Applied Energy 362 (2024) 123006

Contents lists available at ScienceDirect

Applied Energy

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

Data-driven approach for day-ahead System Non-Synchronous Penetration forecasting: A comprehensive framework, model development and analysis

Javier Cardo-Miota [a,b], Rohit Trivedi [a], Sandipan Patra [a], Shafi Khadem [a,*], Mohamed Bahloul [a]

[a] International Energy Research Centre, Tyndall National Institute, UCC, Cork, T12 RC5P, Ireland
[b] Department of Industrial Systems Engineering and Design, Universitat Jaume I, Castelló de la Plana, Spain

ARTICLE INFO

Keywords:
SNSP
Machine learning
Data-driven analysis
Day-ahead forecasting
Neural networks
Deep learning

ABSTRACT

This article presents a comprehensive, innovative, and data-driven approach for predicting System Non-Synchronous Penetration (SNSP) levels. It consists of iterative steps that involve data analytics and forecasting model development to overcome the challenges associated with forecasting, such as data mining or overfitting. The approach starts by defining the problem domain and identifying relevant features using the Pearson correlation method. The framework ensures that all forecasting models carry out data pre-processing uniformly. The hyperparameters, understood as adjustable external factors not learned during the training process that affect the performance and predictive ability of the forecasting model are optimized using the random search algorithm to enhance the models' performance. The study compares the performance of classical models, such as Random Forest and Light Gradient Boosting, with advanced machine learning-based models, such as Feed-forward, Gate Recurrent Unit, Short-Term Long Memory, and Convolutional Neural Network. Data from the Irish power system is chosen as a case study. The results indicate that the Feed-forward model produces the lowest errors. It has a Mean Absolute Error of about 4.09, a Root Mean Squared Error of 5.37 and a Mean Absolute Percentage Error of 18.17% respectively. This systematic and practical approach can be applied to other regions with similar challenges. This study also highlights the potential of advanced machine learning-based models in improving SNSP forecasting accuracy. The approach is beneficial for network and market operators, and ancillary service providers in smart grid network operations, with a 15-minute resolution. It provides a promising direction for future research in this area.

1. Introduction

Following the significant European Union commitments on climate action, the Irish climate action plan (CAP) 2021 [1] sets out a roadmap for reducing the emissions to half by 2030 and reaching net zero by 2050 as a commitment to the Programme for Government [2]. Among the most important measures in this plan is to reach the proportion of renewable electricity on the system up to 80% and achieve the system operation at SNSP of 95% by 2030 while keeping the curtailment levels to a minimum and a significant reduction in numbers of conventional units online [3]. Likewise, the 'Path to Net Zero Energy' report [4] released by Northern Ireland Executive plans out their ambitious renewable energy targets, including meeting at least 70% of electricity consumption from diverse renewable sources by 2030 and doubling the size of the low carbon and renewable energy economy to more than £2 billion turnover, among other objectives, to achieve the UK government's climate action goals.

The CAP also outlines the facilitation of the development of 5GW wind connections in the system by 2030. With increasing installed non-synchronous generation capacity, including interconnections, it becomes necessary to measure and limit the System Non-Synchronous Penetration (SNSP) to ensure the safe and reliable operation of the power system. SNSP is defined as a ratio of the real-time generation from non-synchronous sources (NSG) and net imports to total demand plus net exports, thus formulated as in Eq. (1) [6]:

$$SNSP_{(t)}(\%) = \frac{P_{NSG(t)} + P_{HDVC(Imports)(t)}}{P_{Demand(t)} + P_{HDVC(Exports)(t)}} \cdot 100 \tag{1}$$

Where P_{NSG} refers to the non-synchronous generation, P_{Demand} is the system demand, $P_{HDVC(Imports)}$ and $P_{HDVC(Exports)}$ and are the power imported/exported through HVDC interconnection.

Fig. 1 describes the evolution of SNSP in Ireland beginning in 2015 with a 50% ratio that increased gradually to 60% in 2017. In March

* Corresponding author.
E-mail addresses: jcardo@uji.es (J. Cardo-Miota), rohit.trivedi@tyndall.ie (R. Trivedi), sandipan.patra@tyndall.ie (S. Patra), shafi.khadem@tyndall.ie (S. Khadem), mohamed.bahloul@tyndall.ie (M. Bahloul).

https://doi.org/10.1016/j.apenergy.2024.123006
Received 15 December 2023; Received in revised form 1 March 2024; Accepted 8 March 2024
Available online 20 March 2024

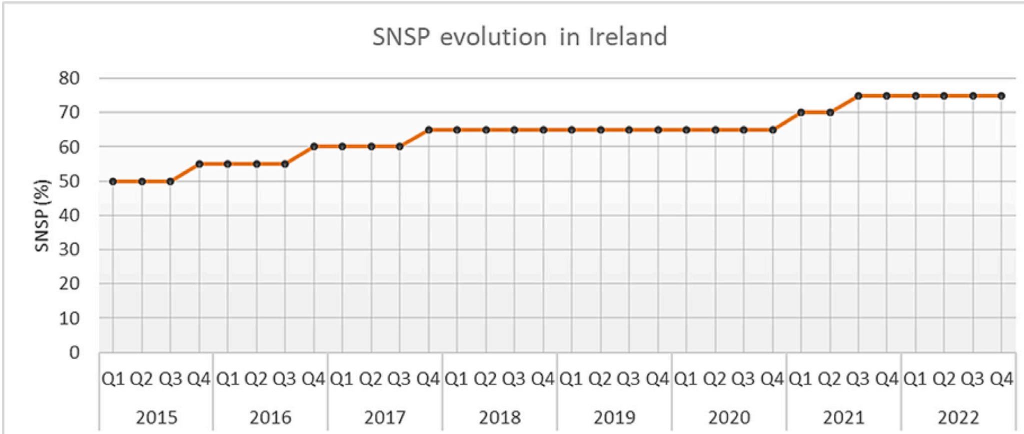

Fig. 1. SNSP evolution in Ireland.
Source: Data extracted from [5]

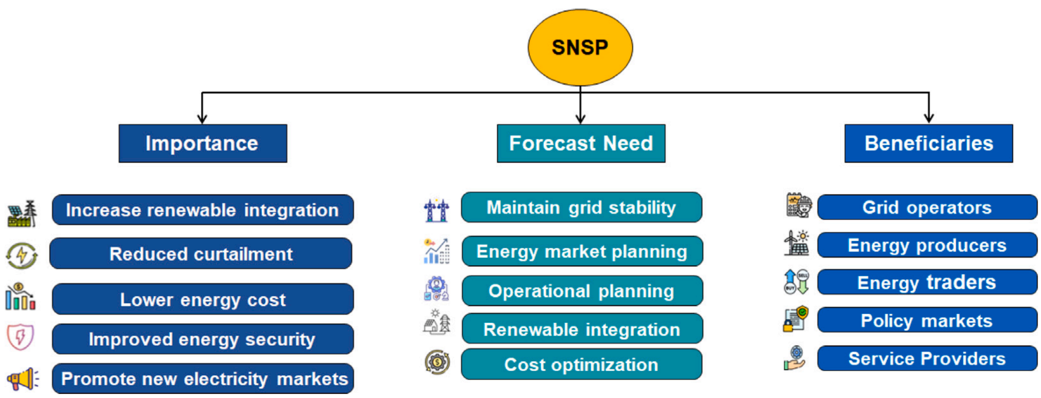

Fig. 2. Understanding the significance of the SNSP.

2018, a remarkable DS3 programme milestone was achieved when the SNSP limit on the all-island system was increased from 60% to 65%. Further increment in SNSP level to 70% occurred in March 2021. Recently, as indicating in [7], EirGrid has completed a comprehensive and detailed trial on operating the all-island power system at an SNSP level of up to 75% wherein the trial period involves 200 h of system operation with SNSP levels of more than 70%. None of the system operational insecurities during the trial were found to be a direct consequence of system operation beyond 70% SNSP levels. Hence, it is recommended to implement 75% SNSP on the system on an enduring basis.

Up to now, little effort has been made to fully comprehend the significance of precise SNSP forecasting and how it can benefit various stakeholders. This paper addresses this issue by discussing the importance of SNSP forecasting, its advantages, and the groups that can benefit from it. Section 2 presents the authors' key contributions to improving SNSP prediction accuracy, and the context within this work is developed. The literature review is included in Section 3. A discussion about the role of the SNSP in the DS3 programme is presented in Section 4. The overall methodology is explained in Section 5. In Section 6, various machine learning (ML) and deep learning (DL) models are evaluated, and their accuracy is discussed. Finally, conclusions are drawn in Section 7.

Fig. 2 illustrates the significance of SNSP in three segments: importance, need for SNSP forecast, and stakeholders who benefit from SNSP data. Subsequently, each segment is explained in detail.

1. **Importance of SNSP ratio**
   • **Increased renewable energy integration**: A higher SNSP ratio indicates a higher level of integration of vRES, such as wind and solar power, into the grid. This can reduce reliance on fossil fuel-based generation.
   • **Reduced curtailment of renewable energy**: By providing additional flexibility to power systems, higher SNSP ratio and technologies can help absorb more energy from vRES and reduce the need for curtailment [8–10].
   • **Lower costs**: A higher SNSP ratio can help lower the cost of electricity generation by reducing the need for backup generation capacity and improving energy efficiency. This can help make vRES more competitive with conventional sources of electricity [11].
   • **Improved energy security**: By reducing dependence on imported fossil fuels, a higher SNSP ratio can help improve energy security and reduce the risk of supply disruptions.
   • **Promote new electricity market framework**: A novel SNSP-based market reformulation for both the capacity and wholesale market mechanisms and the grid-supporting services market can provide a fundamental market reform framework to "bed down" these market arrangements to allow a sufficient period to be applied and deliver better results and benefits.
2. **Benefits of accurate SNSP forecast**
   • **Maintain grid stability**: Accurate SNSP forecasts can help grid operators anticipate and manage the vRES variability and unpredictability challenges [12,13] by providing insights into the expected levels of vRES generation in the grid.
   • **Energy market planning**: Energy market planning requires accurate electricity demand and supply forecasting to ensure that supply meets demand at all time-frames [14]. Accurate SNSP forecasts can help energy market planners account for the

expected levels of vRES generation and ensure that electricity supply meets demand stably and reliably.
• **Operational planning**: Power system operators must plan and manage their operations effectively to maintain grid stability and avoid blackouts [15]. SNSP forecast can help operators anticipate electricity supply and demand changes and adjust their operations accordingly.
• **Energy Policy**: SNSP forecast can help policymakers and energy stakeholders plan and manage the integration of renewable energy into the grid in a way that maximizes its benefits while minimizing its challenges [16,17].
• **Cost optimization**: Accurate SNSP forecast can help energy stakeholders optimize their operations and reduce costs by avoiding the need for expensive backup generation capacity and minimizing the curtailment of excess renewable energy [18,19].

3. **Beneficiaries of SNSP data**
• **Grid operators**: Forecasted SNSP data can help grid operators plan and manage their operations by providing insights into the expected levels of non-synchronous penetration in the grid [20]. This can help them anticipate and respond to changes in the electricity supply and demand, maintain grid stability, and avoid blackouts.
• **Independent Power Producers**: IPPs, particularly those who generate electricity from renewable sources, can benefit from forecasted SNSP data by optimizing their production schedules. They can adjust their production schedules to avoid periods of low demand and reduce curtailment of excess renewable energy [19,21].
• **Energy traders**: Forecasted SNSP data can help them make informed decisions about when to buy and sell electricity based on the expected levels of vRES generation and the resulting electricity prices [22].
• **Policymaker and Regulation Authority**: By understanding the expected levels of non-synchronous penetration in the grid, they can design policies that encourage the integration of vRES into the grid and support the transition to a more sustainable decarbonized energy system [17,23].
• **Ancillary services providers**: Grid services providers can use SNSP forecasts to efficiently and effectively manage energy production and flexibility resources operation, which helps improve the electricity system's reliability and efficiency [19].

2. Key contributions & paper context

2.1. Key contribution

To the best of our knowledge, no research paper has been published that focuses solely on forecasting SNSP. However, some studies have been conducted on the factors contributing to SNSP, such as wind energy generation and system load demand [24–27]. These studies have been carried out independently, without considering the SNSP ratio. Consequently, there is an urgent need to develop an all-inclusive SNSP forecast model that can accurately predict the SNSP for the power system, both in the day-ahead and near real-time format. Such a model would enable power system operators to manage SNSP more efficiently and optimize the use of renewable energy resources, ultimately leading to a more sustainable power grid. Therefore, the key contributions offered by this work are as follows:

• This paper presents an ML-based, data-centric method for SNSP forecasting, the first of its kind, that does not interfere with current forecast methodologies.

• An advantage of this approach is that it can be easily integrated into existing forecasting systems due to its simplicity. The data used in this study has already been forecasted by EirGrid using their established methodologies. The required data, including lagged timestamps, can be conveniently obtained from the [5].

• The proposed approach is user-friendly and can be seamlessly integrated for efficient forecasting of 24h-ahead SNSP on the same portal, thus making it a valuable tool.

2.2. Paper context

The proposed method, which utilizes machine learning and data-centric techniques, has the potential to provide highly accurate and reliable SNSP forecasts. This will enable power system operators, service providers and market operators to make better-informed decisions and optimize the use of renewable energy resources. By utilizing these estimations, service providers can create an advanced business model and gain a competitive edge by better managing their position in the market.

In Ireland, the DS3 market is considered one of the primary tools for EirGrid (TSO, Ireland) and SONI (TSO, Northern Ireland, UK) to achieve national and European targets. The procurement of DS3 services is ensured by contracting energy storage systems connected to the MV grid under the Capped and Uncapped procurement. The payment rates for Uncapped Procurement, one of the main procurement pillars, use a Scaling Factor that combines the Temporal Scarcity Scalar, which is based on SNSP forecast, with other scalars (Product Scalar, Locational Scalar, Temporal Scarcity Scalar, Performance Scalar, etc.).

The TSO calculates the Temporal Scarcity Scalar for every 30-minute interval and communicates the value to the providing units at least 2 h ahead of real-time [28]. The service provider can also define other scalars internally. Two business cases for ESS-based service providers can be developed based on the above.

The first case uses ESS to deliver only DS3 services. The service provider can manage ESS resources based on the communicated Temporal Scarcity Scalar or, in other words, the 2 h SNSP forecast from the TSO. This is depicted in Fig. 3 and colored in green.

The second business case is based on a 24 h SNSP forecast. Here, the ESS operator can develop an optimal day-ahead operation strategy to manage their position in the DS3 market better, as well as other energy trading markets. This will allow the business to stack revenues and benefits from higher market exposure, as shown in orange, as shown in Fig. 3, colored in orange.

Future work will highlight the importance of SNSP prediction and its implementation in ESS control for these business cases.

3. Literature review

3.1. vRES forecasting research

Over the years, Ireland has made significant strides in increasing its power system's share of vRES, and consequently, the SNSP ratio. As explained before, the dependency of the SNSP ratio on the RES production is well demonstrated. Hence, having an accurate prediction of the RES generation is essential to increase the capacity of estimating the SNSP ratio with high performance.

Some studies have been developed on the forecasting of vRES generation in Ireland. In [24], the authors propose a novel evolutionary covariance matrix adaptation strategy to train a neural network to obtain a short-term estimate of wind generation, demand and carbon dioxide intensity level in Ireland. This approach based on evolutionary neural networks significantly outperforms 7 traditional forecasting methods and shows its robustness when predicting with a time horizon of 2.5 h.

In [25], an ensemble variational mode decomposition (VMD) based on extreme learning machine (ELM) is proposed to predict the generation of different wind farms of Ireland. Time series data collected at turbine level are used as model inputs. Results with a prediction horizon up to 8 h ahead show that the use of the turbine-level data increase the model performance by around 20%.

Regarding the forecast of solar generation in the Irish island, hardly any works have been found due to the little amount of installed PV capacity and its current unattractive economic feasibility situation [29]. The authors in [26] proposed two classical statistical forecasting methods to obtain a short-term power generation estimation of several

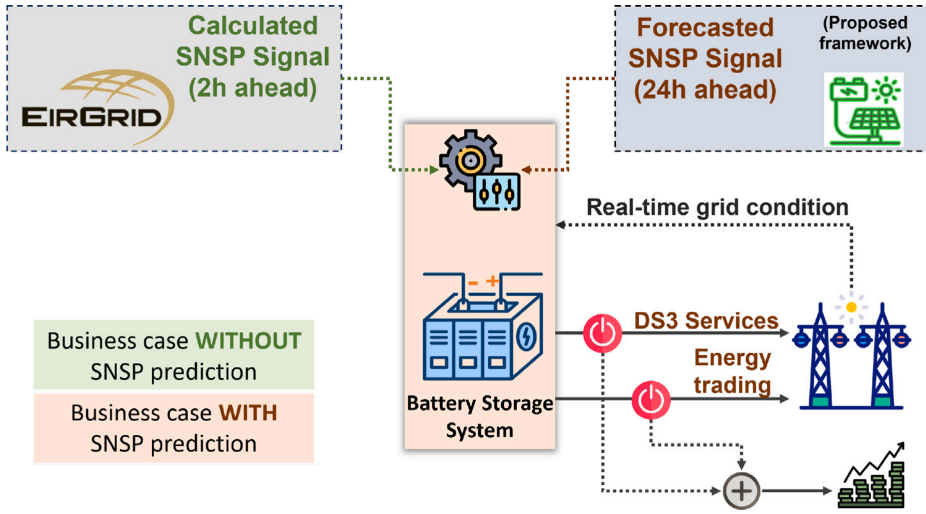

Fig. 3. Possible business cases for ESS-based service provider with and without the proposed SNSP predictions.

large-scale solar PV farms across Northern Ireland based on meteorological and historic PV generation data. The results show a relatively low error metrics, finding days with a MAE value less than 5 MW, validating the performance of the models.

3.2. SNSP research

The SNSP plays a critical role in the operation of the power system and has far-reaching benefits for stakeholders. Currently, the SNSP limit in Ireland is at 75%. However, the country aims to further increase the SNSP to 95% by 2030 [3]. Therefore, accurate forecasting of the SNSP is essential for ensuring reliable and sustainable power system operations.

In [30], the authors examined the effect of errors in offshore wind energy forecasting on the operation and management of a pool-based electricity market. The study's forecast capacity model for 2050 revealed that forecast errors could significantly affect onshore wind curtailment. In fact, the majority (80%) of dispatch-down events in 2011 were due to the SNSP operational limit. The base scenario, which assumed zero forecast error, resulted in a significant 25% curtailment of wind generation. These findings highlight the urgent need for higher SNSP limits and increased inter-connector capacity to manage the impact of forecast errors and reduce curtailment of wind generation.

The recent Ireland Capacity Outlook 2022–2031 report [27] noted that a multi-year linear regression model is currently being used for demand forecasting in Ireland and Northern Ireland. This model is based on changes in economic and historical parameters, including Real Modified Gross National Income, weather data, heat pump installations, smart meter installations, and electric vehicle installations. While this method has been effective to date, it would be interesting to explore the potential impact of machine learning (ML) models on forecasting these parameters. Such an investigation could lead to even more accurate and reliable demand forecasting.

In a recent study [31], the authors discovered that increasing SNSP limits to 75% significantly impacted curtailment in the years 2020–2021. Specifically, the study revealed that less than 20% of curtailment during this period was due to the SNSP limit. These findings suggest that raising the SNSP limit could be an effective strategy for reducing curtailment and optimizing the use of vRES.

In [32], the authors proposed an analysis of historical SNSP data in the South Australian power system to determine its impact on frequency response. They concluded that, under certain operating conditions, variation of SNSP may result in load shedding at system encounters or violation of the recommended ROCOF threshold. Consequently, it is necessary to limit SNSP to a certain value to ensure system stability and safety regarding frequency response.

In [33], the authors discussed using primary frequency response metrics to evaluate the dynamics of frequency perturbation data in Ireland. The results reflected the large influence of SNSP variation on the performance of different primary frequency response metrics such as nadir frequency, inertial and primary frequency response and rate of change of frequency.

On the other hand, the authors in [34] presented an in-depth analysis of the locational effect of the grid following (GFL) and grid forming (GFM) converters based on fast frequency response metrics in the Irish power system assuming the projected 90% SNSP limit in 2030.

4. Role of the SNSP in the Irish DS3 programme

The DS3 programme of the Irish electricity market was established in 2011 with the goal of producing at least 16% of all energy consumed by 2020 from renewable sources. This programme involves 14 ancillary services (classified into reserve, inertia, ramping, reactive power, and fast-acting services). The main target of these services is to ensure grid stability under acceptable quality and safety conditions. These services are provided through two types of contracting: Volume Capped and Volume Uncapped [28].

The trading period payment for the participation in the services under the uncapped volume procedure is calculated as described in Eq. (2).

$$\begin{aligned}\text{Trading Period Payment} &= \text{Available Volume} \cdot \text{Payment Rate} \cdot \\ &\cdot \text{Scaling Factor} \cdot \text{Trading Period Duration}\end{aligned} \quad (2)$$

Where *Payment Rates* term for each service are defined in the DS3 System Service Statement of Payments [35]. Depending on the service, the *Scaling Factor* consist of either only the Temporal Scarcity Scalar or this scalar multiplied by another scalar type such as Product Scalar, Locational Scalar, Continuous Scalar, Fast Response Scalar, Wattless Scalar and Performance Scalar.

At this point, the SNSP ratio plays an essential role within the DS3 programme: as shown in Table 1, the value of the Temporal Scarcity Scalar for each service depends solely on time-weighted average of the SNSP value for each 30-minute settlement trading period. This scalar is then applied to the tariff of the uncapped services. The TSO publish the forecast of SNSP levels at least 2 h ahead of real-time in order to help the service providers manage their units according to these values. However, a 24h-ahead SNSP estimate, such as the one proposed in this paper, would give service providers a competitive advantage, allowing

Table 1
Temporal Scarcity Scalar for the DS3 system services depending on the SNSP interval [6].

System service	Temporal Scarcity Scalar			
	SNSP <50%	*50% ≤ SNSP <60%*	*60% ≤ SNSP <70%*	*70% ≤ SNSP*
Primary Operating Reserve (POR)	1	1	4.7	6.3
Secondary Operating Reserve (SOR)	1	1	4.7	6.3
Tertiary Operating Reserve 1 (TOR 1)	1	1	4.7	6.3
Tertiary Operating Reserve 2 (TOR 2)	1	1	4.7	6.3
Replacement Reserve – Synchronized (RRS)	1	1	4.7	6.3
Replacement Reserve – Desynchronized (RRD)	1	1	4.7	6.3
Ramping Margin 1 (RM1)	1	1	4.7	6.3
Ramping Margin 3 (RM3)	1	1	4.7	6.3
Ramping Margin 8 (RM8)	1	1	4.7	6.3
Synchronous Inertial Response (SIR)	1	1	4.7	6.3
Steady State Reactive Power (SSRP)	1	1	4.7	6.3
Fast Frequency Response (FFR)	0	1	0	6.3
Dynamic Reactive Response (DRR)	0	0	0	6.3
Fast Post Fault Active Power Recovery (FPFAPR)	0	0	0	6.3

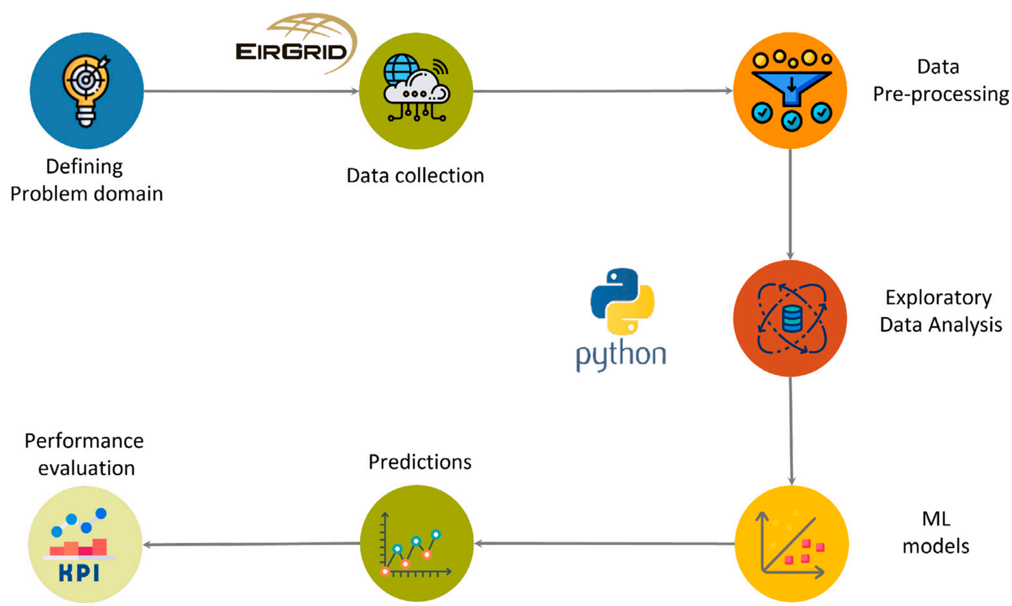

Fig. 4. SNSP data analysis and forecasting methodology.

them to define better strategies to manage their position in different electricity markets, as discussed in Fig. 3.

5. SNSP forecasting methodology

To explore the predictor capability of the different forecasting techniques presented above, a comprehensive methodology, summarized in Fig. 4, has been adopted and updated from [36]. This working framework consists of different consecutive stages that are explained below.

5.1. Defining problem domain

In the first step, the scope and objectives of the analysis are clearly defined. This involves identifying the need for exploratory data analysis to uncover patterns in the data and forecast SNSP levels. Furthermore, following the idea presented in Fig. 3 and discussed above, the current research focuses on day-ahead SNSP forecasting.

5.2. Data collection

The second step involves gathering the necessary data from reputable sources such as the EirGrid website. Historical SNSP data and other relevant variables such as load demand, wind generation (actual and forecasted), total actual system generation, and inter-connector import and export are collected in this step.

5.3. Data pre-processing

In this step, the data collected undergoes cleaning, filtering, and formatting to prepare it for analysis. The process involves eliminating any missing or inaccurate information, verifying for inconsistencies, and transforming the data into a format that is easy to analyze.

Once the data-cleansing process is complete, the inputs are prepared for models. The dataset is categorized into three broad groups based on the time series nature (dependent or independent) and available information at the time of prediction: past independent variables, future exogenous input variables, and the dependent input (target variable). The past independent variables comprise load demand, wind generation, and interconnector imports and exports. The future input variable is the forecasted wind generation data, and the dependent input is the actual historical SNSP.

In addition, the input data are normalized using the Min-Max normalization technique [37] to help the models better understand the relationships between the different inputs by transforming the data into values between 0 and 1. The mathematical expression of this technique

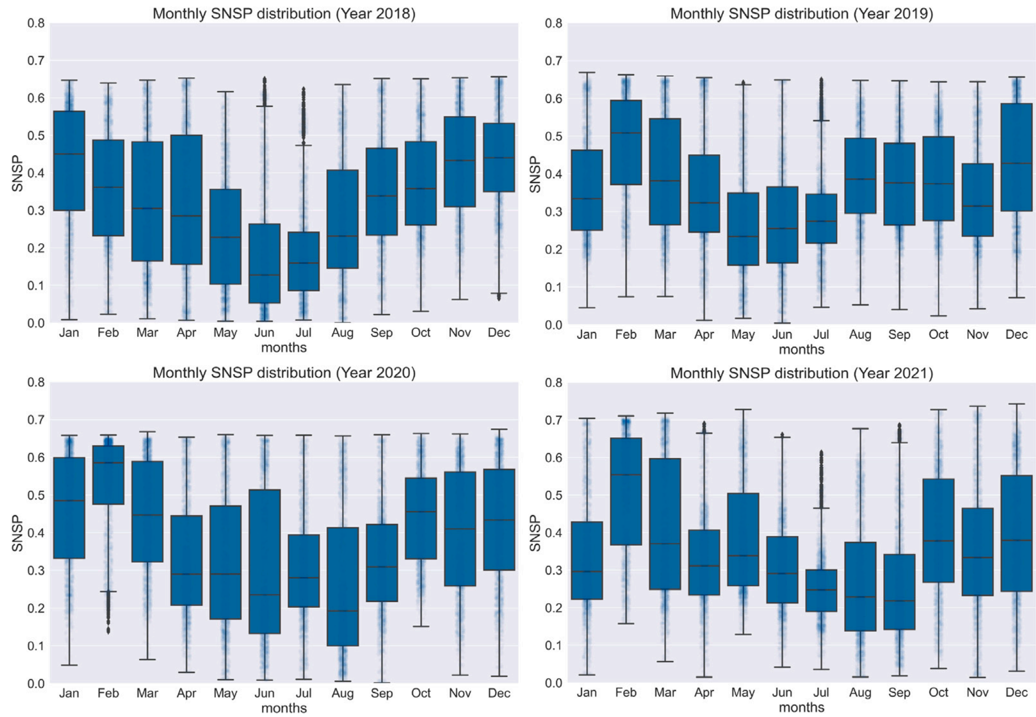

Fig. 5. Box plot for monthly SNSP distribution for a 4-year period.

is shown in Eq. (3).

$$X_{normalized} = \frac{X - X_{min}}{X_{min} - X_{max}} \tag{3}$$

5.4. Exploratory data analysis (EDA)

Once the data has been pre-processed, the subsequent step involves exploring and analyzing it. A correlation study is conducted to gain insights into the correlations among the variables, particularly concerning SNSP. Additionally, seasonality and trend analysis are also carried out.

Four years' worth of SNSP data is displayed in box plots (Fig. 5). The box plots represent the distribution of the data and show the minimum and maximum values, the median, and the first and third quartiles. The analysis focuses on identifying any patterns or trends in the data, with a particular emphasis on seasonal changes.

Analyzing the box plot of SNSP data reveals a similar pattern of seasonal changes. In each year, the SNSP ratio is low during the summer season and high during the winter season. This pattern suggests that the SNSP ratio is affected by seasonal changes, such as temperature or weather patterns. In all years except 2020, the SNSP values for the month of July consistently exceed the upper whisker or third-quartile range. This suggests that there are significant deviations from the median or typical values during this month. Additionally, the SNSP values within the inter-quartile range are lower than other seasons each year, indicating a relatively narrow distribution. However, in June 2020, the SNSP ratio exhibited a wider range of values, ranging from 15% to 52%, which is the highest compared to the SNSP ratio in June for other years. This suggests that the SNSP ratio for June 2020 was more variable and less consistent than in previous years. The analysis of SNSP data shows a wide distribution of values during winter months, particularly in January 2018, 2019 and December 2019 and 2021. This suggests that the SNSP ratio during these months exhibits significant deviations from the median or typical values, indicating that several factors may affect the ratio during these periods. Moreover, the months of February and March in 2019 and 2021 also demonstrate a wide spectrum of SNSP values, indicating that the variability in the SNSP ratio is not limited to winter months alone.

These findings underscore the importance of understanding the seasonal variations and other factors affecting the SNSP ratio. Accurate forecasting of the SNSP ratio is critical to ensure a reliable and sustainable power supply, particularly during periods of high variability in the ratio.

Fig. 6 provides valuable insights into Pearson's correlation coefficient of the SNSP ratio with other variables. The results demonstrate a clear positive correlation between the SNSP ratio and the actual and forecasted wind generation, with correlation coefficients of 0.94 and 0.92, respectively. This indicates that wind generation has a significant impact on the SNSP ratio and plays a crucial role in maintaining a stable power supply system. On the other hand, interconnected imports, exports, and load demand exhibit a slight negative correlation with the SNSP ratio and thus have a limited impact on the SNSP ratio, and their negative correlation may be attributed to other factors influencing the power supply system. The total actual generation, which includes power generation from all resources on the system, exhibits a positive correlation of approximately 0.18 with the SNSP ratio. While the correlation coefficient is relatively low, it still suggests that the total actual generation contributes to the overall power supply system and somewhat influences the SNSP ratio.

Overall, these findings demonstrate the importance of accurate forecasting of wind generation and its impact on the SNSP ratio to ensure a reliable and sustainable power supply. Understanding the correlation between variables can help develop effective strategies for managing and optimizing the power supply system. Fig. 7 depicts the distribution of load demand, wind generation, and SNSP ratio for the year 2022 through a box plot. The results indicate that the SNSP ratio exhibits almost the same variations as wind generation, suggesting that wind generation has a high impact on the SNSP ratio. However, the load demand does not demonstrate a significant pattern similarity with the SNSP ratio, as it gradually increases from summer months to winter and vice versa. The scatter plot between SNSP and wind generation reveals a very strong correlation between the two variables, as all the data points lie close to the diagonal line. This implies that the SNSP ratio

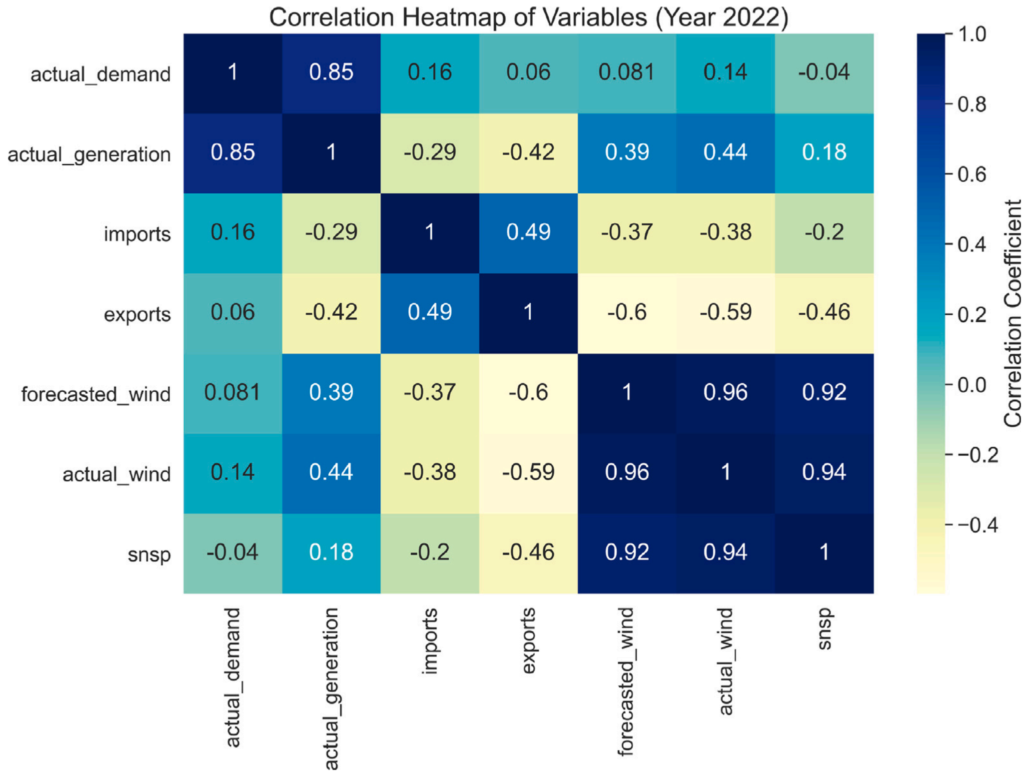

Fig. 6. Correlation heatmap of variables from Eirgrid and entso-e dataset for the year 2022.

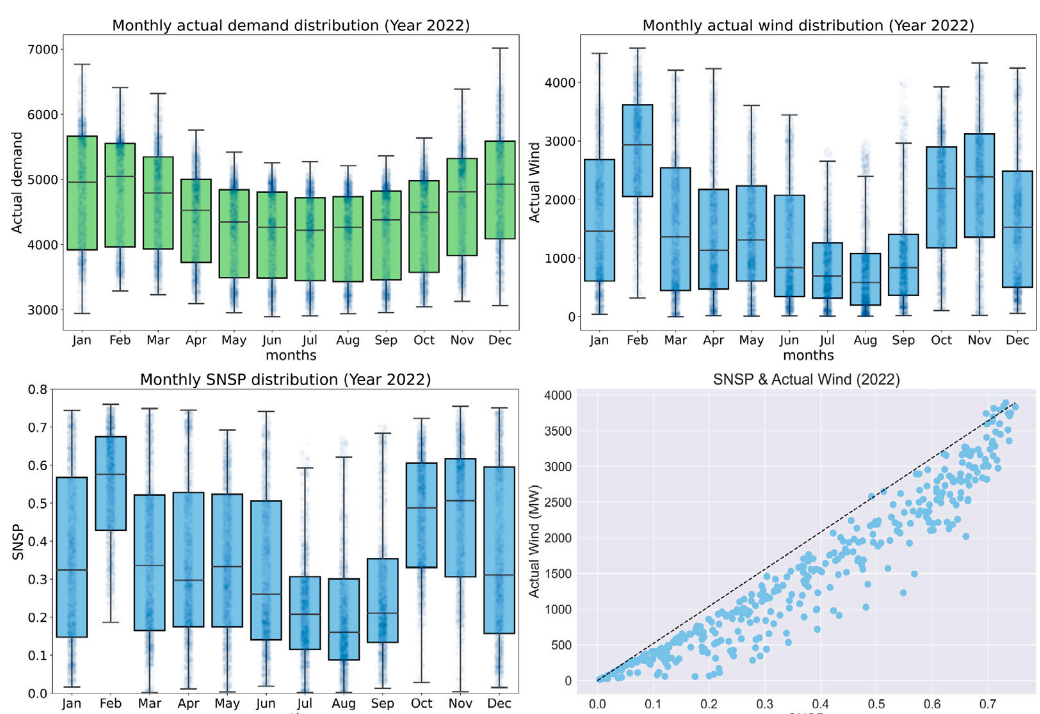

Fig. 7. Box plot distribution for load demand, wind generation and SNSP for the year 2022 with a scatter plot between SNSP and actual wind generation.

is highly dependent on wind generation, which has a direct influence on the power supply.

In general, these results emphasize the significance of precise wind generation forecasting and its influence on the SNSP ratio to guarantee a reliable and sustainable power source. Comprehending the link between variables can aid in devising efficient approaches for managing the network effectively.

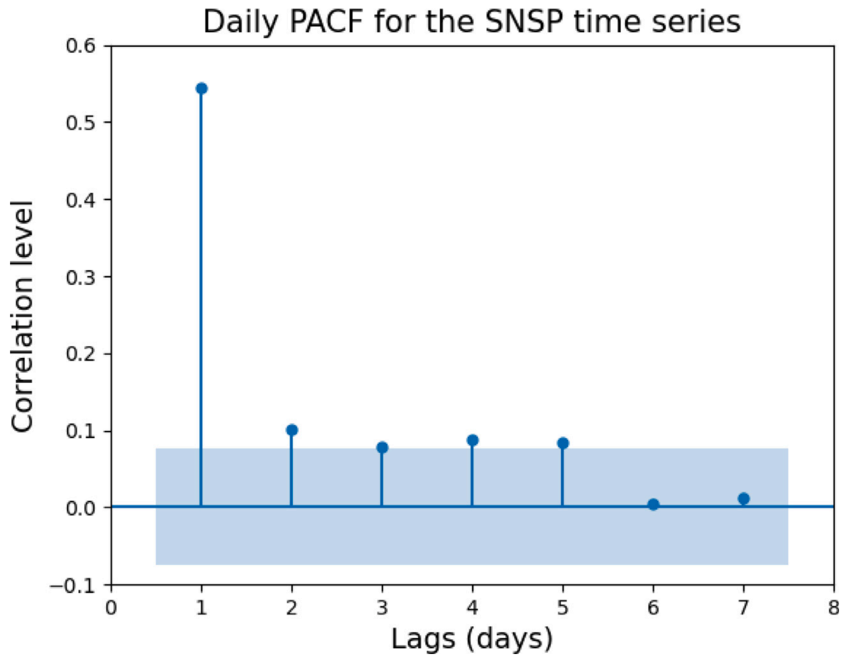

Fig. 8. Daily PACF for the SNSP time series.

A PACF analysis is conducted to assess the predictive influence of previous values in the SNSP time series. The time series data is resampled daily to determine the number of preceding days that should be incorporated into different models. Fig. 8 shows the PACF result. The X-axis shows the number of previous daily samples, or lags, considered, whereas the Y-axis shows the correlation level that the time series presents with its previous lags. Note that a 95% confidence interval is shown in shaded blue. Based on the correlation value, we can determine whether the previous lag impacts the current day. If the value falls within the interval, there is no influence. If it falls outside, there is an influence. Our findings suggest that while the previous 5 days do have a notable auto-correlation, the most significant influence is from the day before.

Finally, a stationary analysis has been performed in this work. A stationary time series keeps its statistical properties constant over time, making the prediction problem easier to analyze and deal with [38]. There are some statistical tests to determine if a time series is stationary or not, the Augmented Dickey–Fuller (ADF) test, is one of the most widely used [39]. It consists of a contrast hypothesis test where the null hypothesis (H_0) indicates that the time series is non-stationary, while the alternative hypothesis (H_1) implies the opposite. Thus, an ADF test has been used in this work to determine whether the SNSP data is stationary or not. The resultant p-value was $9.15 \cdot 10^{-27}$, considerably lower than the significance level (0.05). This result rejects H_0, concluding that the SNSP variable is a stationary time series. Therefore it is not necessary to perform any transformation on the data set before introducing it into the different forecasting models.

5.5. Build ML models

ML models, including classical and DL, are developed to analyze the data and predict future SNSP levels. This involves selecting appropriate algorithms, training the models using historical data, and fine-tuning the models to improve their accuracy. All models are programmed using Python 3.9. and Tensorflow 2.3.0 and using a laptop with an Intel Core i7 processor and NVIDIA GeForce RTX 3060 GPU.

5.5.1. Classical ML models

ML is a branch of artificial intelligence (AI) that is stirring up with high success in many fields particularly in time series forecasting [40, 41]. Within this field, Random Forest (RF) and Light Gradient Boosting (LGB) are some of the most widely used techniques in regression problems [42–44].

Fig. 9 shows the diagram followed to build the RF model. To forecast the day-ahead SNSP, the training process involves preparing a sampling dataset using the past 24 h (t-24 h) available values for each feature and future available values (t+24 h) for the forecasted wind generation feature. In the training dataset, a random forest is built with ‘n’ number of trees, where each tree is constructed by recursively splitting the data until there are fewer or equal numbers of samples in each node. In each decision tree node, the number of features is randomly selected.

The RF algorithm creates ‘k’ subsets of the training dataset. Samples that do not appear in any of these subsets are referred to as ‘out-of-bag’ samples. Each subset trains a separate model, resulting in ‘k’ trained models in the ensemble. The final prediction is generated by averaging the predictions of all individual trees in the random forest ensemble.

Hyperparameter tuning is essential to ensure the selection of the optimal model architecture. Hyperparameters are adjustable external factors not learned during the training process that affect the performance and predictive ability of the forecasting model. In particular, two hyperparameters are tuned: the number of estimators, representing the number of trees, and the node size, representing the minimum samples for node split. In this work, a random search has been performed using the Python package Scikit-learn [45] for the hyperparameter tuning of the classical ML models. The selected criterion consists of minimizing the Mean Absolute Error (MAE) for the validation metrics. The resulting hyperparameters are 400 estimators and a node size of 5.

While RF is classified as a bagging algorithm, LightGBM (LGB) operates as a boosting algorithm. The boosting process involves iteratively training decision trees on the residuals of the preceding tree. Subsequently, the predictions from all the trees are aggregated to produce a final prediction. LGB models are designed to offer high efficiency and scalability, making them particularly suitable for large-scale time series regression problems. Once again, a hyperparameter tuning process is required to obtain the model with the best architecture possible. Following the same criteria as in the RF model, the random search has been focused on three parameters: the number of estimators,

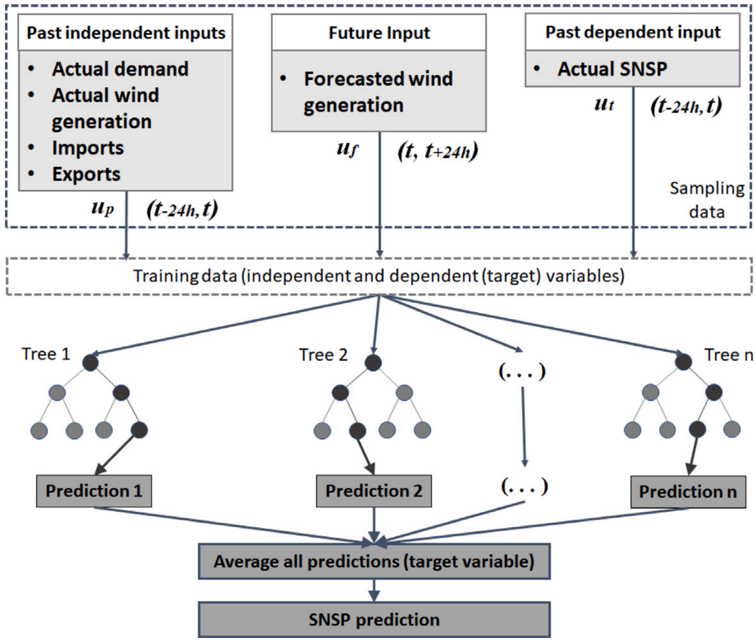

Fig. 9. RF regression tree model.

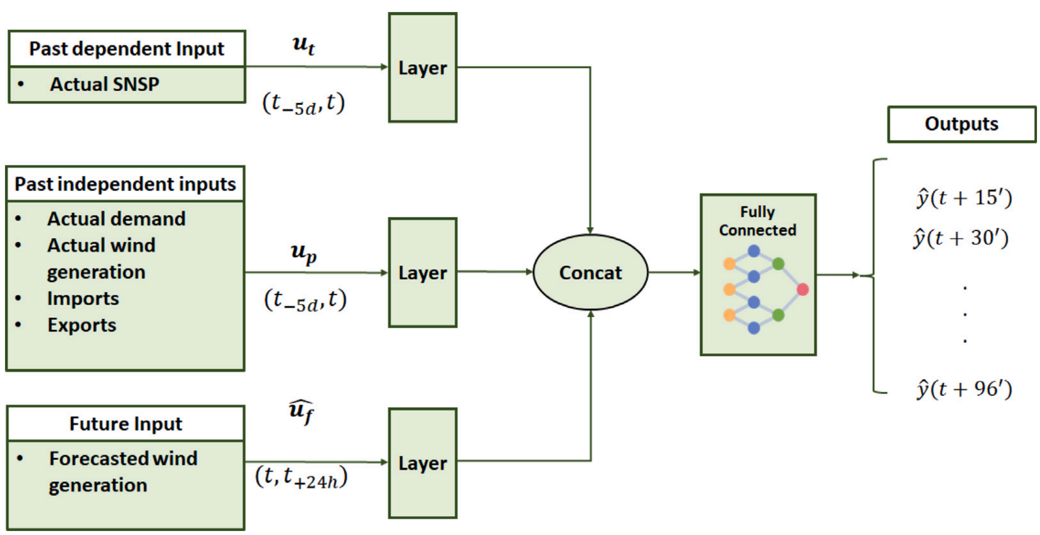

Fig. 10. Data processing diagram for deep learning models.

resulting in 1120, the maximum tree depth for the base decision trees, resulting in 12, and the learning rate, resulting in 0.1.

5.5.2. DL models

DL is an ML approach that has steadily increased over the last decades in many research fields, particularly in time series forecasting [46]. This approach based on artificial neural networks (ANN) consists of applying linear and nonlinear operations to the inputs through a series of stacked layers to obtain the outputs. Subsequently, a supervised learning process is developed to adjust the weights of the different layers of the model. Within them, feed-forward ANN [47], convolutional ANN [48] and recurrent ANN [49] such as gated recurrent unit (GRU) [50] and long-term short-term memory (LSTM) [51] are the most widely used in the field of time series forecasting.

Fig. 10 shows the data processing diagram that the DL models follow. Based on the PACF study, the five previous days (t-5 d) of SNSP are considered model inputs. In addition, the five previous days (t-5 d) of the past independent inputs and the 24 future values (t+24 h) for the forecasted wing generation feature are also considered.

Within the model development, Fig. 11 shows the different layers that are used depending on the DL model: hidden layers in the case of the feed-forward model, 1D-convolutional together with a max-pooling to reduce the computational time and the number of parameters to be learned for the convolutional model. Here, GRU and LSTM for the two different RNN models are analyzed. Subsequently, the resulting vectors per input type are flattened, concatenated and introduced into a fully connected layer before obtaining the output for the next 24 h every 15 min. Dropout layers are also used in the models to avoid over-fitting.

Finally, a hyperparameter-tuning process is performed to choose the best architecture possible. Choosing the hyperparameters that best fit the validation data is essential to ensure a model that learns the different relationships between the inputs and outputs [52,53]. In this work, a random search has been performed using the Python package Keras tuner [54] for the hyperparameter tuning of the DL models. Particularly, different numbers of layers, number of neurons or filters

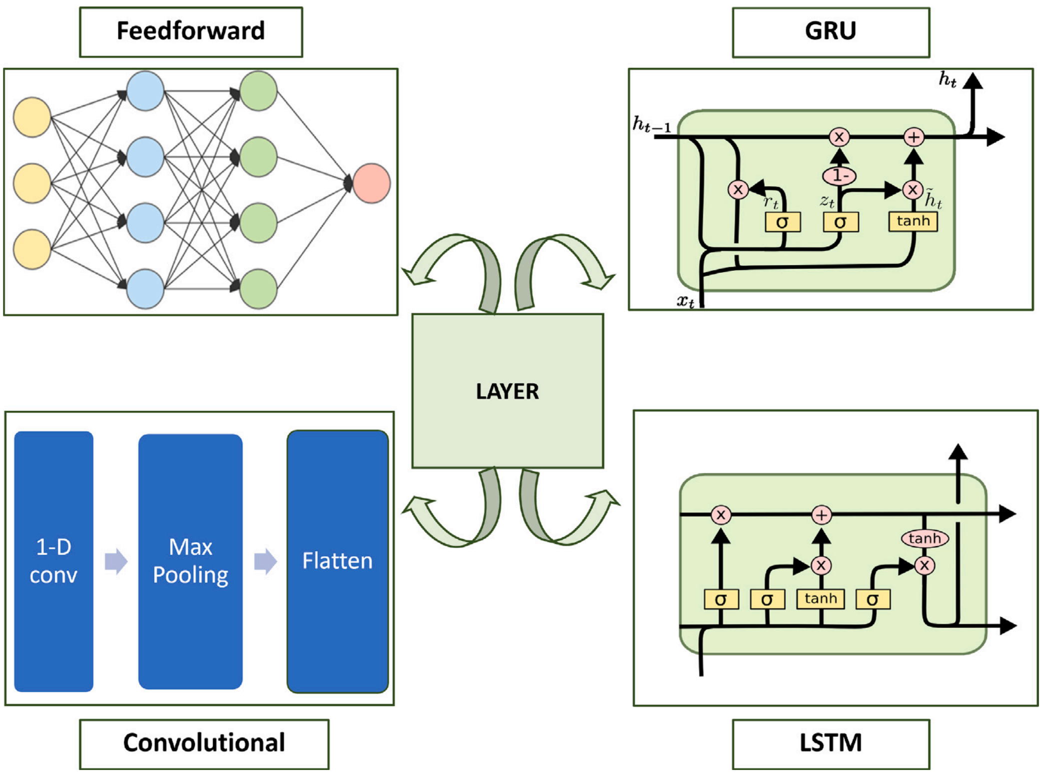

Fig. 11. The different layers used by the deep learning models.

Table 2
Hyperparameters and validation MAE of the deep learning models.

	Feedforward	1D-Conv	GRU	LSTM
Layers: Type & number	Dense: 1	1D-Conv: 1	GRU: 1	LSTM: 1
Units & Activation	75 Linear	64 Linear	40 Tanh	80 Tanh
Layer Dropout	0.3	0.4	0.05	0.05
Fully connected: Layers, units & activation	–	25 Linear	40 Linear	40 Linear, 60 Linear
Fully connected: Dropout	–	0.3	0.05	0.05
Learning rate	10^{-4}	10^{-3}	10^{-4}	10^{-4}
Validation MAE	0.0703	0.0727	0.075	0.07365

and activation functions (Linear, ReLU, Tanh) have been tested. Note that the activation function is crucial as it introduces non-linearities to the model, enabling the neural network to learn complex patterns and enhance its ability to represent non-linear relationships in the data. Using the Adam optimizer, the selected criterion minimizes the MAE for the validation metrics. Table 2 shows both the best hyperparameters and validation MAE obtained for each deep learning approach.

6. Results & discussion

In this section, the final two stages of the proposed methodology have been presented and discussed. These stages include predictions and performance evaluation. After selecting the features and hyperparameters, the resulting models are trained and tested. Note that although four years of SNSP data are presented in Fig. 5, only 2021 and 2022 data are used because the input data (specifically the import and export data) are only available for these periods. The dataset is divided into 50% for the training process, and the remaining 44% and 6% are used for validation and testing, respectively. To ensure the accuracy of the classical ML models, the distribution of the residuals has been analyzed to test their goodness of fit. Additionally, we use four error metrics to evaluate the performance of the ML models: the mean absolute error (MAE) Eq. (4), root mean squared error (RMSE) Eq. (5), the mean absolute percentage error (MAPE) Eq. (6) and adjusted coefficient of determination ($R^2_{adjusted}$) Eq. (7), where R^2 is the coefficient of determination calculated following Eq. (8) and p represents the number of independent variables.

$$MAE = \frac{1}{N}\sum_{k=1}^{N}|\hat{y}_t - y_k| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{y}_t - y_k)^2} \tag{5}$$

$$MAPE = \frac{1}{n}\sum_{k=1}^{N}\left|\frac{\hat{y}_t - y_k}{y_k}\right| \cdot 100 \tag{6}$$

$$R^2_{adjusted} = 1 - \frac{(1-R^2)\cdot(N-1)}{(N-p-1)} \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \mu)^2} \tag{8}$$

Figs. 12 and 13 present the results obtained using the classical models. Fig. 12 illustrates the LGBM-based distribution of SNSP residuals for the train and test sets. The plot shows that the residuals are densely scattered around the horizontal line, indicating that the model performs well in capturing the variability of the data. The $R^2_{adjusted}$ value for the test and train sets is 0.91 and 0.92, respectively. The scatter plot

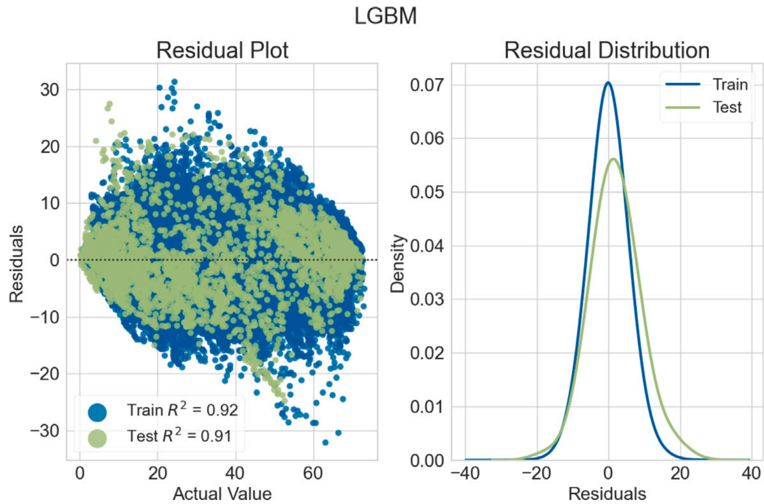

Fig. 12. Residuals for LGBM model.

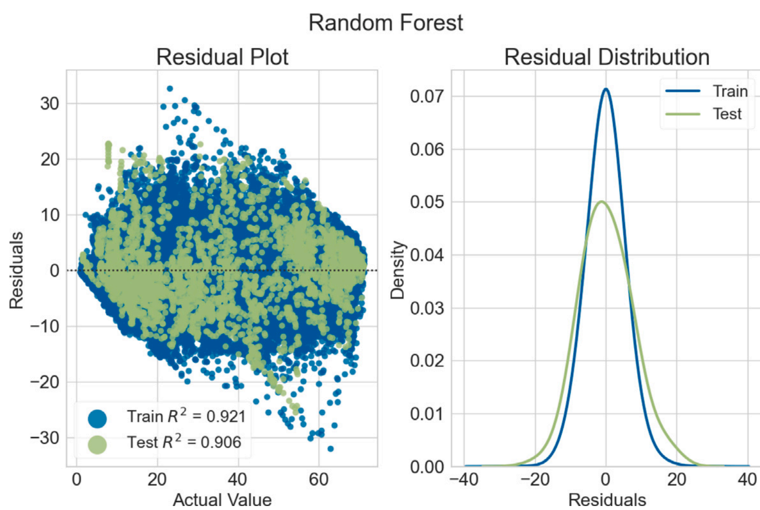

Fig. 13. Residuals for RF model.

does not exhibit any obvious pattern, suggesting that the residuals are random and independent of each other. These findings support the suitability of the model for forecasting purposes. Furthermore, the kernel density estimation (KDE) plot of residuals demonstrates a nearly normal distribution. The bell-shaped curve is symmetric and evenly distributed around zero, indicating that the residuals are unbiased and have constant variance. This KDE plot provides additional evidence that the model is appropriate for predicting future values of SNSP.

Fig. 13 depicts the residual plot and KDE of residuals for the RF model. The plot shows that the residual values are generally scattered around zero, with only a few negative residual values beyond 20. The model has a high $R^2_{adjusted}$ value of 0.906 and 0.921 for the test and training set, indicating a good fit. It also represents a similar performance than LGBM. The KDE plot on the right-hand side of the figure shows a bell-shaped curve for both the training and test data, indicating a symmetric distribution around zero. These results suggest that the RF model has good predictive power, with few outliers and a reasonably symmetrical distribution of residuals.

Fig. 14 compares the performance of SNSP forecasting accuracy by the different DL models. The dashed black diagonal line indicates a perfect prediction, whereas the blue circles show the predictions of each model, over/underestimating the value of SNSP, respectively. Furthermore, the value of the $R^2_{adjusted}$ metric of the corresponding model is highlighted in the upper left corner.

Table 3
Metrics results for test data.

Models	MAE	RMSE	MAPE (%)	MAE (%)	RMSE (%)
RF	5.71	7.17	34.40%	15.84%	19.89%
LGB	5.41	7.06	27.45%	14.99%	19.57%
Feed-forward	4.09	5.37	18.17%	11.26%	14.78%
CNN	4.42	5.71	20.40%	12.17%	15.72%
GRU	4.39	5.87	18.40%	12.00%	16.10%
LSTM	4.48	5.91	19.20%	12.20%	16.20%

The regression curves of the various models show no significant differences, as indicated by the $R^2_{adjusted}$ metrics. A common trend among the models is that they tend to underestimate, particularly for SNSP values ranging from 40 to 50%. However, it is important to note that the feedforward and convolutional models provide more accurate predictions when SNSP values exceed 60%, while the GRU and LSTM models exhibit weaker performance at higher SNSP values.

Table 3 presents the error metrics for all models on the test set. The feed-forward model produces the most accurate results, with the lowest error metrics among all models, including MAE at 4.09, RMSE at 5.37 and MAPE at 18.17%. The GRU model produces the second-lowest MAE and MAPE of 4.39 and 18.40%, respectively, further demonstrating its strong predictive capabilities.

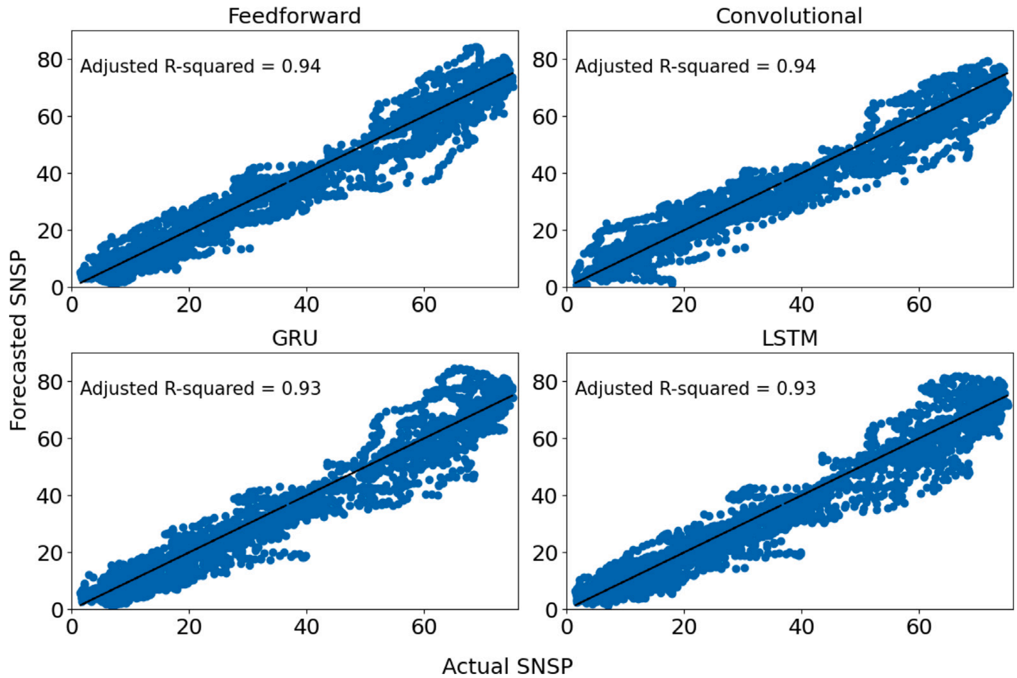

Fig. 14. Regression curves for deep learning models.

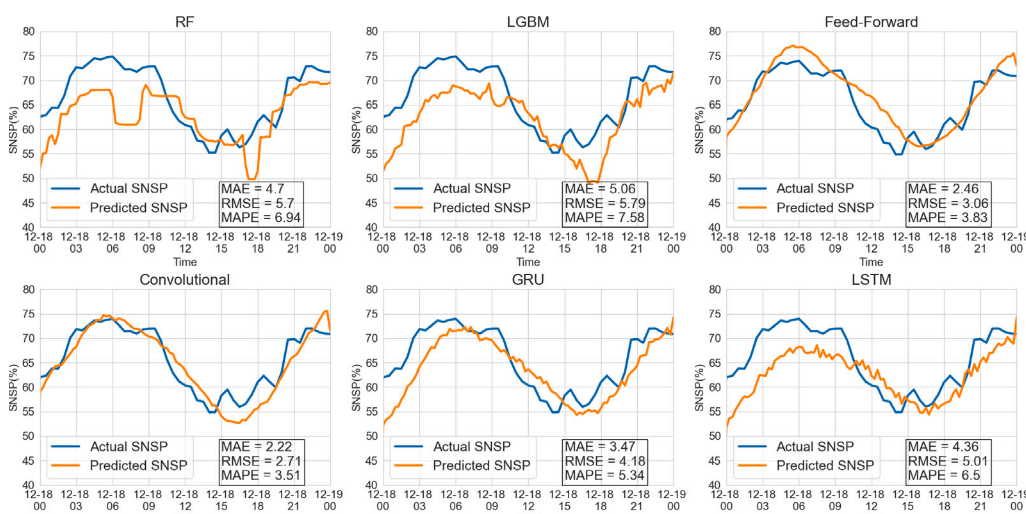

Fig. 15. 24h-ahead SNSP prediction for 18th December 2022.

Fig. 15 presents the results from all models for a 24h-ahead SNSP prediction for December 18, 2022. Error metrics are also annotated on the graph. The convolutional model exhibits the highest accuracy, with minimal errors and predictions closely following the actual values and patterns. Moreover, the feed-forward model also delivers exceptional outcomes compared to other models, demonstrating lower errors and robust predictive abilities.

As discussed in the above sections, the main goal of having a 24h-ahead SNSP prediction is to be able to estimate in advance the trading period payment for the different DS3 programme services to decide the position of the ancillary services providers in the different electricity markets. To this end, based on the SNSP predictions presented in this section, it is calculated the *Temporal Scarcity Scalar* estimations on the test data by following the dependency relation between this scalar and the SNSP interval that Table 1 presents. To evaluate these estimations, the accuracy metric, defined as the ratio between the number of correct predictions and the total number of predictions is calculated. Table 4 summarizes the obtained results. Each of the groups defined therein is conformed by the DS3 programme services that have the same scarcity scalar according to the SNSP interval:

- **Group 1**: conformed by the POR, SOR, TOR1, TOR2, RRS, RRD, RM1, RM3, RM8, SIR and SSRP services.
- **Group 2**: conformed by the FFR service.
- **Group 3**: conformed by the DRR and FPFAPR services.

Note that all models predict worse the temporal scarcity scalar of DS3 services belonging to group 2 than those belonging to group 1 or 2. As expected, ML models perform worse than DL models. The feed-forward model achieves the highest accuracy in predicting the scarcity factor, while the LGB model obtains the worst accuracy in estimating the scarcity factor.

Table 4
Accuracy metric results for the Temporal Scarcity Scalar prediction on test data.

Models	Accuracy (%)		
	Group 1	Group 2	Group 3
RF	84.21%	79.10%	91.97%
LGB	82.76%	77.05%	91.57%
Feed-forward	87.23%	83.70%	94.32%
CNN	84.98%	81.38%	93.65%
GRU	87.06%	83.50%	93.41%
LSTM	86.12%	82.83%	93.04%

7. Conclusion

This paper presents a comprehensive framework, model development and analysis for a data-driven approach to estimate system non-synchronous penetration (SNSP) level in power system networks. Real data from Ireland's power system has been considered for a case study. Six data-driven ML techniques have been used. The approach can be followed by other countries and networks in developing effective solutions to address the challenges associated with the power supply system and ensure a reliable and sustainable energy future.

In the case of Ireland, analysis of four years' worth of SNSP data indicates that wind generation plays a critical role in the stability and prediction of the power supply system. Although other variables showed weaker correlations with the SNSP ratio, they were still included in the models due to their association with actual and forecasted variables.

The statistical analysis revealed that the SNSP ratio is lower in summer due to decreased wind availability. To address this, implementing grid-scale solar projects could augment renewable energy sources and improve the SNSP ratio. However, operational constraints such as transmission limitations, grid stability issues, and power system constraints need to be considered when designing strategies for SNSP improvement.

In terms of model architectures, the hyperparameter configurations of ML models were optimized using Keras Tuner. The models' performance on test data was evaluated using four error metrics. Results showed that the DL models outperformed the classical ML models. The feed-forward model demonstrated strong predictive capabilities with the lowest MAE (4.09), RMSE (5.37) and MAPE (18.17%), while the RF model achieves the highest MAE (5.71), RMSE (7.17) and MAPE (34.40%). The remaining models have similar error metric values, with the LGB model slightly under-performing.

The higher performance of the feedforward model over the rest of ML and DL models can be attributed to various factors. If the relationship between inputs and outputs is predominantly simple and not heavily reliant on complex temporal patterns, the feedforward model may be more efficient in capturing these direct relationships. The absence of intricate temporal dependencies in input data and careful tuning of hyperparameters to align with the specific nature of the dataset can also contribute to its superior performance. Additionally, the simplicity of the feedforward model may help prevent overfitting compared to more complex models such as LSTM and GRU. These findings underscore the importance of selecting and fine-tuning models based on data characteristics to achieve optimal performance in time series forecasting applications.

Finally, future work on this topic will focus on the implementation of these SNSP estimations in a new day-ahead bidding strategy algorithm based on Reinforcement Learning that will allow the analysis of the potential and optimal participation of ESS in the Irish DS3 programme. Also, the future market reorganization to be experienced in the coming years within the European context will open the possibility of testing and improving these models in a new framework. Other deep learning algorithms, such as Encoder-Decoder, Transformers or Attention Mechanisms techniques will equally be considered for comparison in future works.

CRediT authorship contribution statement

Javier Cardo-Miota: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Rohit Trivedi:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sandipan Patra:** Writing – review & editing. **Shafi Khadem:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization. **Mohamed Bahloul:** Writing – review & editing, Methodology, 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

Data will be made available on request.

Acknowledgments

This work is a part of the COSTORE project, grant no 21/RDD/690 and the authors in IERC, Tyndall thankfully acknowledge the financial support from the Sustainable Energy Authority of Ireland (SEAI). The authors would also like to thank the financial support provided by the Universitat Jaume I (Spain) within the context of grant number, PREDOC/2020/35 and research stay grant E-2022-26.

References

[1] Goverment of Ireland. Climate action plan 2021: Securing our future. Technical Report, EirGrid; 2021, p. 207.

[2] Government Information Service. 2020 Programme for government. Technical Report June, Department of the Taoiseach; 2020, p. 128, URL https://www.gov.ie/en/publication/7e05d-programme-for-government-our-shared-future/.

[3] EirGrid, SONI. Shaping our electricity future. Technical Report, EirGrid & SONI; 2021, URL https://www.eirgridgroup.com/site-files/library/EirGrid/Shaping_Our_Electricity_Future_Roadmap.pdf.

[4] DfE. Energy Strategy for Northern Ireland: The Path to Net Zero Energy. Technical Report December, Northern Ireland Executive; 2021, p. 1–57, URL https://www.economy-ni.gov.uk/publications/energy-strategy-path-net-zero-energy.

[5] EirGrid Group. Smart grid dashboard. 2023, URL https://www.smartgriddashboard.com/. Accessed 19 September 2023.

[6] EirGrid. System non-synchronous penetration definition and formulation. Technical Report August, EirGrid; 2018, URL http://www.eirgridgroup.com/site-files/library/EirGrid/SNSP-Formula-External-Publication.pdf.

[7] Bagchi A, Nedic D, Kennedy E, Fagan E. Overview of technical studies conducted for facilitating increased renewable penetration on the island of Ireland. In: 2022 IEEE sustainable power and energy conference. 2022, p. 1–5. http://dx.doi.org/10.1109/iSPEC54162.2022.10033007.

[8] Impram S, Varbak Nese S, Oral B. Challenges of renewable energy penetration on power system flexibility: A survey. Energy Strategy Rev 2020;31:100539. http://dx.doi.org/10.1016/j.esr.2020.100539, URL https://www.sciencedirect.com/science/article/pii/S2211467X20300924.

[9] Bahloul M, Horan D, Khadem SK. BESS viability analysis for PV power plant clipping loss minimisation. In: 2023 international conference on future energy solutions. 2023, p. 1–6. http://dx.doi.org/10.1109/FES57669.2023.10182932.

[10] Bahloul M, Horan D, Khadem SK. Energy storage sizing analysis and its viability for PV power plant clipping losses minimisation. Electr Power Syst Res 2023;225:109837.

[11] EirGrid, SONI. Potential solutions for mitigating technical challenges arising from high RES-E penetration on the island of Ireland A technical assessment of 2030 study outcomes. Technical Report December, EirGrid - SONI; 2021.

[12] Al kez D, Foley AM, McIlwaine N, Morrow DJ, Hayes BP, Zehir MA, et al. A critical evaluation of grid stability and codes, energy storage and smart loads in power systems with wind generation. Energy 2020;205:117671. http://dx.doi.org/10.1016/j.energy.2020.117671, URL https://www.sciencedirect.com/science/article/pii/S0360544220307787.

[13] O'Sullivan J, Rogers A, Flynn D, Smith P, Mullane A, O'Malley M. Studying the maximum instantaneous non-synchronous generation in an island system—Frequency stability challenges in Ireland. IEEE Trans Power Syst 2014;29(6):2943–51. http://dx.doi.org/10.1109/TPWRS.2014.2316974.

[14] Eggleston J, Halley A, Mancarella P. System non-synchronous penetration (SNSP) metric and potential alternatives in low-carbon grids. In: 11th solar & storage power system integration workshop. (SIW 2021), vol. 2021, 2021, p. 30–4. http://dx.doi.org/10.1049/icp.2021.2480.

[15] Nedd M, Bukhsh W, MacIver C, Bell K. Metrics for determining the frequency stability limits of a power system: A GB case study. Electr Power Syst Res 2021;190:106553. http://dx.doi.org/10.1016/j.epsr.2020.106553, URL https://www.sciencedirect.com/science/article/pii/S0378779620303576.

[16] Raman G, Raman G, Peng JC-H. Coupled power generators require stability buffers in addition to inertia. Sci Rep 2022;12(1):13714. http://dx.doi.org/10.1038/s41598-022-17065-7.

[17] Khadem S, Bahloul M, Morch A, Papadimitriou C, Nouri A, Carroll P, Shalaby M, Stanev R, Mutule A, Efthymiou V. A dynamic process to identify the national smart grid research &innovation status and priorities. In: 2022 22nd international scientific conference on electric power engineering. IEEE; 2022, p. 1–6.

[18] Vorushylo I, Keatley P, Shah N, Green R, Hewitt N. How heat pumps and thermal energy storage can be used to manage wind power: A study of Ireland. Energy 2018;157:539–49. http://dx.doi.org/10.1016/j.energy.2018.03.001, URL https://www.sciencedirect.com/science/article/pii/S0360544218303931.

[19] Bahloul M, Daoud M, Khadem SK. Optimal dispatch of battery energy storage for multi-service provision in a collocated PV power plant considering battery ageing. Energy 2024;293:130744.

[20] Boyle J, Littler T, Foley A. Review of frequency stability services for grid balancing with wind generation. J Eng 2018;2018(15):1061–5. http://dx.doi.org/10.1049/joe.2018.0276.

[21] Drew DR, Coker PJ, Bloomfield HC, Brayshaw DJ, Barlow JF, Richards A. Sunny windy sundays. Renew Energy 2019;138:870–5. http://dx.doi.org/10.1016/j.renene.2019.02.029, URL https://www.sciencedirect.com/science/article/pii/S0960148119301739.

[22] Newbery D. Tales of two islands – lessons for EU energy policy from electricity market reforms in Britain and Ireland. Energy Policy 2017;105:597–607. http://dx.doi.org/10.1016/j.enpol.2016.10.015, URL https://www.sciencedirect.com/science/article/pii/S0301421516305602.

[23] Newbery D. National energy and climate plans for the island of Ireland: wind curtailment, interconnectors and storage. Energy Policy 2021;158:112513. http://dx.doi.org/10.1016/j.enpol.2021.112513, URL https://www.sciencedirect.com/science/article/pii/S0301421521003839.

[24] Mason K, Duggan J, Howley E. Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. Energy 2018;155:705–20. http://dx.doi.org/10.1016/j.energy.2018.04.192, URL https://www.sciencedirect.com/science/article/pii/S036054421830817X.

[25] Sopeña JMG, Pakrashi V, Ghosh B. Can we improve short-term wind power forecasts using turbine-level data? A case study in Ireland. In: 2021 IEEE madrid powerTech. 2021, p. 1–6. http://dx.doi.org/10.1109/PowerTech46648.2021.9494805.

[26] Cowan K, Liu X. Data driven solar forecasting model for Northern Ireland. In: 2021 56th international universities power engineering conference. 2021, p. 1–6. http://dx.doi.org/10.1109/UPEC50034.2021.9548211.

[27] EirGrid Group, SONI Ltd. Ireland capacity outlook. Technical Report October, EirGrid Group and SONI Ltd; 2022.

[28] Bahloul M, Daoud M, Khadem SK. A bottom-up approach for techno-economic analysis of battery energy storage system for Irish grid DS3 service provision. Energy 2022;245:123229. http://dx.doi.org/10.1016/j.energy.2022.123229, URL https://www.sciencedirect.com/science/article/pii/S0360544222001323.

[29] Assereto M, Byrne J. No real option for solar in Ireland: A real option valuation of utility scale solar investment in Ireland. Renew Sustain Energy Rev 2021;143:110892. http://dx.doi.org/10.1016/j.rser.2021.110892, URL https://www.sciencedirect.com/science/article/pii/S1364032121001866.

[30] Higgins P, Foley A, Douglas R, Li K. Impact of offshore wind power forecast error in a carbon constraint electricity market. Energy 2014;76:187–97. http://dx.doi.org/10.1016/j.energy.2014.06.037, URL https://www.sciencedirect.com/science/article/pii/S036054421400735X.

[31] Hurtado M, Kerci T, Tweed S, Kennedy E, Kamaluddin N, Milano F. Analysis of wind energy curtailment in the Ireland and Northern Ireland power systems. 2023, arXiv:2302.07143.

[32] Nahid-Al-Masood, Modi N, Saha TK, Yan R. Investigation of non-synchronous penetration level and its impact on frequency response in a wind dominated power system. In: 2016 IEEE power and energy society general meeting. 2016, p. 1–5. http://dx.doi.org/10.1109/PESGM.2016.7741587.

[33] Iswadi H, Best RJ, Morrow DJ. Irish power system primary frequency response metrics during different system non synchronous penetration. In: 2015 IEEE eindhoven powerTech. 2015, p. 1–6. http://dx.doi.org/10.1109/PTC.2015.7232425.

[34] Kez DA, Foley AM, Morrow DJ. Analysis of fast frequency response allocations in power systems with high system non-synchronous penetrations. IEEE Trans Ind Appl 2022;58(3):3087–101. http://dx.doi.org/10.1109/TIA.2022.3160997.

[35] EirGrid. DS3 system services protocol regulated arrangements DS3 system services implementation projecte1st may 2019 version 2.0. EirGrid; 2019.

[36] Trivedi R, Patra S, Khadem S. A data-driven short-term PV generation and load forecasting approach for microgrid applications. IEEE J Emerg Sel Top Ind Electron 2022;3(4):911–9. http://dx.doi.org/10.1109/JESTIE.2022.3179961.

[37] Patel C, Pandey A, Wadhvani R, Patil D. Forecasting nonstationary wind data using adaptive min-max normalization. In: 2022 1st international conference on sustainable technology for power and energy systems. 2022, p. 1–6. http://dx.doi.org/10.1109/STPES54845.2022.10006473.

[38] Nason GP. Stationary and non-stationary time series. Stat Volcanol 2006;60.

[39] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Amer Statist Assoc 1979;74(366a):427–31. http://dx.doi.org/10.1080/01621459.1979.10482531.

[40] Masini RP, Medeiros MC, Mendes EF. Machine learning advances for time series forecasting. J Econ Surv 2023;37(1):76–111. http://dx.doi.org/10.1111/joes.12429, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12429.

[41] Shinde PP, Shah S. A review of machine learning and deep learning applications. In: 2018 fourth international conference on computing communication control and automation. 2018, p. 1–6. http://dx.doi.org/10.1109/ICCUBEA.2018.8697857.

[42] Biau G. Analysis of a random forests model. J Mach Learn Res 2012;13(1):1063–95.

[43] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y. Lightgbm: A highly efficient gradient boosting decision tree. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, editors. In: Advances in neural information processing systems, vol. 30, Curran Associates, Inc.; 2017, URL https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

[44] Aksoy N, Genc I. Predictive models development using gradient boosting based methods for solar power plants. J Comput Sci 2023;67:101958. http://dx.doi.org/10.1016/j.jocs.2023.101958, URL https://www.sciencedirect.com/science/article/pii/S1877750323000182.

[45] 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. Scikit-learn: Machine learning in python. J Mach Learn Res 2011;12:2825–30.

[46] Lim B, Zohren S. Time-series forecasting with deep learning: A survey. Phil Trans R Soc A 2021;379(2194):20200209. http://dx.doi.org/10.1098/rsta.2020.0209, URL https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2020.0209.

[47] Hamzaçebi C. Improving artificial neural networks' performance in seasonal time series forecasting. Inform Sci 2008;178(23):4550–9. http://dx.doi.org/10.1016/j.ins.2008.07.024, URL https://www.sciencedirect.com/science/article/pii/S0020025508002958, Including Special Section: Genetic and Evolutionary Computing.

[48] Cardo-Miota J, Pérez E, Beltran H. Deep learning-based forecasting of the automatic frequency reserve restoration band price in the iberian electricity market. Sustain Energy, Grids Netw 2023;35:101110. http://dx.doi.org/10.1016/j.segan.2023.101110, URL https://www.sciencedirect.com/science/article/pii/S2352467723001182.

[49] Bianchi FM, Maiorino E, Kampffmeyer MC, Rizzi A, Jenssen R. Recurrent neural networks for short-term load forecasting. Springer International Publishing; 2017, http://dx.doi.org/10.1007/978-3-319-70338-1.

[50] Saini U, Kumar R, Jain V, Krishnajith M. Univariant time series forecasting of agriculture load by using LSTM and GRU RNNs. In: 2020 IEEE students conference on engineering & systems. 2020, p. 1–6. http://dx.doi.org/10.1109/SCES50439.2020.9236695.

[51] Liu X, Lin Z, Feng Z. Short-term offshore wind speed forecast by seasonal ARIMA - A comparison against GRU and LSTM. Energy 2021;227:120492. http://dx.doi.org/10.1016/j.energy.2021.120492, URL https://www.sciencedirect.com/science/article/pii/S0360544221007416.

[52] Probst P, Bischl B, Boulesteix A-L. Tunability: Importance of hyperparameters of machine learning algorithms. 2018, arXiv:1802.09596.

[53] Weerts HJP, Mueller AC, Vanschoren J. Importance of tuning hyperparameters of machine learning algorithms. 2020, arXiv:2007.07588.

[54] O'Malley T, Bursztein E, Long J, Chollet F, Jin H, Invernizzi L, et al. Kerastuner. 2019, https://github.com/keras-team/keras-tuner.