36	13.62	13.79	15.46	16.10	17.07	17.99	19.20
		R^2	**0.98**	**0.98**	**0.98**	**0.97**	**0.97**	**0.96**	**0.96**	**0.95**	0.94	**0.94**	0.92	0.91	0.90	0.89	0.88
		RMSD	18.50	18.38	19.79	21.09	**22.43**	24.05	**25.51**	28.67	30.90	32.21	35.79	37.30	39.47	42.28	44.42
TNN	33.96	MAD	9.05	9.32	9.12	9.79	10.99	10.99	11.79	12.56	13.36	14.35	15.05	16.81	17.15	18.23	19.07
		R^2	**0.98**	**0.98**	**0.98**	**0.97**	0.96	**0.96**	**0.96**	0.94	0.94	0.93	0.92	0.90	0.90	0.88	0.87
		RMSD	19.40	19.53	19.30	20.96	24.44	25.35	26.88	29.87	31.68	33.60	35.61	39.76	41.10	43.86	46.36

Fig. 9 visually compares the models' prediction interval performance in terms of PINAW and CRPS for the *D-1min* dataset. The continuous lines in Fig. 9 represent the prediction errors of the various models evaluated on the *PI test set*. On the other hand, the dashed lines correspond to the results of the models for the *PI Evaluation set* (*PI-EVAL*). As illustrated in Fig. 9, the PI performance of the models in terms of PINAW for the three studied confidence levels (97.5%, 95% and 90%) was superior when evaluated on the *PI Evaluation set* (dashed lines in Fig. 9) compared to the *PI test set* (continuous lines in Fig. 9). Additionally, Fig. 9 depicts that the errors remained nearly constant as the prediction horizon increased, i.e. there was almost no degradation of the PI forecasting performance as the horizon enlarged. Furthermore, better results can be observed in terms of PINAW and CRPS as the confidence level decreased: the models' errors were lower for the 90% confidence level than those for the 95% and 97.5% cases. The RF model demonstrated the lowest PI performance in terms of CRPS and PINAW when assessed on both the *PI test set* and the *PI Evaluation set* for confidence levels of 97.5% and 95%. For the 90% confidence level, the worse performance of the RF model became more evident in terms of both PINAW and CRPS after a prediction horizon of 60 min. The other five models showed similar PI prediction performance in terms of both metrics (PINAW and CRPS).

The models' prediction interval performance in terms of PINAW and CRPS for the *D-3min* case are visually compared in Fig. 10. Consistent with what occurred for the *D-1min* case, the PI performance in terms of PINAW was superior when evaluated on the *PI Evaluation set* compared to the *PI test set*, as shown in Fig. 10. Additionally, the PI performance in terms of CRPS was better when evaluated on the *PI test set* than on the *PI Evaluation set*. Furthermore, the models' PI performance in terms of PINAW and CRPS improved as the confidence level decreased. Nonetheless, unlike the *D-1min* dataset, the performance of the models decreased as the prediction horizon enlarged, with the errors in terms of PINAW and CRPS growing as the horizon increased. The RF model demonstrated the worst PI performance for the 97.5% confidence level, particularly in terms of PINAW. This model also exhibited the poorest PI performance for the 95% and 90% confidence levels for prediction horizons larger than 60 min. On the other hand, the SVR model showed the best PI performance for prediction horizons beyond 75 min for the three studied confidence levels (97.5%, 95% and 90%).

Fig. 11 provides a visual comparison of the PI performance of the models in terms of PINAW and CRPS for the *D-5min* dataset. In line with the observations done for the *D-1min* and *D-3min* case, the models' PI performance in terms of CRPS was superior when assessed on the *PI test set* (represented with continuous lines in Fig. 11) compared to the *PI Evaluation set* (represented with dashed lines in Fig. 11). In contrast, in terms of PINAW, the models showed the opposite behaviour (better results when evaluated on the *PI Evaluation set* than on the *PI test set*). Furthermore, the models' PI performance in terms of PINAW and CRPS improved as the confidence level decreased. Additionally, similar to the *D-3min* dataset, there was a noticeable degradation in the models' performance as the prediction horizon increased. The RF model exhibited the lowest performance in terms of PINAW and CRPS for the three analysed confidence levels, both when evaluated on the *PI test set* and the *PI Evaluation set*. The RF worst performance was more evident in terms of PINAW. On the contrary, the remaining models showed comparable results for these confidence levels.

The models' PI performance in terms of PINAW and CRPS for the *D-15min* case are visually compared in Fig. 12. Consistent with the results obtained for the *D-1min*, *D-3min*, and *D-5min* cases, the models' PI performance in terms of PINAW was superior when assessed on the *PI Evaluation set* than on the *PI test set*. Instead, the models' CRPS results were better when evaluated on the *PI test set*. Furthermore, the models' performance worsened as the prediction horizon increased. Among the evaluated models, the RF model demonstrated, once again, the worst performance in terms of both PINAW and CRPS for all three confidence levels. The LSTM VEC model exhibited overall good performance, especially for forecasting horizons larger than 180 min, where it outperformed the other models for the three confidence levels. Nonetheless, the SVR model showed comparable performance to the LSTM VEC, even surpassing it for some sporadic cases. For example, the SVR model outperformed the LSTM VEC in terms of CRPS for prediction horizons shorter than 165 min, as illustrated in Fig. 12, making the SVR model a valid competitor for the LSTM VEC.

Tables A.14 to A.37 in Appendix report more details about the results on the uncertainty analysis. The Tables demonstrate that all the models of the different cases satisfied the desired confidence levels (PICP greater or equal to the confidence level), with the only exception of the RF model of the *D-1min* case evaluated on the *PI test set*. The results clearly showed that the models' performance improved as the confidence level decreased. Finally, there can be observed an improvement of the models' performance as the downsampling increased, which is easily distinguishable by comparing the CRPS results of the *D-15min* with the other cases. For instance, for the 90% confidence level of the *D-15min* case, the models' CRPS errors on the *PI Evaluation set* were below 15% for prediction horizons shorter than 60 min as shown in Fig. 12(c). In contrast, for the *D-5min* case, the CRPS errors for all models were already about 15% at a prediction horizon of 15 min as depicted in Fig. 11(c).

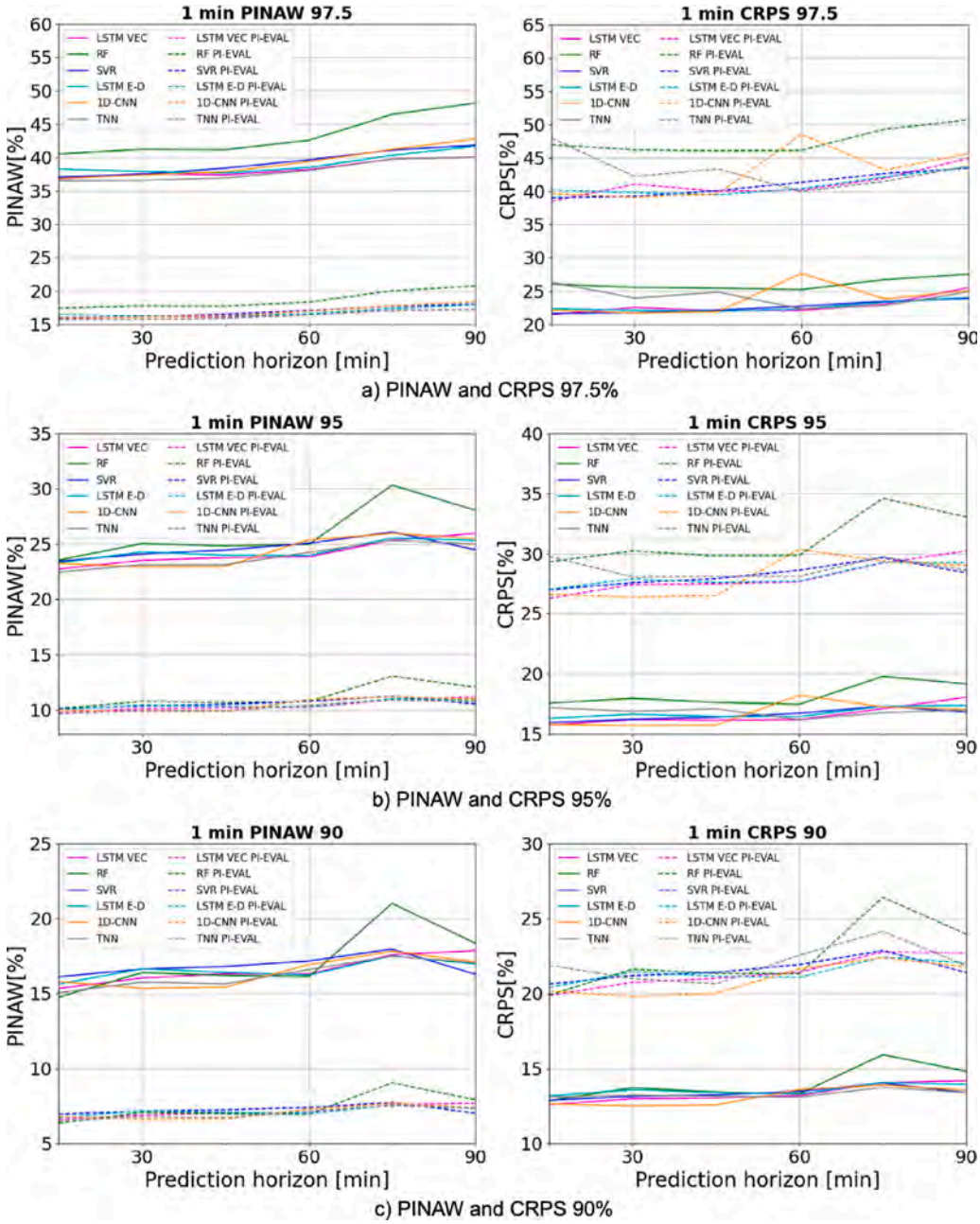


Fig. 9. Comparison of models Prediction Interval forecasting performance in terms of PINAW (left) and CRPS (right) for the D-1 min PI Test Set (continuous lines) and PI Evaluation Set (dashed lines).

5. Conclusions

In this work, we introduced a novel methodology to forecast in short-term horizons (i.e. 15 min to 240 min) the power delivered to the grid of the ISWEC device. Our main goal was to obtain accurate forecasts suitable for smart grid advanced monitoring and control techniques, such as State Estimation and Demand-Response. To this aim, we computed and compared the prediction performance of five Machine Learning models used for time-series point forecasting which had shown promising results for wave power and time series forecasting: Random Forest (RF), Support Vector Regression (SVR), 1 Dimensional Convolutional Neural Network (1D-CNN), Transformer Neural Network (TNN) and a Long Short-Term Memory (LSTM) neural network in two different forms (Vector and Encoder-Decoder). Compared with the traditional physical approaches, performing the predictions employing these machine learning algorithms provides significant advantages in terms of time and cost (Mousavi et al., 2021). Additionally, we studied the effectiveness of the downsampling technique by aggregating the original dataset sampled every 0.1 s in time windows of 1 min, 3 min, 5 min and 15 min, referring to these datasets as *D-1min*, *D-3min*, *D-5min* and *D-15min*. We compared the prediction performance of the different models for these different downsampled versions of the dataset (i.e. *D-1min*, *D-3min*, *D-5min* and *D-15min*). Furthermore, to calculate the inherent uncertainties associated with the point forecasting models, we computed the Prediction Intervals using Non-Parameteric Kernel Density Estimation (NPKDE).

The point forecasting results showed a considerable improvement in the models' performance as the downsample increased, which was noticeable with the enlargement of the prediction horizon for bigger downsamplings: 90 min for the *D-1min* case; 165 min for *D-3min* case; 195 min for the *D-5min* case; and 225 min for the *D-15min* case. The RF model was the worst-performing model in all the cases.

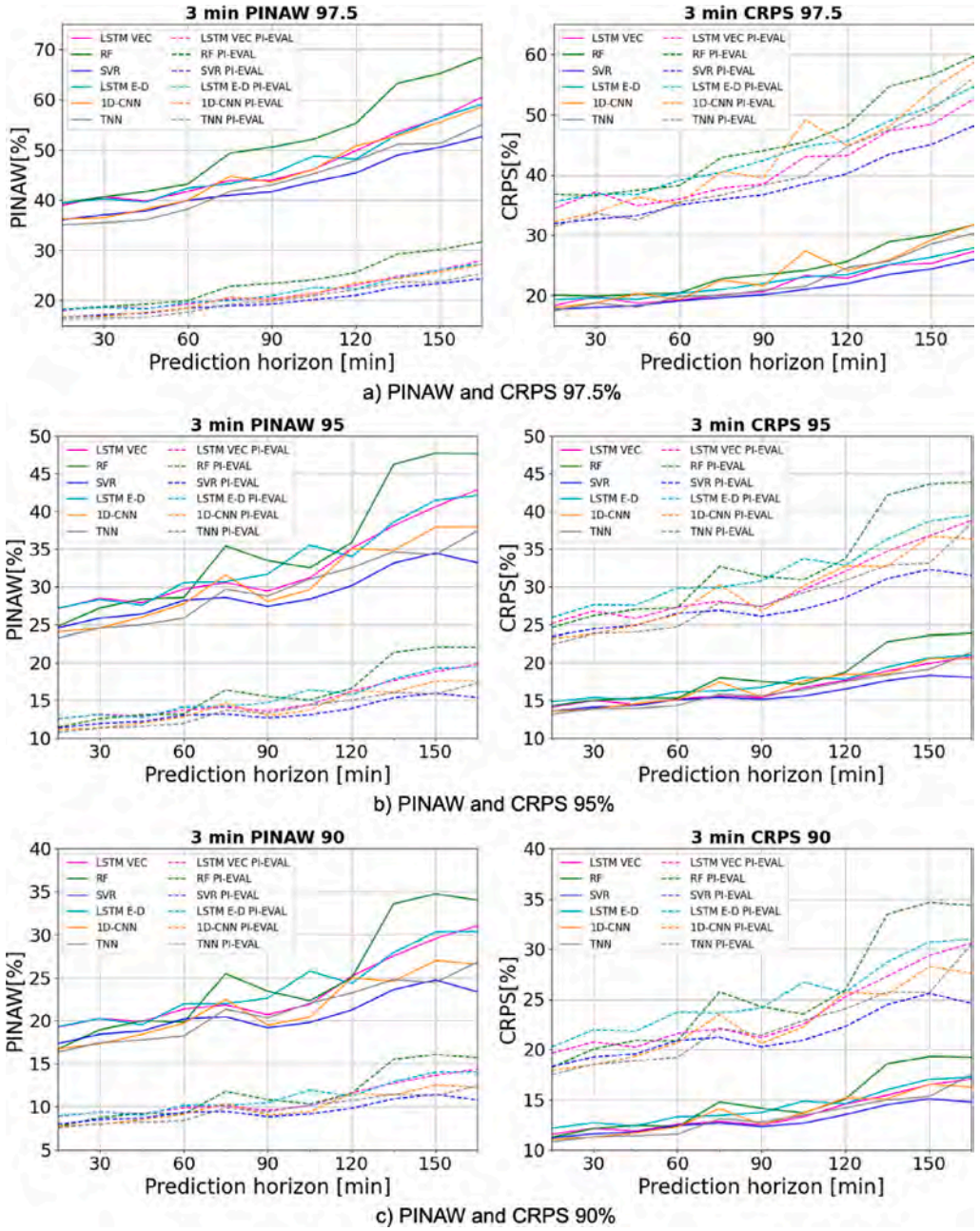


Fig. 10. Comparison of models Prediction Interval forecasting performance in terms of PINAW (left) and CRPS (right) for the D-3 min PI Test Set (continuous lines) and PI Evaluation Set (dashed lines).

On the contrary, no unique ML model performed better for all the downsamplings. For the *D-1min* case, the 1D-CNN model demonstrated the best prediction performance. Furthermore, the LSTM VEC and SVR models showed promising behaviour for the *D-3min* case. Nonetheless, the shorter training time of the LSTM VEC compared to the SVR made the first a better alternative. Moreover, the TNN presented comparable performance to the LSTM VEC for prediction horizons up to 60 min, making it a valid competitor for forecasts with this horizon. For the *D-5min* case, the LSTM E-D model demonstrated the best performance. In the *D-15min* case, the SVR model presented an overall good performance for the entire prediction horizon. Moreover, the LSTM E-D model showed comparable performance to the SVR for horizons bigger than 165 min, making it a feasible alternative for these larger horizons.

In line with what occurred in point forecasting, the uncertainty analysis results showed that the performance of the models' PIs improved for bigger downsamplings. Additionally, as the confidence level decreased, the performance of the models increased. Once again, the RF presented the worst performance among the models.

As discussed above, the point forecasting and uncertainty analysis results showed that the models' performance improved as the downsampling increased. This behaviour can be explained by the reductions in the data fluctuations and the peaks smoothing the downsampling technique performs, making it easier for the models to follow the data trends, especially in medium and long-term predictions. Another reason for this improvement can be that the models are simplified when aggregating data in bigger time steps since the number of data points is reduced. This also implies that for forecasting a particular time ahead (e.g. 4 h), the number of steps ahead that the model needs to predict can be significantly reduced. For example, for a 4-h ahead prediction (i.e. 240 min), with a downsampling of 1 min, the algorithm performs predictions of 240 steps ahead, while with a downsampling of 15 min, the steps ahead are "only" 16. Nonetheless, it is crucial to understand

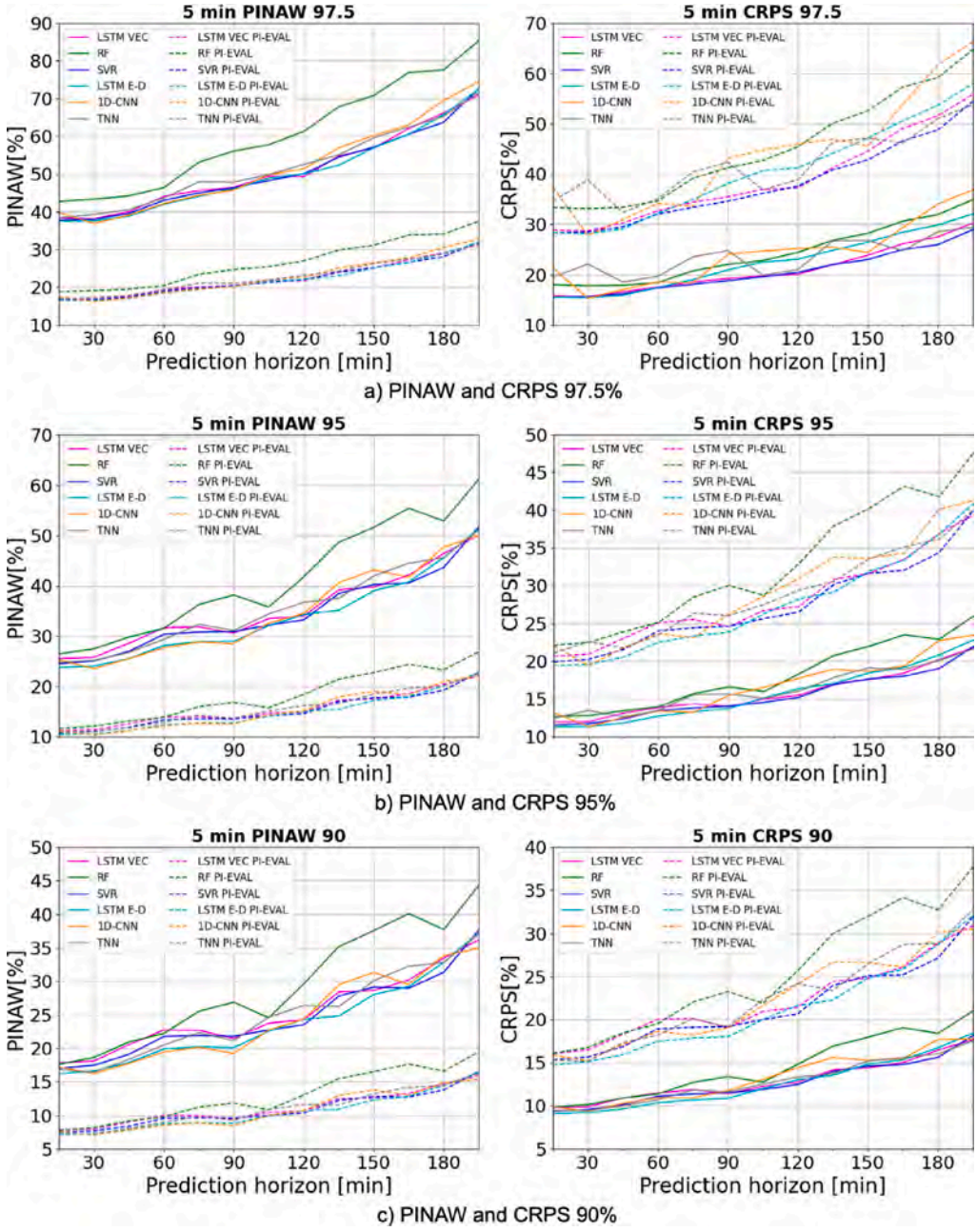


Fig. 11. Comparison of models Prediction Interval forecasting performance in terms of PINAW (left) and CRPS (right) for the D-5 min PI Test Set (continuous lines) and PI Evaluation Set (dashed lines).

the predictions' final objective before employing the downsampling technique in a forecasting scenario. This is because downsampling is strongly related to the use case for which the forecasting is intended. For instance, if the goal is to employ the forecasts for State Estimation, where nowcasting is typically used, the downsampling is restricted to prevent smoothing out important information.

Furthermore, it is important to highlight that in many cases (e.g. the *D-3min* LSTM VEC model), the best-performing point-forecasting model did not show good uncertainty analysis results. This means that it is crucial to understand the final objective of the forecasting to make proper decisions regarding the downsampling to implement (as previously discussed) and, for the cases where the associated uncertainties need to be taken into account, compute the uncertainties analysis of the models and compare the models' performance based on these results. For example, the *D-1min* case can be associated with WEC control, where very short-term forecasts are usually employed, and downsampling is restricted to prevent smoothing out valuable data. Instead, the *D-15min* case can be associated with Demand-Response techniques for efficient grid dispatching, where short-term forecasts (i.e. 15 min to 4 h) are usually employed. In this scenario, the computation of the uncertainty analysis may also play a key role in providing important information to the grid operators.

In the proposed methodology, we employed the *ISWEC Wave-to-Grid Simulator* to obtain a realistic dataset of the power delivered to the grid due to the lack of real-world data. This simulator had been extensively validated against experimental tests, with an average percentage error

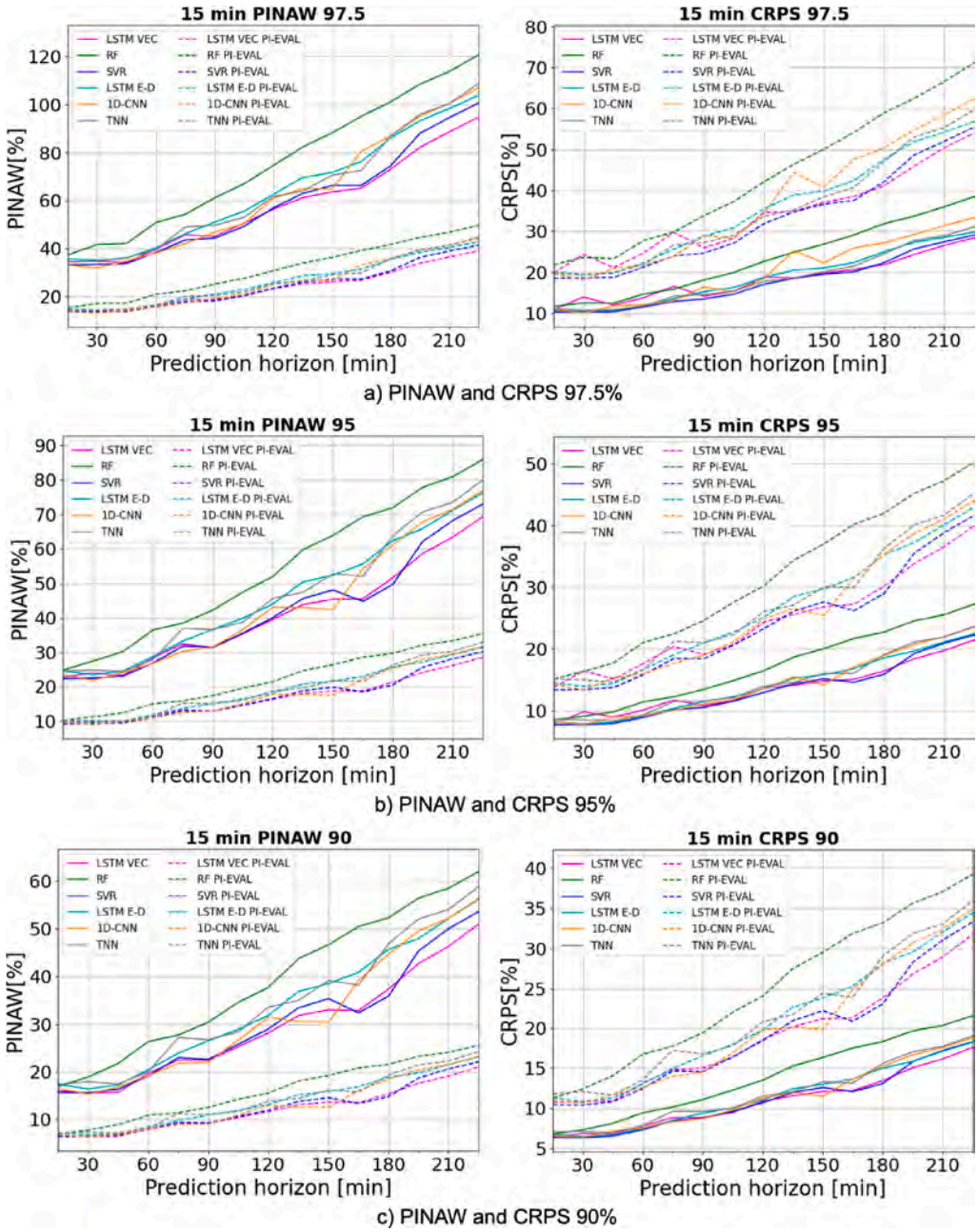


Fig. 12. Comparison of models Prediction Interval forecasting performance in terms of PINAW (left) and CRPS (right) for the D-15 min PI Test Set (continuous lines) and PI Evaluation Set (dashed lines).

of 11.2% of the delivered power to the grid. Therefore, the prediction accuracies of the models are inherently limited by these errors. Nonetheless, using the simulator mitigated the significant challenge of data scarcity for the case under study. Moreover, the usage of this simulator unlocks the possibility to generalize the methodology to other sea regions, "relaunching" the *ISWEC Wave-to-Grid Simulator* using the new region of interest as input to obtain the "new" dataset and following the proposed methodology. On the contrary, using the *ISWEC Wave-to-Grid Simulator* for other wave energy converters is not possible because the simulator is WEC-dependent. Nonetheless, in the case of employing different classes of WECs, our methodology can be updated, replacing only the *ISWEC Wave-to-Grid Simulator* with the corresponding simulator for the desired WEC class (e.g. point absorbers or attenuators). In contrast, the other modules of the methodology (i.e. *dataset*, *point forecasting* and *uncertainty analysis*) remain unchanged. When generalizing the methodology to other sea regions or employing the forecasting modules for other WECs, the prediction performance of the algorithms may worsen due to potential new "unseen" trends or patterns arising from different wave patterns and weather conditions. Transfer learning can be employed to address these challenges. In this context, the forecasting algorithms can be pre-trained with the dataset generated by the *ISWEC Wave-to-Grid Simulator*. Afterwards, the algorithms can be fine-tuned with the new sea region or WEC data. Transfer Learning also becomes crucial once real-world data of WEC's power delivered to the grid becomes available. In such a scenario, the algorithms can be fine-tuned by employing real-world datasets, demonstrating that the

Table A.14
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI test set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	99.19	99.26	99.18	99.28	99.53	99.57
	PINAW	40.54	41.26	41.19	42.59	46.47	48.16
	CRPS	26.05	25.55	25.43	25.20	26.72	27.52
SVR	PICP	99.33	99.36	99.44	99.51	99.53	99.55
	PINAW	37.09	37.49	38.41	39.67	41.13	41.87
	CRPS	21.64	21.75	22.17	22.73	23.47	**23.90**
LSTM VEC	PICP	99.10	99.04	99.17	99.36	99.38	99.51
	PINAW	36.77	37.59	37.40	38.20	39.79	40.08
	CRPS	**21.48**	22.55	22.06	**22.10**	23.04	25.49
LSTM E-D	PICP	99.18	99.28	99.19	99.28	99.39	99.46
	PINAW	38.28	37.91	37.67	38.51	40.35	41.70
	CRPS	22.36	22.13	21.98	22.32	23.31	24.07
1D-CNN	PICP	99.33	99.30	99.33	99.50	99.55	99.67
	PINAW	36.88	37.39	37.86	39.39	41.23	42.80
	CRPS	22.14	**21.69**	**21.92**	27.63	23.83	25.06
TNN	PICP	99.33	99.36	99.40	99.47	99.46	99.55
	PINAW	**36.51**	**36.59**	**36.95**	**38.07**	**39.77**	**40.02**
	CRPS	26.30	23.95	24.81	22.34	**22.83**	24.86

proposed methodology can be applied to provide realistic forecasts. The fact of using the simulated data for training enables to fine-tune the algorithms and perform the forecast without the need for a large real-world dataset, which should report even years of collected data. This is a crucial advantage in a scenario where real-world data is scarce.

In this study, we compute the associated uncertainties with the NPKDE method. Future works may include studying and comparing other techniques (e.g. Quantile Regression). Moreover, we aim to study the performance of other downsampling methods like exponential weighting. Furthermore, in this work, we focused solely on the power delivered to the grid, employing it as the only input variable for the models to perform the predictions and comparing the prediction of the models using different downsample versions of the original dataset (i.e. 1 min, 3 min, 5 min, and 15 min). As previously discussed, the results improved as the downsampling increased. Nonetheless, the models' performance may be improved by utilizing other exogenous inputs in the forecasting algorithms without modifying the downsampling. For this reason, we plan to investigate in future work the effectiveness of employing other input variables to forecast the power delivered to the grid, such as, for example, motion measurements of the ISWEC (e.g. accelerations and angular velocities), as well as other exogenous inputs like weather forecasts. The study and analysis of the effectiveness of employing these other inputs to perform forecasting will be accompanied by the corresponding sensitivity analysis of the impact of these variables on the outcomes. We aim to identify the most relevant inputs correlated with the power delivered to the grid phenomena by applying and comparing different feature selection techniques. Finally, we plan to compare the performance of multivariate models with the univariate implementations presented in this work.

CRediT authorship contribution statement

Rafael Natalio Fontana Crespo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Alessandro Aliberti:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Lorenzo Bottaccioli:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Edoardo Pasta:** Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Sergej Antonello Sirigu:** Writing – review & editing, Supervision, Project administration. **Enrico Macii:** Supervision, Project administration. **Giuliana Mattiazzo:** Supervision, Project administration. **Edoardo Patti:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This publication is part of the following projects:

- PNRR-NGEU which has received funding from the MUR, Italy – DM 352/2022;
- "Network 4 Energy Sustainable Transition – NEST", Project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca (MUR), CUP E13C2 2001890001. Project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3 - Call for tender No. 341 of 15.03.2022 of MUR; funded by the European Union – NextGenerationEU.

Appendix. Model Prediction Interval performance in terms of PICP [%], PINAW [%] and CRPS [%]

See Tables A.14–A.37.

Table A.15
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI test set with confidence level 95%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	95.46	96.25	95.79	95.93	97.56	96.33
	PINAW	23.54	25.03	24.82	24.97	30.32	28.05
	CRPS	17.56	17.94	17.62	17.46	19.78	19.16
SVR	PICP	96.62	96.82	96.89	96.97	97.21	96.41
	PINAW	23.44	24.09	24.45	25.06	26.09	**24.47**
	CRPS	15.94	16.22	16.39	16.72	17.31	**16.81**
LSTM VEC	PICP	96.13	96.42	96.55	96.67	96.97	97.01
	PINAW	22.72	23.53	23.76	**23.83**	25.35	25.96
	CRPS	**15.74**	16.19	16.15	16.19	17.09	18.04
LSTM E-D	PICP	96.01	96.52	96.36	96.31	96.75	96.49
	PINAW	23.29	24.28	23.98	23.95	25.54	25.33
	CRPS	16.29	16.64	16.45	16.44	17.30	17.37
1D-CNN	PICP	96.78	96.49	96.45	97.40	97.30	97.08
	PINAW	23.24	**22.92**	**22.97**	25.37	25.96	25.50
	CRPS	15.80	**15.68**	**15.73**	18.22	17.25	17.03
TNN	PICP	96.61	96.72	96.64	96.84	96.98	96.92
	PINAW	**22.45**	23.12	23.13	24.23	**25.31**	24.99
	CRPS	17.18	16.87	17.08	**16.18**	**16.77**	16.98

Table A.16
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI test set with confidence level 90%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	*88.36*	90.03	90.04	90.25	93.84	90.92
	PINAW	**14.75**	16.40	16.22	**16.13**	21.01	18.36
	CRPS	12.83	13.72	13.42	13.31	15.91	14.80
SVR	PICP	92.07	92.56	92.56	92.60	92.98	90.87
	PINAW	16.12	16.64	16.81	17.18	17.97	**16.29**
	CRPS	12.87	13.13	13.23	13.47	14.02	**13.38**
LSTM VEC	PICP	91.07	91.27	91.95	92.09	92.63	92.12
	PINAW	15.34	16.04	16.32	16.31	17.60	17.86
	CRPS	**12.58**	12.97	13.02	13.19	14.05	14.18
LSTM E-D	PICP	90.09	91.33	91.26	91.08	91.96	90.78
	PINAW	15.63	16.66	16.42	16.24	17.53	16.98
	CRPS	13.16	13.58	13.39	13.28	14.06	13.93
1D-CNN	PICP	91.96	91.13	91.04	92.48	92.89	92.16
	PINAW	15.80	**15.35**	**15.43**	16.94	17.82	17.12
	CRPS	12.63	**12.50**	**12.57**	13.59	13.89	13.55
TNN	PICP	91.27	91.92	91.56	92.37	92.68	92.23
	PINAW	15.05	15.77	15.65	16.61	**17.50**	17.06
	CRPS	13.08	13.21	13.14	**13.08**	**13.74**	13.39

Table A.17
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI evaluation set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	99.04	99.14	99.10	99.16	99.25	99.25
	PINAW	17.44	17.75	17.72	18.33	20.00	20.72
	CRPS	47.00	46.24	46.11	46.17	49.32	50.74
SVR	PICP	99.14	99.17	99.21	99.26	99.29	99.28
	PINAW	15.96	16.13	16.53	17.07	17.70	18.02
	CRPS	39.01	39.24	40.10	41.33	42.66	**43.47**
LSTM VEC	PICP	99.07	99.05	99.13	99.17	99.11	99.12
	PINAW	15.82	16.17	16.09	16.44	17.12	17.24
	CRPS	**38.51**	41.05	39.99	40.18	41.97	44.95
LSTM E-D	PICP	99.13	99.14	99.12	99.14	99.19	99.16
	PINAW	16.47	16.31	16.21	16.57	17.36	17.94
	CRPS	40.16	39.82	**39.50**	40.37	42.16	43.63
1D-CNN	PICP	99.21	99.14	99.19	99.27	99.28	99.35
	PINAW	15.87	16.09	16.29	16.95	17.74	18.42
	CRPS	39.65	**39.08**	39.64	48.55	43.21	45.67
TNN	PICP	99.23	99.19	99.23	99.23	99.23	99.21
	PINAW	**15.71**	**15.74**	**15.90**	**16.38**	**17.11**	**17.22**
	CRPS	48.01	42.23	43.34	**39.96**	**41.49**	43.77

Table A.18
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI evaluation set with confidence level 95%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	97.17	97.56	97.51	97.43	98.11	97.61
	PINAW	10.13	10.77	10.68	10.75	13.05	12.07
	CRPS	29.32	30.21	29.81	29.85	34.60	33.05
SVR	PICP	97.81	97.98	98.02	98.01	98.10	97.69
	PINAW	10.09	10.37	10.52	10.78	11.23	**10.53**
	CRPS	26.97	27.56	27.91	28.61	29.66	**28.38**
LSTM VEC	PICP	97.60	97.70	97.83	97.72	97.83	97.95
	PINAW	9.78	10.12	10.22	**10.25**	10.91	11.17
	CRPS	**26.25**	27.42	27.48	**27.64**	**29.21**	30.21
LSTM E-D	PICP	97.46	97.73	97.76	97.63	97.81	97.60
	PINAW	10.02	10.45	10.32	10.31	10.99	10.90
	CRPS	27.01	27.89	27.57	27.70	29.26	29.25
1D-CNN	PICP	97.84	97.71	97.66	98.18	98.11	97.97
	PINAW	10.00	**9.86**	**9.88**	10.92	11.17	10.97
	CRPS	26.59	**26.36**	**26.51**	30.35	29.37	29.01
TNN	PICP	97.80	97.88	97.88	97.94	97.99	97.96
	PINAW	**9.66**	9.95	9.95	10.43	**10.89**	10.75
	CRPS	29.76	28.05	28.10	28.10	29.74	28.60

Table A.19
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-1min PI evaluation set with confidence level 90%.

Model	Metric	Prediction Horizon [min]					
		15	30	45	60	75	90
RF	PICP	94.01	94.88	94.87	94.61	96.29	95.09
	PINAW	**6.35**	7.06	6.98	**6.94**	9.04	7.90
	CRPS	19.91	21.63	21.35	21.33	26.41	23.97
SVR	PICP	95.59	95.86	95.89	95.90	96.14	95.28
	PINAW	6.94	7.16	7.23	7.39	7.73	**7.01**
	CRPS	20.64	21.20	21.42	21.94	22.87	**21.41**
LSTM VEC	PICP	95.22	95.46	95.76	95.62	95.89	95.85
	PINAW	6.60	6.90	7.02	7.02	7.57	7.68
	CRPS	**19.88**	20.75	21.01	21.44	22.78	22.72
LSTM E-D	PICP	94.86	95.43	95.36	95.10	95.57	95.05
	PINAW	6.72	7.17	7.06	6.99	7.54	7.31
	CRPS	20.42	21.45	21.15	**21.09**	22.44	22.04
1D-CNN	PICP	95.61	95.19	95.15	95.98	96.10	95.72
	PINAW	6.80	**6.61**	**6.64**	7.29	7.67	7.37
	CRPS	20.17	**19.81**	**20.02**	21.65	**22.40**	21.78
TNN	PICP	95.42	95.73	95.63	95.90	96.11	95.85
	PINAW	6.48	6.78	6.73	7.15	**7.53**	7.34
	CRPS	21.90	21.00	20.65	22.57	24.15	21.87

Table A.20
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI test set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	99.33	99.48	99.48	99.55	99.69	99.64	99.71	99.69	99.81	99.79	99.83
	PINAW	39.38	40.63	41.69	43.17	49.33	50.55	52.14	55.30	63.22	65.13	68.41
	CRPS	20.03	19.84	20.16	20.39	22.78	23.44	24.13	25.61	28.90	29.94	31.67
SVR	PICP	99.27	99.53	99.58	99.67	99.65	99.62	99.69	99.67	99.76	99.72	99.69
	PINAW	36.07	37.05	37.80	39.93	**40.94**	**41.65**	**43.67**	**45.37**	**48.96**	**50.51**	**52.58**
	CRPS	17.67	**18.00**	18.29	**19.11**	**19.64**	**20.09**	**20.98**	**21.94**	**23.49**	**24.36**	**25.92**
LSTM VEC	PICP	99.44	99.52	99.59	99.71	99.76	99.73	99.83	99.81	99.86	99.88	99.90
	PINAW	38.94	40.63	39.79	41.71	43.89	43.94	46.14	49.84	53.57	56.34	60.42
	CRPS	18.30	19.84	18.69	19.20	20.08	20.57	23.30	22.84	25.05	25.30	27.28
LSTM E-D	PICP	99.34	99.56	99.25	99.56	99.54	99.39	99.71	99.56	99.73	99.85	99.83
	PINAW	39.36	40.26	39.58	42.37	43.24	45.20	48.79	48.15	53.09	56.38	59.02
	CRPS	19.32	19.59	19.37	20.34	20.96	21.97	23.06	23.46	25.13	26.34	27.85
1D-CNN	PICP	99.37	99.44	99.48	99.60	99.65	99.55	99.74	99.74	99.79	99.81	99.91
	PINAW	36.23	36.46	38.18	39.91	44.74	43.59	46.13	50.74	52.81	55.45	58.51
	CRPS	17.94	18.77	20.29	19.29	22.54	21.60	27.37	24.09	25.89	29.20	31.76
TNN	PICP	99.30	99.46	99.39	99.65	99.62	99.67	99.77	99.84	99.88	99.86	99.86
	PINAW	**35.04**	**35.48**	**36.06**	**38.18**	41.69	42.99	45.26	47.88	51.10	51.26	54.98
	CRPS	**17.46**	18.73	**18.07**	19.81	19.96	20.84	21.47	24.59	25.66	28.53	30.28

Table A.21
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI test set with confidence level 95%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	96.88	97.64	97.88	97.95	98.95	98.14	97.81	98.07	99.05	99.05	98.62
	PINAW	24.73	27.18	28.36	28.57	35.41	33.49	32.54	35.87	46.21	47.72	47.65
	CRPS	14.27	14.98	15.29	15.27	17.97	17.50	17.13	18.70	22.77	23.59	23.82
SVR	PICP	97.67	98.00	98.07	98.38	98.35	97.69	97.93	97.86	98.31	98.26	97.95
	PINAW	24.55	25.83	26.37	28.19	**28.58**	**27.40**	**28.35**	**30.15**	**33.15**	34.48	**33.20**
	CRPS	13.61	14.08	14.31	15.05	**15.37**	**15.12**	**15.55**	**16.49**	**17.60**	**18.31**	**18.02**
LSTM VEC	PICP	98.21	98.33	98.26	98.74	98.74	98.36	98.60	98.94	99.23	99.32	99.32
	PINAW	27.07	28.47	27.89	29.69	30.56	29.43	31.20	35.09	38.10	40.63	42.86
	CRPS	14.08	14.94	14.45	15.16	15.58	15.36	16.68	17.65	18.88	19.89	20.77
LSTM E-D	PICP	97.96	98.23	97.72	98.37	98.28	98.03	98.91	98.64	99.10	99.05	99.17
	PINAW	27.16	28.32	27.54	30.56	30.68	31.69	35.54	34.02	38.57	41.47	42.08
	CRPS	14.81	15.34	15.11	16.09	16.24	16.72	18.02	17.78	19.41	20.61	20.97
1D-CNN	PICP	97.46	97.68	98.00	98.22	98.69	97.79	98.12	98.78	98.66	99.04	98.87
	PINAW	24.05	24.53	25.97	27.74	31.59	28.04	29.63	35.04	34.83	37.91	37.95
	CRPS	13.58	**13.86**	14.54	15.04	17.43	15.45	17.56	18.46	18.45	20.53	20.51
TNN	PICP	97.44	97.79	97.77	98.00	98.52	98.26	98.50	98.62	98.87	98.36	98.71
	PINAW	**23.26**	**24.51**	**24.96**	**25.83**	29.68	28.76	31.04	32.54	34.65	**34.27**	37.44
	CRPS	**13.17**	13.90	**13.92**	**14.30**	15.81	15.59	16.42	17.49	18.37	19.11	21.36

Table A.22
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI test set with confidence level 90%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	91.81	94.00	94.31	94.05	96.69	95.05	94.33	94.67	97.26	97.02	96.67
	PINAW	16.71	18.94	20.00	19.84	25.47	23.44	22.34	25.00	33.57	34.72	34.00
	CRPS	11.25	12.12	12.45	12.31	14.76	14.13	13.67	15.04	18.57	19.27	19.21
SVR	PICP	94.35	94.85	95.01	95.86	95.04	94.05	94.33	94.49	95.98	95.51	94.68
	PINAW	17.33	18.44	18.80	20.22	**20.47**	**19.18**	**19.81**	**21.25**	**23.62**	24.74	**23.36**
	CRPS	11.18	11.61	11.76	12.42	**12.68**	**12.32**	**12.67**	**13.50**	**14.49**	**15.09**	**14.75**
LSTM VEC	PICP	95.17	95.80	95.65	96.18	96.21	95.05	95.70	96.52	97.54	97.46	97.80
	PINAW	19.24	20.25	19.90	21.36	21.85	20.71	21.91	25.18	27.54	29.55	31.05
	CRPS	11.56	12.14	11.86	12.55	12.84	12.49	13.28	14.54	15.44	16.50	17.03
LSTM E-D	PICP	94.47	95.15	93.88	95.61	94.90	95.19	96.65	95.29	96.97	97.38	97.26
	PINAW	19.32	20.26	19.54	21.98	21.98	22.63	25.76	24.35	27.94	30.30	30.36
	CRPS	12.15	12.71	12.38	13.31	13.41	13.70	14.84	14.58	15.99	17.02	17.22
1D-CNN	PICP	93.71	94.15	94.74	95.19	95.94	94.27	94.55	96.31	96.20	96.43	96.03
	PINAW	16.75	**17.31**	18.42	19.72	22.47	19.50	20.43	24.97	24.62	27.00	26.59
	CRPS	11.03	**11.27**	11.71	12.30	14.10	12.52	13.64	15.15	15.09	16.54	16.24
TNN	PICP	93.26	94.51	94.39	94.55	96.08	95.31	95.92	96.01	96.64	95.54	96.50
	PINAW	**16.35**	17.44	**17.78**	**18.22**	21.34	20.34	22.05	23.23	24.79	**24.44**	26.85
	CRPS	**10.82**	11.29	**11.41**	**11.56**	13.11	12.78	13.44	14.18	14.98	15.31	17.50

Table A.23
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI evaluation set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	99.16	99.23	99.26	99.24	99.37	99.36	99.26	99.31	99.45	99.49	99.46
	PINAW	18.20	18.78	19.27	19.96	22.80	23.37	24.10	25.56	29.22	30.11	31.62
	CRPS	36.78	36.61	37.46	38.22	42.92	44.02	45.43	48.15	54.67	56.45	59.62
SVR	PICP	99.24	99.27	99.28	99.38	99.38	99.32	99.34	99.32	99.41	99.40	99.35
	PINAW	16.67	17.13	17.47	18.46	**18.92**	**19.25**	**20.18**	**20.97**	**22.63**	**23.35**	**24.31**
	CRPS	31.93	**32.64**	33.25	**35.03**	**35.94**	**36.70**	**38.51**	**40.20**	**43.46**	**45.04**	**48.14**
LSTM VEC	PICP	99.26	99.29	99.29	99.35	99.37	99.37	99.40	99.45	99.56	99.54	99.55
	PINAW	18.00	18.78	18.39	19.28	20.29	20.31	21.33	23.04	24.76	26.04	27.93
	CRPS	34.36	37.17	34.89	36.01	37.80	38.48	43.04	43.20	47.36	48.36	52.51
LSTM E-D	PICP	99.30	99.26	99.24	99.34	99.29	99.30	99.45	99.36	99.50	99.54	99.53
	PINAW	18.19	18.61	18.30	19.59	19.99	20.90	22.55	22.26	24.54	26.06	27.28
	CRPS	35.54	36.81	36.72	39.13	40.43	42.43	44.76	45.63	48.97	51.30	54.58
1D-CNN	PICP	99.16	99.26	99.36	99.32	99.34	99.36	99.35	99.38	99.42	99.42	99.47
	PINAW	16.75	16.85	17.65	18.45	20.68	20.15	21.32	23.45	24.41	25.63	27.05
	CRPS	32.25	33.76	36.29	35.24	40.59	39.50	49.18	44.80	48.03	54.06	58.61
TNN	PICP	99.19	99.25	99.25	99.27	99.29	99.27	99.28	99.36	99.38	99.39	99.41
	PINAW	**16.20**	**16.40**	**16.67**	**17.65**	19.27	19.87	20.92	22.13	23.62	23.69	25.41
	CRPS	**31.39**	33.62	**32.51**	35.49	36.64	38.33	39.77	44.74	47.52	50.72	56.04

Table A.24
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI evaluation set with confidence level 95%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	97.78	98.25	98.39	98.23	98.72	98.43	98.02	98.27	98.91	98.88	98.80
	PINAW	11.43	12.57	13.11	13.20	16.37	15.48	15.04	16.58	21.36	22.06	22.03
	CRPS	24.59	26.17	27.00	27.22	32.72	31.41	30.92	33.71	42.17	43.60	43.89
SVR	PICP	98.27	98.53	98.54	98.64	98.55	98.38	98.30	98.38	98.70	98.64	98.40
	PINAW	11.35	11.94	12.19	13.03	**13.21**	**12.66**	**13.10**	**13.93**	**15.32**	15.94	**15.35**
	CRPS	23.44	24.42	24.87	26.47	**26.88**	**26.09**	**26.99**	**28.55**	**31.09**	**32.32**	**31.53**
LSTM VEC	PICP	98.55	98.74	98.65	98.73	98.75	98.53	98.65	98.82	98.95	99.00	98.98
	PINAW	12.51	13.16	12.89	13.72	14.13	13.60	14.42	16.22	17.61	18.78	19.81
	CRPS	25.13	26.80	25.80	27.30	28.08	27.39	29.59	32.07	34.71	36.78	38.77
LSTM E-D	PICP	98.36	98.52	98.35	98.52	98.46	98.57	98.86	98.67	98.97	99.06	99.02
	PINAW	12.55	13.09	12.73	14.13	14.18	14.65	16.43	15.73	17.83	19.17	19.45
	CRPS	25.92	27.62	27.53	29.83	29.93	30.77	33.74	32.86	36.36	38.68	39.50
1D-CNN	PICP	98.16	98.29	98.52	98.54	98.70	98.32	98.51	98.65	98.60	98.72	98.63
	PINAW	11.12	11.34	12.00	12.82	14.60	12.96	13.70	16.20	16.10	17.52	17.54
	CRPS	23.19	**23.75**	24.96	26.32	30.23	26.76	30.10	32.76	32.68	36.68	36.39
TNN	PICP	98.10	98.40	98.45	98.47	98.56	98.40	98.42	98.57	98.59	98.56	98.59
	PINAW	**10.75**	**11.33**	**11.54**	**11.94**	13.72	13.30	14.35	15.04	16.02	**15.84**	17.31
	CRPS	**22.38**	23.83	**24.00**	**24.66**	27.89	27.39	29.22	30.83	32.88	33.16	38.38

Table A.25
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-3min PI evaluation set with confidence level 90%.

Model	Metric	Prediction Horizon [min]										
		15	30	45	60	75	90	105	120	135	150	165
RF	PICP	95.44	96.47	96.84	96.65	97.50	96.82	96.15	96.52	97.87	97.86	97.38
	PINAW	7.72	8.75	9.24	9.17	11.78	10.84	10.33	11.56	15.52	16.05	15.71
	CRPS	18.19	20.04	20.90	20.86	25.73	24.25	23.52	25.93	33.44	34.62	34.34
SVR	PICP	96.55	96.97	97.03	97.26	97.16	96.68	96.58	96.86	97.17	97.29	96.68
	PINAW	8.01	8.52	8.69	9.35	**9.46**	**8.86**	**9.16**	**9.82**	**10.92**	11.44	**10.80**
	CRPS	18.29	19.23	19.55	20.90	**21.20**	**20.25**	**20.92**	**22.28**	**24.46**	**25.58**	**24.60**
LSTM VEC	PICP	97.22	97.48	97.41	97.60	97.50	97.15	97.20	97.73	97.98	98.12	98.12
	PINAW	8.89	9.36	9.20	9.87	10.10	9.57	10.13	11.64	12.73	13.66	14.35
	CRPS	19.61	20.74	20.22	21.59	22.05	21.17	22.51	25.24	27.30	29.36	30.60
LSTM E-D	PICP	96.77	96.94	96.71	97.20	97.05	97.17	97.78	97.31	97.83	98.06	97.79
	PINAW	8.93	9.36	9.03	10.16	10.16	10.46	11.91	11.25	12.91	14.01	14.04
	CRPS	20.23	21.96	21.81	23.72	23.59	24.12	26.70	25.67	28.71	30.71	30.99
1D-CNN	PICP	96.17	96.49	96.91	96.87	97.30	96.50	96.60	97.53	97.32	97.45	97.24
	PINAW	7.74	**8.00**	8.51	9.12	10.38	9.01	9.44	11.54	11.38	12.48	12.29
	CRPS	17.93	**18.46**	19.31	20.67	23.47	20.60	22.24	25.70	25.47	28.28	27.58
TNN	PICP	96.31	96.69	96.71	96.81	97.30	96.83	96.99	97.25	97.37	96.99	97.28
	PINAW	**7.56**	8.06	**8.22**	**8.42**	9.86	9.40	10.19	10.74	11.46	**11.30**	12.41
	CRPS	**17.46**	18.53	**18.86**	**19.19**	22.12	21.42	22.88	24.03	25.72	25.66	30.47

Table A.26
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI test set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	99.56	99.68	99.68	99.68	99.84	99.84	99.68	99.84	99.80	99.76	99.88	99.72	99.88
	PINAW	42.72	43.37	44.22	46.41	53.15	56.13	57.79	61.34	67.85	70.81	76.89	77.57	85.31
	CRPS	18.00	17.78	17.86	18.40	20.72	22.01	22.79	24.39	26.80	28.19	30.63	31.93	34.91
SVR	PICP	99.57	99.57	99.65	99.69	99.73	99.57	99.76	99.76	99.84	99.88	99.84	99.80	100.00
	PINAW	37.70	38.02	39.54	43.06	44.98	46.39	**48.27**	50.08	54.52	57.20	60.60	**63.77**	72.61
	CRPS	**15.51**	15.63	16.14	**17.32**	**18.06**	**18.79**	**19.55**	20.42	21.97	**22.98**	24.97	**25.94**	**28.93**
LSTM VEC	PICP	99.64	99.61	99.64	99.64	99.76	99.61	99.84	99.76	99.92	99.88	99.96	99.88	99.96
	PINAW	38.37	38.14	40.01	44.09	45.60	46.48	49.46	**49.29**	54.83	**56.91**	62.11	66.23	**70.97**
	CRPS	15.78	15.60	16.55	17.52	18.53	19.21	19.74	**20.06**	**21.95**	23.83	26.12	27.49	30.24
LSTM E-D	PICP	99.45	99.61	99.69	99.69	99.80	99.76	99.84	99.92	99.92	99.92	99.92	99.92	99.92
	PINAW	**37.57**	37.59	**38.81**	**42.00**	**44.09**	46.07	48.89	50.07	**52.39**	57.03	**60.54**	65.81	72.48
	CRPS	15.55	15.41	**15.89**	17.41	18.98	20.92	22.49	23.03	24.81	26.44	28.40	29.87	32.05
1D-CNN	PICP	99.57	99.41	99.45	99.45	99.65	99.61	99.73	99.80	99.88	99.84	99.96	99.96	100.00
	PINAW	39.81	**36.92**	39.08	42.20	44.36	**45.76**	49.83	51.42	57.01	60.27	63.07	69.53	74.44
	CRPS	21.31	**15.29**	16.98	18.55	18.32	24.04	24.67	25.16	25.44	24.39	29.36	34.06	36.86
TNN	PICP	99.57	99.61	99.73	99.76	99.88	99.92	99.84	99.96	99.96	99.88	99.92	99.92	99.88
	PINAW	38.41	39.26	40.58	43.63	47.93	47.89	49.84	52.57	55.11	59.69	62.48	65.32	71.56
	CRPS	19.47	22.07	18.52	19.62	23.51	24.75	19.87	20.96	26.70	26.76	**24.75**	28.51	29.38

Table A.27
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI test set with confidence level 95%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	97.15	97.43	97.98	98.22	98.81	98.57	97.86	98.34	98.73	98.77	98.81	98.34	98.65
	PINAW	26.40	27.44	29.79	31.50	36.26	38.17	35.76	41.78	48.68	51.60	55.42	52.86	61.09
	CRPS	12.66	12.78	13.44	13.95	15.68	16.57	15.95	18.28	20.73	21.95	23.49	22.89	26.02
SVR	PICP	97.53	97.84	97.92	98.86	99.14	98.55	98.39	98.27	98.90	98.90	98.63	98.82	99.25
	PINAW	24.45	25.02	26.93	30.33	30.76	30.87	32.15	**33.22**	38.46	40.26	**40.55**	**43.64**	51.72
	CRPS	11.54	11.72	12.36	13.46	13.80	14.03	**14.52**	**15.18**	**16.87**	17.64	**17.98**	**19.01**	22.00
LSTM VEC	PICP	97.79	98.06	98.46	99.05	98.97	98.42	98.70	98.50	99.05	99.21	99.13	99.17	99.45
	PINAW	25.48	25.75	28.55	31.62	31.79	30.57	33.46	33.98	39.19	39.74	42.13	46.41	50.07
	CRPS	11.92	11.98	13.13	13.86	14.32	13.98	14.94	15.45	17.04	**17.55**	18.39	20.21	21.79
LSTM E-D	PICP	97.37	97.53	97.76	98.55	98.55	98.51	98.67	98.82	98.71	99.10	99.10	99.02	99.33
	PINAW	**23.69**	23.98	25.50	28.11	28.88	28.94	**32.02**	34.28	**35.11**	**39.01**	40.78	45.61	51.35
	CRPS	**11.31**	11.34	**11.81**	**12.72**	13.30	**13.74**	15.09	16.34	17.04	18.41	19.17	20.82	22.84
1D-CNN	PICP	97.76	97.25	97.73	98.24	98.43	98.00	98.59	98.63	99.18	99.18	98.94	99.25	99.41
	PINAW	25.26	**23.49**	**25.47**	**27.66**	**28.80**	**28.48**	32.56	34.47	40.64	43.14	41.69	47.68	**50.02**
	CRPS	13.18	**11.32**	12.55	13.41	**13.23**	15.46	16.51	17.73	18.92	18.62	19.33	22.73	23.46
TNN	PICP	97.61	98.04	98.12	98.75	98.94	98.67	98.90	98.98	99.02	99.22	99.02	99.06	99.25
	PINAW	24.71	25.15	26.72	29.34	32.31	31.09	34.38	36.69	37.50	41.98	44.49	45.63	51.09
	CRPS	12.45	13.47	12.62	13.48	15.54	15.62	15.07	16.00	17.82	19.11	18.87	20.08	**21.65**

Table A.28
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI test set with confidence level 90%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	92.83	93.78	94.46	95.05	96.00	95.53	94.10	95.33	96.75	96.75	96.60	96.08	96.99
	PINAW	17.60	18.61	20.95	22.22	25.53	26.88	24.56	29.66	35.20	37.56	40.11	37.72	44.27
	CRPS	9.83	10.13	10.89	11.36	12.70	13.38	12.73	14.80	16.90	17.95	19.06	18.39	20.99
SVR	PICP	93.84	93.88	95.06	96.59	96.31	95.57	95.88	95.14	97.06	96.86	96.20	97.25	97.80
	PINAW	17.00	17.46	19.09	21.77	21.90	21.79	22.73	**23.51**	27.81	29.14	**28.98**	**31.35**	37.63
	CRPS	9.36	9.55	10.13	11.09	11.35	11.50	**11.88**	**12.47**	13.96	14.59	**14.76**	**15.61**	18.14
LSTM VEC	PICP	94.08	94.27	95.62	96.76	96.72	94.83	96.13	95.10	97.51	96.84	97.27	97.59	97.71
	PINAW	17.84	18.09	20.54	22.74	22.69	21.48	23.75	24.28	28.43	28.64	30.20	33.47	36.21
	CRPS	9.76	9.87	10.89	11.48	11.86	11.39	12.24	12.75	14.15	**14.38**	15.01	16.45	17.67
LSTM E-D	PICP	93.06	93.76	94.51	95.80	95.73	94.16	95.73	95.61	95.37	95.80	96.20	97.14	97.92
	PINAW	**16.22**	16.58	17.87	19.89	20.23	20.10	**22.55**	24.34	**24.83**	**28.05**	29.21	32.84	37.28
	CRPS	**9.07**	**9.20**	**9.63**	**10.34**	**10.69**	**10.86**	11.99	13.01	**13.53**	14.82	15.33	16.72	18.43
1D-CNN	PICP	93.80	93.10	94.39	94.94	94.82	94.08	95.92	96.16	97.53	97.57	96.67	97.57	97.65
	PINAW	17.20	**16.21**	**17.74**	**19.45**	**20.15**	**19.23**	22.72	24.37	29.51	31.34	29.54	33.83	**34.93**
	CRPS	9.86	9.35	10.26	10.95	10.92	11.76	13.06	14.40	15.58	15.27	15.43	17.70	17.75
TNN	PICP	93.29	93.10	93.53	95.18	95.18	94.08	96.39	96.39	95.45	96.31	97.49	97.14	97.69
	PINAW	16.81	16.42	18.34	20.49	22.34	21.16	24.65	26.41	26.25	30.10	32.27	32.81	37.02
	CRPS	9.45	9.45	9.98	10.70	11.82	11.61	12.50	13.32	13.71	15.18	15.58	16.02	**17.52**

Table A.29
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI evaluation set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	99.24	99.21	99.30	99.30	99.37	99.37	99.35	99.28	99.44	99.45	99.48	99.39	99.47
	PINAW	18.78	19.07	19.44	20.40	23.36	24.68	25.40	26.97	29.82	31.13	33.80	34.10	37.50
	CRPS	33.35	33.14	33.40	34.65	39.16	41.27	42.73	45.43	50.13	52.56	57.27	59.16	64.72
SVR	PICP	99.22	99.20	99.31	99.41	99.37	99.34	99.36	99.31	99.39	99.45	99.39	99.42	99.54
	PINAW	16.57	16.71	17.38	18.93	19.77	20.39	**21.22**	22.01	23.96	25.14	26.64	**28.03**	31.92
	CRPS	**28.26**	28.48	29.51	**32.04**	**33.39**	**34.62**	**36.16**	37.62	**40.85**	**42.82**	46.63	**48.79**	54.46
LSTM VEC	PICP	99.24	99.19	99.34	99.37	99.40	99.34	99.33	99.29	99.41	99.44	99.43	99.46	99.52
	PINAW	16.87	16.77	17.59	19.38	20.05	20.43	21.74	**21.67**	24.10	**25.02**	27.30	29.11	**31.20**
	CRPS	28.80	28.68	30.26	32.67	34.41	35.48	36.97	**37.25**	41.28	44.53	49.12	51.60	55.83
LSTM E-D	PICP	99.22	99.23	99.35	99.36	99.41	99.34	99.33	99.31	99.33	99.41	99.34	99.34	99.42
	PINAW	**16.51**	16.52	**17.06**	**18.46**	**19.38**	20.25	21.49	22.01	**23.03**	25.07	**26.61**	28.93	31.86
	CRPS	28.32	28.10	**29.08**	32.19	34.90	38.11	40.67	41.26	44.12	47.17	50.62	53.67	57.96
1D-CNN	PICP	99.35	99.22	99.27	99.33	99.34	99.37	99.36	99.42	99.49	99.42	99.42	99.57	99.60
	PINAW	17.50	**16.23**	17.18	18.55	19.50	**20.12**	21.90	22.60	25.06	26.50	27.73	30.56	32.72
	CRPS	37.38	**27.76**	30.95	34.19	33.54	43.21	44.80	45.98	46.96	45.58	53.93	61.69	66.28
TNN	PICP	99.26	99.31	99.37	99.43	99.44	99.42	99.39	99.42	99.53	99.55	99.49	99.51	99.54
	PINAW	16.89	17.26	17.84	19.18	21.07	21.05	21.91	23.11	24.23	26.24	27.47	28.71	31.46
	CRPS	34.61	38.84	32.36	35.09	40.42	42.36	36.88	38.87	46.40	47.18	**46.19**	51.03	**53.76**

Table A.30
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI evaluation set with confidence level 95%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	97.91	98.19	98.36	98.50	98.60	98.58	98.16	98.59	98.81	98.83	98.86	98.61	98.90
	PINAW	11.60	12.06	13.09	13.85	15.94	16.78	15.72	18.37	21.40	22.68	24.36	23.24	26.86
	CRPS	22.11	22.48	23.89	25.14	28.52	30.04	28.73	32.93	37.91	40.22	43.19	41.79	47.66
SVR	PICP	98.17	98.34	98.50	98.78	98.60	98.54	98.54	98.52	98.83	98.77	98.72	98.74	99.11
	PINAW	10.75	11.00	11.84	13.33	13.52	13.57	14.13	**14.60**	16.91	17.70	**17.82**	**19.18**	22.73
	CRPS	19.89	20.24	21.55	24.02	24.41	24.63	**25.65**	**26.50**	30.20	31.66	**32.07**	**34.41**	40.12
LSTM VEC	PICP	98.32	98.33	98.55	98.78	98.61	98.43	98.58	98.48	98.83	98.78	98.76	98.78	98.95
	PINAW	11.20	11.32	12.55	13.90	13.97	13.44	14.71	14.94	17.23	17.47	18.52	20.40	22.01
	CRPS	20.63	20.92	23.03	25.09	25.53	24.55	26.73	27.23	30.85	**31.62**	33.52	36.91	**39.40**
LSTM E-D	PICP	98.05	98.16	98.25	98.56	98.43	98.39	98.50	98.49	98.63	98.58	98.50	98.58	98.81
	PINAW	**10.41**	10.54	11.21	12.36	12.70	12.72	**14.08**	15.07	**15.44**	**17.15**	17.93	20.05	22.57
	CRPS	**19.40**	19.56	**20.54**	**22.50**	23.33	**23.82**	26.29	28.25	**29.16**	31.92	33.41	36.87	40.94
1D-CNN	PICP	98.40	98.13	98.29	98.37	98.43	98.30	98.61	98.73	98.89	98.87	98.77	99.00	98.95
	PINAW	11.11	**10.32**	**11.19**	**12.16**	**12.66**	**12.52**	14.31	15.15	17.87	18.96	18.33	20.96	**21.99**
	CRPS	22.05	**19.37**	21.79	23.63	**23.11**	26.24	28.64	30.98	33.81	33.70	34.30	40.14	41.29
TNN	PICP	98.33	98.44	98.56	98.75	98.79	98.66	98.77	98.80	98.78	98.83	98.90	98.87	99.06
	PINAW	10.86	11.06	11.75	12.90	14.20	13.67	15.11	16.13	16.49	18.45	19.56	20.06	22.46
	CRPS	21.03	22.59	21.72	23.48	26.34	26.03	27.50	29.41	30.58	33.57	35.14	36.18	40.13

Table A.31
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-5min PI evaluation set with confidence level 90%.

Model	Metric	Prediction Horizon [min]												
		15	30	45	60	75	90	105	120	135	150	165	180	195
RF	PICP	95.87	96.32	96.90	96.96	97.38	97.27	96.32	97.05	97.72	97.78	97.66	97.11	97.81
	PINAW	7.74	8.18	9.21	9.77	11.22	11.82	10.80	13.04	15.47	16.51	17.63	16.58	19.46
	CRPS	16.09	16.76	18.44	19.53	22.04	23.27	21.88	25.69	29.96	31.95	34.15	32.70	37.66
SVR	PICP	96.30	96.49	97.06	97.51	97.39	97.14	97.05	97.08	97.81	97.79	97.41	97.49	98.16
	PINAW	7.47	7.68	8.39	9.57	9.63	9.58	9.99	**10.33**	12.22	12.81	**12.74**	**13.78**	16.54
	CRPS	15.32	15.66	16.85	18.94	19.10	19.19	**20.01**	**20.68**	23.95	25.11	**25.21**	**27.14**	31.89
LSTM VEC	PICP	96.53	96.73	97.26	97.62	97.44	96.96	97.16	97.22	97.79	97.70	97.44	97.66	97.94
	PINAW	7.84	7.95	9.03	10.00	9.97	9.44	10.44	10.67	12.50	12.59	13.27	14.71	15.92
	CRPS	16.06	16.49	18.29	20.03	20.15	19.04	20.96	21.42	24.44	**24.80**	26.12	28.88	30.89
LSTM E-D	PICP	95.96	96.28	96.77	97.17	97.00	96.75	97.17	97.31	97.08	97.39	97.14	97.50	97.80
	PINAW	**7.13**	7.29	7.86	8.74	8.89	8.83	**9.91**	10.70	**10.91**	**12.33**	12.84	14.44	16.39
	CRPS	**14.72**	**15.10**	**15.93**	**17.48**	**17.86**	**18.01**	20.03	21.64	**22.27**	24.83	25.85	28.70	32.09
1D-CNN	PICP	96.47	96.04	96.64	96.92	97.08	96.36	97.28	97.43	97.96	97.94	97.55	97.83	97.96
	PINAW	7.56	**7.13**	**7.80**	**8.55**	**8.86**	**8.45**	9.99	10.71	12.97	13.78	12.99	14.87	**15.36**
	CRPS	15.88	15.21	17.14	18.62	18.22	19.05	21.63	24.12	26.71	26.62	26.13	30.13	**30.48**
TNN	PICP	96.35	96.38	96.85	97.25	97.30	96.63	97.51	97.77	97.37	97.69	97.90	97.72	98.06
	PINAW	7.39	7.22	8.06	9.01	9.82	9.30	10.83	11.61	11.54	13.23	14.18	14.42	16.27
	CRPS	15.29	15.16	17.34	18.20	20.12	19.15	22.32	24.18	23.38	26.48	28.71	28.85	32.72

Table A.32
PI results based on PICP [%], PINAW and CRPS for the D-15min PI test set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	99.29	99.64	99.76	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	PINAW	37.38	41.68	42.16	50.89	54.34	61.17	66.98	74.62	82.35	88.26	95.14	101.14	108.44	113.66	120.65
	CRPS	11.71	12.51	12.46	14.65	15.83	18.10	19.94	22.64	24.90	26.85	29.12	31.87	33.72	36.02	38.52
SVR	PICP	99.29	99.18	99.41	99.76	99.76	99.76	99.88	100.00	100.00	100.00	99.88	99.88	100.00	100.00	100.00
	PINAW	33.26	33.47	**33.79**	38.45	43.59	**44.35**	**49.25**	56.93	63.02	66.29	66.33	74.52	88.10	94.57	100.59
	CRPS	**10.27**	10.27	**10.35**	**11.56**	**12.88**	**13.48**	**14.63**	**17.05**	**18.52**	**19.52**	**20.11**	22.32	25.70	27.42	29.11
LSTM VEC	PICP	99.53	99.53	99.41	99.88	99.76	99.76	99.88	99.88	99.88	99.88	99.88	99.88	100.00	99.88	100.00
	PINAW	34.42	34.72	34.15	39.74	45.76	45.08	50.51	**56.43**	**61.24**	**63.64**	**65.20**	**73.16**	**82.22**	**88.64**	**94.76**
	CRPS	10.87	13.94	12.02	13.75	16.57	14.10	15.38	18.82	18.62	19.80	20.67	**21.82**	**24.32**	**26.55**	**28.42**
LSTM E-D	PICP	99.76	99.40	99.52	99.76	99.88	99.88	99.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	PINAW	35.72	35.17	36.08	40.37	46.17	51.04	55.55	62.58	69.62	71.73	76.30	86.20	93.27	98.28	103.88
	CRPS	10.69	10.45	10.66	11.84	13.52	15.22	16.28	18.71	20.49	21.01	22.37	24.97	27.33	28.53	29.79
1D-CNN	PICP	99.41	99.18	99.41	99.65	99.53	99.76	99.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	PINAW	**33.07**	**31.97**	34.90	**38.15**	**42.15**	46.81	50.35	61.06	64.97	64.87	80.53	87.04	95.15	100.46	106.88
	CRPS	10.97	**10.03**	11.85	12.07	13.04	16.40	15.22	18.48	25.27	22.28	25.97	27.16	29.23	31.41	33.35
TNN	PICP	99.41	99.65	99.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	PINAW	34.61	34.85	36.05	39.56	48.98	49.70	53.00	61.71	64.00	70.62	72.61	86.95	95.94	100.34	108.63
	CRPS	11.14	10.72	10.90	11.78	14.13	14.55	15.37	17.73	18.59	20.15	21.48	24.50	27.76	28.91	31.11

Table A.33
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-15min PI test set with confidence level 95%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	97.39	97.51	98.34	99.17	99.05	99.41	99.05	99.05	99.17	99.52	99.52	99.41	99.29	99.29	99.41
	PINAW	24.72	27.40	30.13	36.44	38.44	42.21	47.43	51.97	59.77	63.92	69.28	71.93	77.83	80.87	85.87
	CRPS	8.55	9.17	9.87	11.48	12.27	13.49	14.93	16.51	18.74	20.08	21.75	22.78	24.60	25.60	27.32
SVR	PICP	97.65	97.53	97.88	98.24	99.18	98.35	99.18	98.82	99.53	99.18	98.82	98.82	99.06	99.18	99.53
	PINAW	**22.30**	22.40	**23.03**	26.84	31.73	31.40	35.52	40.04	45.58	48.06	**44.75**	**49.72**	62.16	68.20	72.99
	CRPS	**7.80**	7.79	**7.96**	**8.97**	10.22	**10.55**	**11.58**	**12.99**	14.40	15.18	**14.65**	**15.93**	19.30	20.99	22.31
LSTM VEC	PICP	97.40	97.87	98.11	97.64	99.17	97.87	98.35	99.05	99.29	98.94	98.46	98.82	99.05	99.29	99.53
	PINAW	22.76	23.94	23.21	28.33	32.44	**31.33**	**35.31**	**39.53**	43.82	45.39	45.53	51.62	**58.67**	**63.41**	**69.33**
	CRPS	7.95	9.87	9.05	10.16	11.75	10.84	11.64	13.71	**14.16**	14.83	15.10	16.46	**18.46**	**19.77**	**21.53**
LSTM E-D	PICP	98.33	97.97	97.85	98.80	99.04	99.16	99.40	99.64	99.88	99.76	99.40	99.28	99.52	99.64	99.64
	PINAW	24.46	23.51	24.50	28.72	33.21	36.61	39.86	43.96	50.40	52.38	55.66	62.40	65.91	71.16	76.24
	CRPS	8.26	7.92	8.15	9.23	10.41	11.50	12.30	13.57	15.28	15.87	16.84	18.60	19.70	21.19	22.52
1D-CNN	PICP	98.00	97.06	97.76	97.88	98.59	98.59	99.29	99.53	99.29	98.71	99.29	99.29	99.53	99.53	99.88
	PINAW	23.43	**21.82**	24.08	**26.70**	**30.23**	31.45	36.50	43.09	**42.98**	**42.39**	54.24	61.19	67.85	71.42	76.84
	CRPS	8.32	**7.77**	9.05	9.24	**10.11**	11.13	11.83	13.65	15.48	**14.22**	17.14	18.97	20.76	22.04	23.58
TNN	PICP	98.00	98.12	98.71	98.82	99.65	99.41	99.29	99.65	99.41	99.76	99.29	99.53	99.65	99.53	99.65
	PINAW	24.49	24.80	24.41	27.82	36.91	36.55	38.71	45.63	47.39	52.80	52.06	63.79	70.73	73.47	79.95
	CRPS	8.70	8.56	8.42	9.37	11.51	11.60	12.16	14.00	14.63	16.04	16.10	19.15	21.16	21.98	23.71

Table A.34
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-15min PI test set with confidence level 90%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	93.47	94.42	95.25	97.39	96.44	96.56	96.91	97.27	97.74	97.62	98.22	97.86	97.74	98.22	97.74
	PINAW	16.86	18.91	21.76	26.35	27.72	30.42	34.59	37.71	43.76	46.69	50.47	52.33	56.40	58.50	61.96
	CRPS	6.79	7.32	8.12	9.43	10.14	11.09	12.27	13.51	15.28	16.34	17.53	18.34	19.69	20.37	21.59
SVR	PICP	93.53	94.59	94.82	96.00	97.76	96.12	96.59	96.82	97.76	98.00	96.71	97.41	98.12	98.35	98.59
	PINAW	**15.63**	15.61	16.35	**19.10**	22.96	22.59	25.78	28.99	33.44	35.36	**32.38**	**35.91**	45.37	50.01	53.60
	CRPS	**6.34**	**6.32**	**6.54**	**7.32**	8.45	8.75	9.61	**10.70**	11.96	12.60	**12.08**	**13.08**	15.84	17.24	18.32
LSTM VEC	PICP	93.85	93.26	92.79	95.39	96.81	95.63	96.22	96.57	96.69	97.28	96.45	97.75	97.28	98.11	98.11
	PINAW	16.03	15.53	**15.69**	19.69	22.48	22.40	**25.27**	**28.16**	31.81	33.02	32.84	37.38	**42.73**	**46.33**	**50.96**
	CRPS	6.47	6.85	6.83	7.78	8.85	8.84	**9.47**	10.88	**11.55**	12.10	12.21	13.44	**15.08**	**16.16**	**17.62**
LSTM E-D	PICP	94.26	94.86	95.10	96.29	98.09	97.49	97.85	97.97	98.45	98.33	98.68	98.80	98.80	98.80	98.68
	PINAW	17.35	16.41	17.25	20.43	23.89	26.50	28.93	31.82	36.87	38.53	40.88	45.71	47.96	52.20	56.13
	CRPS	6.73	6.50	6.68	7.52	8.51	9.38	10.04	11.02	12.44	12.98	13.67	15.02	15.86	17.15	18.26
1D-CNN	PICP	94.71	93.29	94.12	95.06	96.00	95.53	96.82	98.59	97.18	96.94	98.35	98.47	98.59	98.94	98.71
	PINAW	16.52	**15.33**	16.82	19.17	**21.80**	**21.99**	26.62	31.45	**30.57**	**30.40**	39.17	44.68	49.57	52.20	56.32
	CRPS	6.62	6.38	7.15	7.57	**8.34**	**8.71**	9.84	11.25	12.00	**11.48**	13.67	15.25	16.69	17.60	18.82
TNN	PICP	94.71	95.29	94.94	96.35	98.82	97.65	98.12	98.35	98.71	98.94	98.47	98.82	98.94	98.94	98.82
	PINAW	17.40	17.89	17.43	20.00	27.18	26.75	28.35	33.59	34.99	39.18	38.20	46.91	52.06	54.03	58.82
	CRPS	7.06	7.11	6.97	7.84	9.61	9.59	10.03	11.52	12.05	13.29	13.14	15.62	17.06	17.75	19.07

Table A.35
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-15min PI evaluation set with confidence level 97.5%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	99.16	99.19	99.27	99.46	99.43	99.57	99.46	99.41	99.46	99.57	99.51	99.62	99.57	99.65	99.70
	PINAW	15.40	17.17	17.37	20.97	22.39	25.20	27.60	30.75	33.93	36.37	39.20	41.67	44.68	46.83	49.71
	CRPS	21.84	23.65	23.29	27.76	29.83	33.86	37.33	42.12	46.42	50.11	54.25	58.98	62.50	66.59	71.20
SVR	PICP	99.35	99.27	99.30	99.30	99.38	99.27	99.38	99.35	99.52	99.52	99.44	99.52	99.65	99.70	99.62
	PINAW	13.70	13.79	**13.92**	15.84	17.96	**18.28**	**20.29**	23.46	25.97	27.31	27.33	30.70	36.30	38.97	41.45
	CRPS	**18.55**	18.61	**18.75**	**21.24**	23.98	**24.65**	**27.14**	**31.88**	**34.79**	**36.67**	**37.59**	42.11	48.70	52.04	55.51
LSTM VEC	PICP	99.27	99.11	99.16	99.38	99.43	99.27	99.33	99.38	99.41	99.41	99.35	99.30	99.51	99.43	99.51
	PINAW	14.18	14.31	14.07	16.37	18.86	18.57	20.81	**23.25**	**25.23**	**26.22**	**26.87**	**30.15**	**33.88**	**36.52**	**39.04**
	CRPS	19.84	24.38	21.11	24.69	29.76	25.93	28.31	34.68	34.86	37.14	38.53	**41.19**	**45.98**	**50.45**	**54.07**
LSTM E-D	PICP	99.40	99.35	99.35	99.27	99.49	99.51	99.57	99.51	99.65	99.62	99.54	99.59	99.73	99.67	99.67
	PINAW	14.72	14.49	14.87	16.63	19.03	21.03	22.89	25.79	28.69	29.55	31.44	35.52	38.43	40.50	42.80
	CRPS	19.69	19.36	19.77	22.32	25.65	28.75	30.92	35.51	38.93	39.86	42.41	47.57	51.99	54.29	56.88
1D-CNN	PICP	99.14	99.11	99.17	99.25	99.30	99.35	99.27	99.41	99.65	99.30	99.52	99.57	99.76	99.68	99.70
	PINAW	**13.62**	**13.17**	14.38	**15.72**	**17.37**	19.29	20.75	25.16	26.77	26.73	33.18	35.87	39.21	41.40	44.04
	CRPS	19.63	**18.10**	20.92	21.76	**23.88**	28.96	28.30	34.44	44.42	40.82	47.62	50.43	54.82	58.53	62.34
TNN	PICP	99.33	99.33	99.30	99.27	99.49	99.38	99.38	99.44	99.60	99.60	99.46	99.52	99.73	99.70	99.68
	PINAW	14.26	14.36	14.86	16.30	20.18	20.48	21.84	25.43	26.37	29.10	29.92	35.83	39.53	41.34	44.76
	CRPS	20.24	19.44	19.97	21.84	26.67	27.24	29.07	33.71	35.27	38.46	40.79	47.11	53.09	55.40	59.71

Table A.36
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-15min PI evaluation set with confidence level 95%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	97.89	98.14	98.51	98.89	98.60	98.51	98.76	98.81	99.08	98.97	99.19	99.08	99.16	99.16	99.19
	PINAW	10.19	11.29	12.42	15.02	15.84	17.39	19.54	21.42	24.63	26.34	28.55	29.64	32.07	33.32	35.38
	CRPS	15.15	16.42	17.76	21.14	22.48	24.66	27.66	30.28	34.48	37.09	40.24	42.03	45.27	47.31	50.31
SVR	PICP	98.39	98.39	98.57	98.74	98.87	98.52	98.65	98.63	98.95	98.92	98.63	98.71	99.03	99.09	99.17
	PINAW	**9.19**	9.23	**9.49**	11.06	13.07	12.94	14.64	16.50	18.78	19.80	**18.44**	**20.49**	25.61	28.10	30.07
	CRPS	**13.42**	13.46	**13.78**	**15.88**	18.48	**18.47**	**20.71**	**23.34**	26.20	27.67	**26.26**	**29.05**	35.57	38.92	41.69
LSTM VEC	PICP	98.09	98.49	98.25	98.79	98.89	98.46	98.60	98.65	98.79	98.73	98.73	98.73	99.03	98.95	99.08
	PINAW	9.38	9.86	9.56	11.67	13.37	**12.91**	**14.55**	**16.29**	18.05	18.70	18.76	21.27	**24.18**	**26.13**	**28.57**
	CRPS	13.81	16.50	15.13	17.72	20.43	19.16	20.79	24.40	**25.64**	26.91	27.26	30.24	**33.90**	**36.62**	**40.03**
LSTM E-D	PICP	98.54	98.32	98.67	98.81	98.94	99.00	98.94	98.86	99.05	99.08	99.05	99.13	99.16	99.21	99.32
	PINAW	10.08	9.69	10.09	11.84	13.68	15.08	16.42	18.11	20.77	21.58	22.94	25.71	27.16	29.32	31.41
	CRPS	14.52	13.98	14.44	16.72	19.11	21.01	22.83	25.19	28.58	29.78	31.66	35.31	37.32	40.19	43.01
1D-CNN	PICP	98.49	97.87	98.44	98.33	98.82	98.55	98.63	98.74	98.87	98.52	98.95	99.06	99.33	99.25	99.27
	PINAW	9.65	**8.99**	9.92	**11.00**	**12.46**	12.96	15.04	17.76	**17.71**	**17.46**	22.35	25.22	27.96	29.43	31.66
	CRPS	14.29	**13.41**	15.36	16.01	**17.75**	19.05	21.26	24.93	26.64	**25.49**	31.54	35.36	38.77	41.10	44.15
TNN	PICP	98.49	98.47	98.47	98.79	99.06	98.95	98.79	98.90	98.79	98.98	98.82	99.03	99.14	99.11	99.22
	PINAW	10.09	10.22	10.06	11.46	15.21	15.06	15.95	18.80	19.53	21.76	21.45	26.28	29.15	30.28	32.94
	CRPS	15.16	15.00	14.69	16.72	21.30	21.08	22.37	26.12	27.18	30.16	30.04	36.31	40.18	41.72	45.29

Table A.37
PI results based on PICP [%], PINAW [%] and CRPS [%] for the D-15min PI evaluation set with confidence level 90%.

Model	Metric	Prediction Horizon [min]														
		15	30	45	60	75	90	105	120	135	150	165	180	195	210	225
RF	PICP	95.84	96.60	97.08	97.57	97.24	97.35	97.46	97.49	98.03	98.27	98.00	97.87	98.08	98.11	98.11
	PINAW	6.95	7.79	8.96	10.86	11.42	12.53	14.25	15.54	18.03	19.24	20.80	21.56	23.24	24.11	25.53
	CRPS	11.30	12.43	14.06	16.77	17.77	19.43	22.04	24.04	27.47	29.50	31.86	33.18	35.63	37.08	39.24
SVR	PICP	96.48	96.61	96.83	97.26	97.87	97.50	97.66	97.58	98.22	98.14	97.52	97.58	98.22	98.39	98.20
	PINAW	**6.44**	6.43	6.74	**7.87**	9.46	9.31	10.62	11.94	13.78	14.57	**13.34**	**14.80**	18.69	20.61	22.08
	CRPS	**10.43**	**10.39**	**10.82**	**12.47**	14.65	14.57	16.46	**18.46**	20.96	22.21	**20.88**	**23.06**	28.29	31.05	33.31
LSTM VEC	PICP	96.33	96.31	96.09	97.12	97.76	97.33	97.36	97.47	97.87	97.76	97.57	97.84	98.09	98.25	98.38
	PINAW	6.61	6.40	**6.47**	8.11	9.26	9.23	**10.41**	**11.60**	13.11	13.61	13.53	15.40	**17.61**	**19.09**	**21.00**
	CRPS	10.74	10.89	10.88	13.17	14.81	15.04	**16.35**	18.63	20.20	21.21	21.32	23.81	**26.80**	**28.95**	**31.75**
LSTM E-D	PICP	96.86	96.45	96.99	97.70	98.08	98.21	98.21	98.27	98.54	98.54	98.43	98.48	98.43	98.48	98.51
	PINAW	7.15	6.76	7.11	8.42	9.84	10.92	11.92	13.11	15.19	15.88	16.85	18.84	19.76	21.51	23.13
	CRPS	11.31	10.84	11.27	13.05	14.98	16.54	18.01	19.77	22.64	23.79	25.27	28.11	29.57	32.08	34.45
1D-CNN	PICP	97.15	96.58	96.45	97.09	97.74	97.34	97.85	98.01	97.74	97.26	98.06	98.04	98.47	98.36	98.44
	PINAW	6.81	**6.32**	6.93	7.90	**8.98**	**9.06**	10.97	12.96	**12.59**	**12.53**	16.14	18.41	20.43	21.51	23.21
	CRPS	10.96	10.63	11.70	12.60	**14.00**	**14.50**	16.95	19.80	**20.17**	**19.81**	24.89	28.00	30.66	32.44	34.89
TNN	PICP	96.77	97.09	96.83	97.74	98.36	98.12	98.01	98.12	98.20	98.49	98.14	98.39	98.52	98.55	98.55
	PINAW	7.17	7.37	7.18	8.24	11.20	11.02	11.68	13.84	14.42	16.14	15.74	19.33	21.45	22.26	24.24
	CRPS	11.81	12.08	11.65	13.52	17.23	16.80	17.82	20.84	21.72	24.38	23.89	29.03	31.87	33.07	35.94

References

Ahmad, F., Rasool, A., Ozsoy, E., Sekar, R., Sabanovic, A., Elitaş, M., 2018. Distribution system state estimation-A step towards smart grid. Renew. Sustain. Energy Rev. 81, 2659–2671.

Babarit, A., Delhommeau, G., 2015. Theoretical and numerical aspects of the open source BEM solver NEMOH. In: 11th European Wave and Tidal Energy Conference. EWTEC2015, In: Proceedings of the 11th European Wave and Tidal Energy Conference, Nantes, France.

Banik, A., Behera, C., Sarathkumar, T.V., Goswami, A.K., 2020. Uncertain wind power forecasting using LSTM-based prediction interval. IET Renew. Power Gener. 14 (14), 2657–2667.

Banos, O., Galvez, J.M., Damas, M., Pomares, H., Rojas, I., 2014. Window size impact in human activity recognition. Sensors 14 (4), 6474–6499.

Bao, Y., Xiong, T., Hu, Z., 2014. Multi-step-ahead time series prediction using multiple-output support vector regression. Neurocomputing 129, 482–493.

Battezzato, A., Bracco, G., Giorcelli, E., Mattiazzo, G., 2015. Performance assessment of a 2 DOF gyroscopic wave energy converter. J. Theoret. Appl. Mech. 53 (1), 195–207.

Bento, P., Pombo, J., Mendes, R., Calado, M., Mariano, S., 2021. Ocean wave energy forecasting using optimised deep learning neural networks. Ocean Eng. 219, 108372.

Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B., 2011. Algorithms for hyper-parameter optimization. Adv. Neural Inf. Process. Syst. 24.

Bracco, G., Cagninei, A., Giorcelli, E., Mattiazzo, G., Poggi, D., Raffero, M., 2016. Experimental validation of the ISWEC wave to PTO model. Ocean Eng. 120, 40–51.

Bracco, G.G., Giorcelli, E., Mattiazzo, G., 2011. ISWEC: A gyroscopic mechanism for wave power exploitation. Mech. Mach. Theory 46 (10), 1411–1424.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brownlee, J., 2018. Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery.

Burramukku, B., 2020. Estimator model for prediction of power output of wave farms using machine learning methods. arXiv preprint arXiv:2011.13130.

Casas, I., Lekube, J., 2023. Forecast of electricity generation in a wave power plant: The mutriku case study. http://dx.doi.org/10.2139/ssrn.4338614, Available at SSRN 4338614.

Castangia, M., Aliberti, A., Bottaccioli, L., Macii, E., Patti, E., 2021. A compound of feature selection techniques to improve solar radiation forecasting. Expert Syst. Appl. 178, 114979.

Cerqueira, V., Torgo, L., Mozetič, I., 2020. Evaluating time series forecasting models: An empirical study on performance estimation methods. Mach. Learn. 109 (11), 1997–2028.

Chollet, F., et al., 2015. Keras. URL: https://keras.io. (Accessed 25 October 2023).

Chu, Y., Coimbra, C.F., 2017. Short-term probabilistic forecasts for direct normal irradiance. Renew. Energy 101, 526–536.

Cummins, W., 1962. The Impulse Response Function and Ship Motions. Technical Report 1661, Technical Report, Department of the Navy, David Taylor model basin, Washington DC.

De Giorgi, M.G., Congedo, P.M., Malvoni, M., Laforgia, D., 2015. Error analysis of hybrid photovoltaic power forecasting models: A case study of mediterranean climate. Energy Convers. Manage. 100, 117–130.

Deberneh, H.M., Kim, I., 2018a. Development of monitoring and classification systems for wave energy. Int. J. Control Autom. 11 (6), 57–66.

Deberneh, H.M., Kim, I., 2018b. Predicting output power for nearshore wave energy harvesting. Appl. Sci. 8 (4), 566.

Drew, B., Plummer, A.R., Sahinkaya, M.N., 2009. A review of wave energy converter technology. In: Proc. of the Institution of Mechanical Engineers, Part a: Journal of Power and Energy. Vol. 223, Sage Publications Sage UK: London, England, pp. 887–902.

Enerdata, 2023. World power consumption: Electricity consumption 2023. URL: https://yearbook.enerdata.net. (Accessed 27 October 2023).

Eni S.p.A, 2023. ISWEC, the energy from the sea. URL: https://www.eni.com/en-IT/operations/iswec-eni.html. (Accessed 25 October 2023).

Falnes, J., 2002. Ocean Waves and Oscillating Systems. Cambridge University Press, ISBN: 9780521782111, http://dx.doi.org/10.1017/CBO9780511754630.

Fontana Crespo, R.N., Aliberti, A., Bottaccioli, L., Macii, E., Fighera, G., Patti, E., 2023. LSTM for grid power forecasting in short-term from wave energy converters. In: 2023 IEEE 47th Annual Computers, Software, and Applications Conference. COMPSAC, IEEE, pp. 1495–1500.

Gu, B., Shen, H., Lei, X., Hu, H., Liu, X., 2021a. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. Appl. Energy 299, 117291.

Gu, B., Zhang, T., Meng, H., Zhang, J., 2021b. Short-term forecasting and uncertainty analysis of wind power based on long short-term memory, cloud model and non-parametric kernel density estimation. Renew. Energy 164, 687–708.

Gueymard, C.A., 2014. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. Renew. Sustain. Energy Rev. 39, 1024–1034.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: A review. Neurocomputing 187, 27–48.

Hatalis, K., Pradhan, P., Kishore, S., Blum, R.S., Lamadrid, A.J., 2014. Multi-step forecasting of wave power using a nonlinear recurrent neural network. In: 2014 IEEE PES General Meeting| Conference & Exposition. IEEE, pp. 1–5.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Huang, C.-M., Huang, Y.-C., Huang, K.-Y., Chen, S.-J., Yang, S.-P., 2017. Deterministic and probabilistic wind power forecasting using a hybrid method. In: 2017 IEEE International Conference on Industrial Technology. ICIT, IEEE, pp. 400–405.

Ibarra-Berastegi, G., Saénz, 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.

International Energy Agency, 2023. Renewables 2022 – Analysis and forecast to 2027. URL: https://www.iea.org/reports/renewables-2022/renewable-electricity. (Accessed 27 October 2023).

Janiesch, C., Zschech, P., Heinrich, K., 2021. Machine learning and deep learning. Electron. Mark. 31 (3), 685–695.

Khalid, R., Javaid, N., 2020. A survey on hyperparameters optimization algorithms of forecasting models in smart grid. Sustainable Cities Soc. 61, 102275.
Khorramdel, B., Chung, C., Safari, N., Price, G., 2018. A fuzzy adaptive probabilistic wind power prediction framework using diffusion kernel density estimators. IEEE Trans. Power Syst. 33 (6), 7109–7121.
Khosravi, A., Nahavandi, S., 2013. Combined nonparametric prediction intervals for wind power generation. IEEE Trans. Sustain. Energy 4 (4), 849–856.
Khosravi, A., Nahavandi, S., Creighton, D., Atiya, A.F., 2010. Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans. Neural Netw. 22 (3), 337–346.
Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J., 2021. 1D convolutional neural networks and applications: A survey. Mech. Syst. Signal Process. 151, 107398.
Kline, D.M., 2004. Methods for multi-step time series forecasting neural networks. In: Neural Networks in Business Forecasting. IGI Global, pp. 226–250.
Koenker, R., Bassett, Jr., G., 1978. Regression quantiles. Econometrica 33–50.
Lauret, P., David, M., Pinson, P., 2019. Verification of solar irradiance probabilistic forecasts. Sol. Energy 194, 254–271.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Li, D., Hussain, A., Yu, X., Liu, S., Yu, X., Zhang, K., 2021. Wind power prediction based on Kalman filter and non-parametric kernel density estimation. In: 2021 IEEE/IAS Industrial and Commercial Power System Asia. I&CPS Asia, IEEE, pp. 1319–1324.
Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., Yan, X., 2019. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. Adv. Neural Inf. Process. Syst. 32.
Liu, X., Xie, C., 2014. Wind power fluctuation interval estimation based on beta distribution. Electr. Power Autom. Equip. 34 (12), 26–30.
Liu, L., Zhao, Y., Chang, D., Xie, J., Ma, Z., Sun, Q., Yin, H., Wennersten, R., 2018. Prediction of short-term PV power output and uncertainty analysis. Appl. Energy 228, 700–711.
Ljung, L., 1999. System Identification: Theory for the User, second ed. Prentice Hall PTR, ISBN: 9780136566953, p. 609.
McKelvey, T., Akcay, H., Ljung, L., 1996. Subspace-based multivariable system identification from frequency response data. IEEE Trans. Autom. Control 41 (7), 960–979.
Van der Meer, D.W., Widén, J., Munkhammar, J., 2018. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renew. Sustain. Energy Rev. 81, 1484–1512.
Merigaud, A., Ringwood, J.V., 2018. Free-surface time-series generation for wave energy applications. IEEE J. Ocean. Eng. 43 (1), 19–35.
Mousavi, S.M., Ghasemi, M., Manshadi, M.D., Mosavi, A., 2021. Deep learning for wave energy converter modeling using long short-term memory. Mathematics 9 (8), 871.
Mukherjee, S., Osuna, E., Girosi, F., 1997. Nonlinear prediction of chaotic time series using support vector machines. In: Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop. IEEE, pp. 511–520.
Nalamati, D., 2021a. Forecasting power output of wave farm using machine learning: LSTM model.
Nalamati, D., 2021b. Forecasting power output of wave farm using machine learning: Multilayer perceptron.
Neshat, M., Nezhad, M.M., Sergiienko, N.Y., Mirjalili, S., Piras, G., Garcia, D.A., 2022. Wave power forecasting using an effective decomposition-based convolutional bi-directional model with equilibrium nelder-mead optimiser. Energy 256, 124623.
Ni, C., 2021. Data-driven models for short-term ocean wave power forecasting. IET Renew. Power Gener. 15 (10), 2228–2236.
Ni, C., Ma, X., Wang, J., 2019. Integrated deep learning model for predicting electrical power generation from wave energy converter. In: Proc. of ICAC 2019. IEEE, pp. 1–6.
Ni, C., Peng, W., 2023. An integrated approach using empirical wavelet transform and a convolutional neural network for wave power prediction. Ocean Eng. 276, 114231.
Niu, D., Sun, L., Yu, M., Wang, K., 2022. Point and interval forecasting of ultra-short-term wind power based on a data-driven method and hybrid deep learning model. Energy 254, 124384.
Paduano, B., Carapellese, F., Pasta, E., Bonfanti, M., Sirigu, S.A., Basile, D., Pizzirusso, D., Faedo, N., Mattiazzo, G., 2024. Experimental and numerical investigation on the performance of a Moored pitching wave energy conversion system. IEEE J. Ocean. Eng. 49 (3), 802–820.
Paduano, B., Giorgi, G., Gomes, R.P., Pasta, E., Henriques, J.C., Gato, L.M., Mattiazzo, G., 2020. Experimental validation and comparison of numerical models for the mooring system of a floating wave energy converter. J. Mar. Sci. Eng. 8 (8).
Pasta, E., Carapellese, F., Faedo, N., Brandimarte, P., 2023. Data-driven control of wave energy systems using random forests and deep neural networks. Appl. Ocean Res. 140, 103749.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.
Pereira, N., Valério, D., Beirão, P., 2018. Control of a wave energy converter using a multi-agent system and machine learning methods. In: Proc. of RENEW 2018, October 8-10, 2018, Lisbon, Portugal. p. 387.
Prechelt, L., 2002. Early stopping-but when? In: Neural Networks: Tricks of the Trade. Springer, pp. 55–69.
Prusty, B.R., Jena, D., 2017. A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach. Renew. Sustain. Energy Rev. 69, 1286–1302.
Qin, Z., Li, W., Xiong, X., 2011. Estimating wind speed probability distribution using kernel density method. Electr. Power Syst. Res. 81 (12), 2139–2146.
Quan, H., Srinivasan, D., Khosravi, A., 2013. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans. Neural Netw. Learn. Syst. 25 (2), 303–315.
Reikard, G., 2013. Integrating wave energy into the power grid: Simulation and forecasting. Ocean Eng. 73, 168–178.
Shadmani, A., Nikoo, M.R., Gandomi, A.H., Wang, R.-Q., Golparvar, B., 2023. A review of machine learning and deep learning applications in wave energy forecasting and WEC optimization. Energy Strategy Rev. 49, 101180.
Shekar, B., Dagnew, G., 2019. Grid search-based hyperparameter tuning and classification of microarray cancer data. In: 2019 Second International Conference on Advanced Computational and Communication Paradigms. ICACCP, IEEE, pp. 1–8.
Shi, Z., Liang, H., Dinavahi, V., 2017. Direct interval forecast of uncertain wind power based on recurrent neural networks. IEEE Trans. Sustain. Energy 9 (3), 1177–1187.
Siano, P., 2014. Demand response and smart grids—A survey. Renew. Sustain. Energy Rev. 30, 461–478.
Soman, S.S., Zareipour, H., Malik, O., Mandal, P., 2010. A review of wind power and wind speed forecasting methods with different time horizons. In: North American Power Symposium 2010. IEEE, pp. 1–8.
Song, X., Wu, N., Song, S., Zhang, Y., Stojanovic, V., 2023. Bipartite synchronization for cooperative-competitive neural networks with reaction–diffusion terms via dual event-triggered mechanism. Neurocomputing 550, 126498.
Srivastava, S., Lessmann, S., 2018. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol. Energy 162, 232–247.
Sugihara, H., Funaki, T., Yamaguchi, N., 2017. Evaluation method for real-time dynamic line ratings based on line current variation model for representing forecast error of intermittent renewable generation. Energies 10 (4).
Vapnik, V., Golowich, S., Smola, A., 1996. Support vector method for function approximation, regression estimation and signal processing. Adv. Neural Inf. Process. Syst. 9, 281–287.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30.
Vissio, G., 2017. ISWEC Toward the Sea (Ph.D. thesis). Ph.D. thesis, Politecnico di Torino.
Vissio, G., Valério, D., Bracco, G., Beirão, P., Pozzi, N., Mattiazzo, G., 2017. ISWEC linear quadratic regulator oscillating control. Renew. Energy 103, 372–382.
Von Loeper, F., Schaumann, P., de Langlard, M., Hess, R., Bäsmann, R., Schmidt, V., 2020. Probabilistic prediction of solar power supply to distribution networks, using forecasts of global horizontal irradiation. Sol. Energy 203, 145–156.
Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. Renew. Energy 105, 569–582.
Wan, H., Luan, X., Stojanovic, V., Liu, F., 2023. Self-triggered finite-time control for discrete-time Markov jump systems. Inform. Sci. 634, 101–121.
Wang, Z., Wang, C., Cheng, L., Li, G., 2022. An approach for day-ahead interval forecasting of photovoltaic power: A novel DCGAN and LSTM based quantile regression modeling method. Energy Rep. 8, 14020–14033.
Wang, R., Zhuang, Z., Tao, H., Paszke, W., Stojanovic, V., 2023. Q-learning based fault estimation and fault tolerant iterative learning control for MIMO systems. ISA Trans. 142, 123–135.
Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., Sun, L., 2022. Transformers in time series: A survey. arXiv preprint arXiv:2202.07125.
Xie, D., Fu, Y., Yang, S., Yang, Y., Chen, M., 2023. Wind power interval prediction based on robust kernel density estimation. In: Journal of Physics: Conference Series. Vol. 2534, IOP Publishing, 012011.
Yuan, X., Chen, C., Jiang, M., Yuan, Y., 2019. Prediction interval of wind power using parameter optimized beta distribution based LSTM model. Appl. Soft Comput. 82, 105550.
Zhao, X., Liu, J., Yu, D., Chang, J., 2018. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data. Energy Convers. Manage. 164, 560–569.