2025 IMEKO TC-6 International Conference on
Metrology and Digital Transformation - **M4DConf 2025**
Benevento, Italy, September 3-5, 2025

On-Line Frequency Forecasting using Convolutional Neural Networks

Théo Chacou Bertoldi[1], Viktoriya Mostova[2], Silvia Iuliano[2], Alfredo Vaccaro[2]

[1]*Electrical Engineering Department, Polytech Clermont, Campus des Cézeaux, Aubière, 63100, France, theochacoubertoldi@gmail.com*
[2]*Department of Engineering, University of Sannio, Piazza Roma 21, 82100, Benvento, v.mostova@studenti.unisannio.it, s.iuliano@studenti.unisannio.it, vaccaro@unisannio.it*

***Abstract* – In modern power systems the replacement of large synchronous generators with distributed inverter-based resources is lowering the power system inertia, making electrical grids more vulnerable to dynamic perturbations. For this purpose the European Network of Transmission System Operators for Electricity promoted the enhancement of the grid operation tools with specific functions for on-line estimation of the power system inertia, which are extremely useful in detecting critical thresholds and triggering proper mitigation strategies. For this purpose, the development of reliable frequency predictive models represents an essential requirement for enabling system inertia estimation. To try and address this issue, this paper explores the role of Convolutional Neural Networks (CNN) for on-line frequency estimation from grid measurements. The main idea is to model the grid frequency deviations by a CNN-based identification technique, which allows inferring the main parameters ruling the power system dynamics. Detailed simulation results obtained on several case studies demonstrate that the CNN model is able to detect data patterns, and discover hidden relationships maintaining low estimation errors even for multi-step ahead frequency predictions, thus offering a valuable tool for inertia estimation in modern power systems, especially with the increasing share of renewable energy.**

I. INTRODUCTION

The large-scale integration of renewables and high-voltage direct current transmission lines, the development of increasingly competitive electricity markets, and the expansion of power systems have posed significant challenges in modern electrical grids operation [1] [2]. In this context, the need for enabling secure and reliable power systems operation has fostered the enhancement of conventional operation tools with monitoring and forecasting functions aimed at estimating the power demands in critical areas, the active power generated by renewable generators, and the active power flows on strategic interconnections [3]. However, as power grids evolve, the overall level of uncertainty increases significantly, requiring the analysis of an ever-increasing number of potentially relevant variables [4]. Among the different variables that can be predicted, it is beneficial to focus on the frequency forecast, which is a key measure to ensure the power system resilience to dynamic perturbations, and to assess the power system adequacy [5]. In particular, frequency forecasting is of primary importance for two main reasons: it allows to identify the location of possible faults, and it allows detecting the errors in measurement or data transmission, thus acting as a backup system in case of significant discrepancies between the measured and the expected value [6]. Two distinct computing paradigms can be adopted to address the problem of frequency prediction, which rely on statistical and machine learning (ML)-based approaches, respectively. The statistical approach is based on the analysis of the system past behavior to identify patterns and trends that allow the prediction of the future values of the variables of interest, typically through the use of time series [7]. Among these methods it is worth mentioning the application of the state space model and basis functions [6], extended Kalman filter [8], Newton-type algorithms [9], discrete Fourier transform [10] and least squares techniques [11]. Although the statistical approach is well-established and widely used in several fields, it has some limitations, including the inability to capture complex non-linear relationships and reduced robustness to variations in data distribution. In this context, the deployment of ML techniques represents an interesting alternative, which is recently attracting increasing interest due to their ability to identify complex correlations that would otherwise be difficult to detect with statistical methods [7]. Indeed ML-based frequency forecasting methods are based on adaptive mathematical algorithms enabling the knowledge discovery from large amounts of data. Through this process, the analyst can identify recurring patterns, correlations, and complex relationships between the input variables. Therefore, this information can be used to build predictive models capable of generating accurate frequency forecasts and dynamically adapting to new data, continuously improving the forecasting performance. References to studies using artificial neural networks [12], genetic algorithms [13] and support vector regression [14] for frequency forecasting are available in the literature. These papers demonstrated that the deployment of ML-based methods could be the most promising enabling methodology for power system frequency prediction. Anyway, in the Authors' opinion, several open problems still need to be addressed for deploying ML-based frequency prediction models in realistic operation scenario. In particular, new methods for enhancing the adaptivity and generalization capabilities of ML-based models are required to effectively address the complex time-varying phenomena characterizing frequency variations in modern power system operation. To try and address this issue, this paper explores the potential role of CNN to progressively learn the complex features of the frequency variations, starting from low-level elements to more abstract concepts, through a series of convolutional layers. It is expected that the distinctive features of CNN, and in particular their highly adaptivity levels, allow inferring the main parameters ruling the frequency variation dynamics. Detailed simulation results obtained on several case studies demonstrate that the CNN model is able to detect data patterns, and discover hidden relationships maintaining low estimation errors even for multi-step ahead frequency predictions, thus offering a valuable tool for inertia estimation in modern power systems, especially with the increasing share of renewable energy.

II. A CNN-BASED MODEL FOR FREQUENCY PREDICTION

A CNN-based method is proposed to predict the frequency evolution over several time frames. The main idea is to deploy a CNN to extract features from frequency time series using convolution and pooling operations. Convolution is a the first step in feature extraction, inspired by signal analysis, which consists of applying a filter (also called kernel) on the input matrix in order to identify local patterns. The result of this operation is a feature map, which highlights the most relevant aspects of the frequency variations. For this purpose, several variants of convolution can be deployed to improve both the efficiency and accuracy of the knowledge extraction process, including standard convolution, deformable convolution, dilated convolution, and padding convolution. In particular, the convolution formula, shown in Eq.(1), has been adopted in this study:

$$S(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n) \cdot K(m,n) \qquad (1)$$

where $S(i,j)$ is the frequency value of the convolution, $I(i+m, j+n)$ is the value of the input variable, $K(m,n)$ is the value of the kernel (or filter) at (m,n), which are the dimensions of the filter. After applying convolution, the resulting feature maps contain a large number of features, which can lead to a risk of overfitting. To reduce this risk, pooling (or downsampling) is introduced, which acts as a dimensionality reduction technique, keeping only the most important information. There are two main types of pooling: Max Pooling and Average Pooling. Max Pooling selects the maximum value in each region of the feature map as shown in Eq.(2):

$$P(i,j) = \max_{m,n \in R(i,j)} I(m,n) \qquad (2)$$

where $P(i,j)$ is the max pooling output value at position (i,j) and $R(i,j)$ is the input $f \times f$ size region considered for pooling. Average Pooling calculates the average of the values present within each region of the feature map, as shown in Eq.(3):

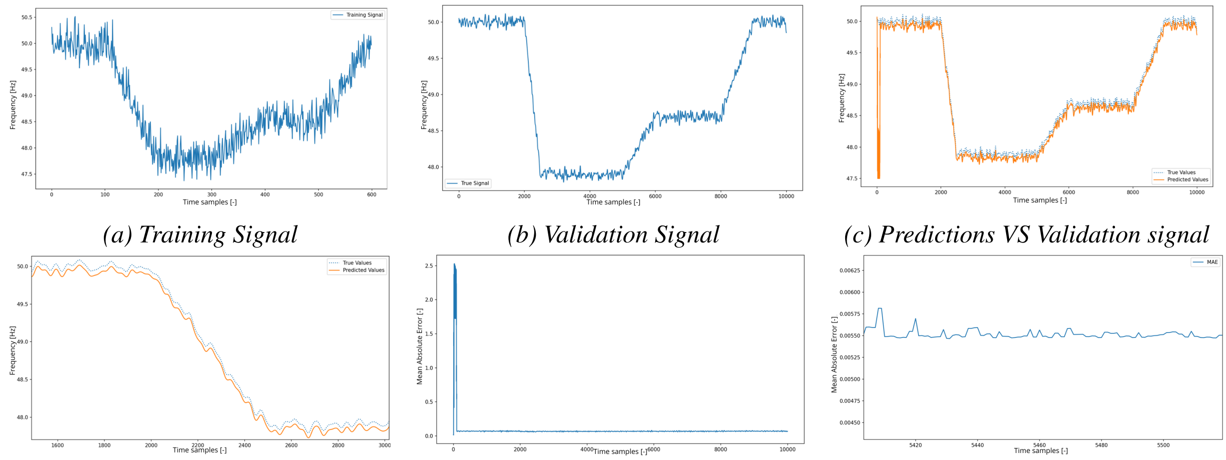
$$P(i,j) = \frac{1}{f^2} \sum_{m,n \in R(i,j)} I(m,n) \qquad (3)$$

where $P(i,j)$ is the average pooling output value at i,j, $R(i,j)$ is the region of dimension $f \times f$ and $1/f^2$ it is the normalization to obtain the mean.

Hence the overall computational process can be summarized as follows:

1. The raw frequency data, reflecting an experimental profile, are received by the input layer;
2. The convolutional layers apply filters to detect features, whose parameters are updated during the learning process. A key step in this process is the selection of the activation function, which introduces nonlinearity into the model, allowing the network to represent more complex relationships;
3. The Pooling Layer reduce the dimensionality of the feature maps, improving computational efficiency;
4. The feature maps are flattened into a vector, which is then processed by a feedforward neural network in the so-called fully connected layers;
5. The output layer produces the final result.

The expected advantages of using this computational process mainly derive from some key features characterizing CNN-based data processing. First, there are local connections; instead of connecting each neuron to all those in the previous layer, each neuron is connected to only a tiny portion of the neurons, which significantly reduces the number of parameters and accelerates the model convergence. Furthermore, CNNs exploit weight sharing, a principle according to which a group of connections can share the exact weights, reducing the number of parameters to be learned. Another important aspect is dimensionality reduction through downsampling; the pooling layers apply the principle of local correlation, which allows reducing the amount of data to be processed and preserving the relevant information.



(a) Training Signal *(b) Validation Signal* *(c) Predictions VS Validation signal*

(d) Zoom on predictions *(e) MAE profile* *(f) Zoom on warm-up zone*

Fig. 1. Ultra Short-term Frequency Prediction: First Approach

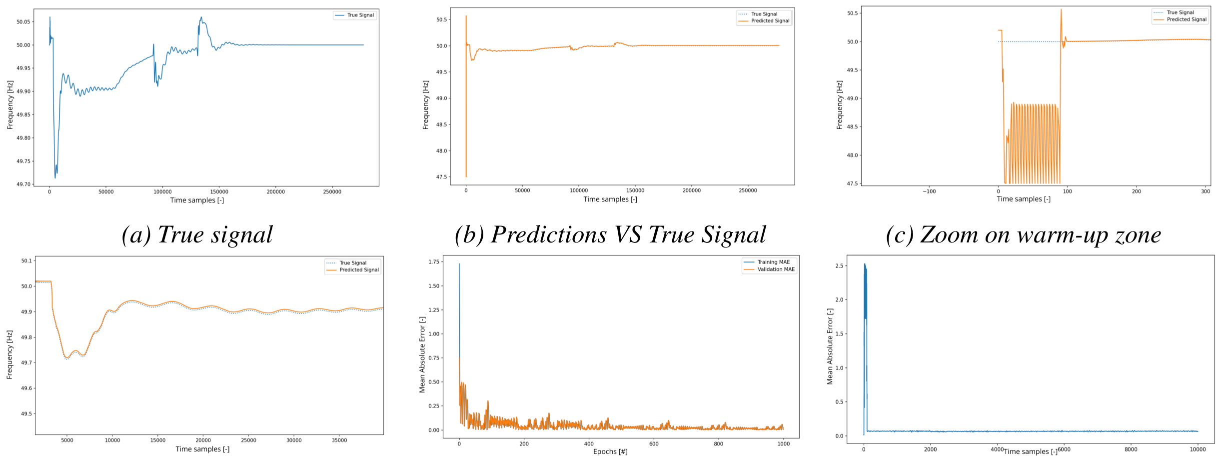
III. CASE STUDY

The CNN-based frequency prediction model has been tested in the task of analyzing sequential frequency data, reflecting an experimental profile, receiving, as input to the neural network, the frequency value at time t. The first convolutional layer uses 16 filters to identify patterns in the data, followed by a subsampling layer, MaxPooling, that reduces complexity by keeping only the most relevant information. This process is repeated with convolutional layers of 32 and 64 filters, alternating with subsampling, progressively refining the analysis. Next, GlobalMaxPooling summarizes the data in one value per filter, and the fully connected layers use this information to recognize patterns and generate predictions or classifications. Therefore, the CNN takes a frequency as input and predicts the next using the direct method.

A. *Ultra Short-term Frequency Prediction*

During the first experimental activity, the training and validation process are separated. Training involves creating, compiling, and training the CNN on a complex signal that simulates variations. The trained model is then saved for validation and tested with accurate data. This method does not allow for hyperparameter optimization, but it does reduce the development time. In particular, the training process starts by generating signals of different natures combined to generate a complex time series. A CNN, composed of multiple convolution and pooling layers, is trained on this data to predict frequency variations, minimizing the mean absolute error (MAE). Once training is complete, the model is saved, and its performance is evaluated through graphical representations that illustrate the trend of the predicted frequencies and the MAE evolution, providing an in-depth analysis of the model's ability to anticipate frequency variations in the system. The signal used to train the CNN varies between 47.5 Hz and 51.5 Hz, representing the physical limits of the actual system. Furthermore, the signal is intentionally characterized by a certain degree of noise to ensure that the CNN can make predictions even in more challenging operating conditions than the standard ones. The graphical representation of the noise signal is shown in Fig. 1a. After completing the training process, the CNN has been validated by generating a variable frequency signal characterized by variations and adding noise to simulate realistic conditions. A deviation of the frequency profile has been introduced into the model that reflects realistic operating conditions, ensuring representative coverage of probable scenarios. In particular, a uniform noise profile has been chosen to simulate random noise, to represent uncertainty without introducing any polarization into the profile. After that, the signal is then filtered using a low-pass filter to attenuate the noise, and the result is represented in Fig. 1b. During the validation phase, the model predicts the subsequent frequency from the filtered signal segments, ensuring the predictions fall within a physically plausible range. The results, including forecasts, actual values, and mean absolute errors, are recorded and analyzed to evaluate the model's performance. In addition, explanatory graphs illustrate the evolution of the signal, the predictions made, and the MAE over time, allowing a visual assessment of the model's accuracy.

At the end of the compilation process, the following results emerge: the graph in Fig. 1c shows the prediction curve (in orange) superimposed on the actual signal (in dotted blue), initially in its entirety and then with an enlargement of the frequency descent phase from 50 Hz to 47.9 Hz, Fig. 1d. The analysis highlights how the predictions correctly follow the validation signal. A relevant aspect is the trend of the MAE as a function of the pre-



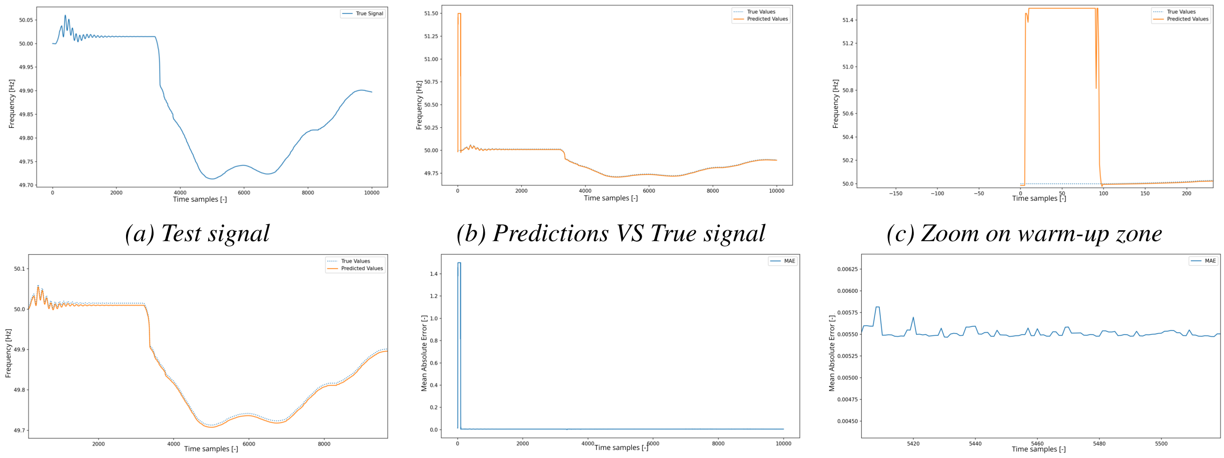
(a) True signal *(b) Predictions VS True Signal* *(c) Zoom on warm-up zone*

(d) Zoom after the warm-up zone *(e) MAE profile* *(f) Training and validation MAE*

Fig. 2. Ultra Short-term Frequency Prediction for a realistic operation scenario

dictions. In particular, by enlarging the area characterized by the highest error rate, Fig. 1f, it is observed that this corresponds to the first 100 points of the signal. This interval represents the warming up phase of the CNN, during which the model calibrates itself and learns the characteristics of the input signal. A synthetic analysis shows that the CNN presents significant errors in the first 100 points out of 10,000, corresponding to 1% of the entire signal. Once the calibration phase is completed, the CNN demonstrates a high predictive capacity, generating forecasts with a constant error close to zero, as highlighted in Fig. 1e. Indeed, after the first 100 forecasts, the maximum error found equals 0.02790. Once the CNN model has been trained and validated, it is essential to verify its prediction ability under more realistic conditions. To this end, the frequency evolution of a real system was analyzed. The compilation process took a significant amount of time due to the length of the signal, which includes 270,000 points, resulting in a processing time of approximately six hours. At the end of the compilation, the Fig. 2a and Fig. 2b were obtained. The warm-up phase occurs again in the first 100 points, as the enlarged Fig. 2c highlights. Beyond this interval, the CNN demonstrates a high predictive capacity, providing accurate estimates of the different frequency variations. Further experimental analysis have been developed by solving the same prediction problem applying the conventional approach used in neural network applications, which involves splitting the data into training, validation, and testing. At the end of each training epoch, the model is evaluated on the validation set to optimize hyperparameters and monitor performance, using metrics such as accuracy and F1 score. This approach prevents overfitting and improves the model ability to generalize the unseen data, thus ensuring reliable and accurate predictions. The input signal to the CNN consists of 80% of the training set, while a validation set distinct from the training set represents the remaining 20%. At the end of the compilation phase, the following results were obtained.

The Fig. 2f illustrates the evolution of the mean absolute error for the training and validation sets during the CNN training epochs. In the initial phases, the errors are high, but they quickly decrease, indicating that the model is effectively learning the characteristics of the signal. The convergence of training and validation errors around 0.1 suggests a good model generalization capability, with no apparent signs of overfitting. However, regular oscillations are observed, especially in the final stages of training, which could arise from fluctuations in the data batches or the nature of the validation set itself. Overall, the model appears stable and efficient after about 200 epochs. After the CNN training and validation phase, it is essential to verify its predictive capabilities under more realistic conditions. For this purpose, the frequency evolution of a real system was recovered. Since the system under consideration includes a high number of samples (270,000), to optimize the design time, it was decided to perform the predictions on the first 10,000 samples (Fig. 3a). The Fig. 3b compares the predicted signal with the actual signal, clearly analyzing the model performance. During the initial warm-up phase, which extends over the first 100 points, the model presents a higher error rate due to the calibration and adaptation process to the input data. However, after this phase, the predictions align closely with the actual signal, demonstrating the model effectiveness in capturing the underlying patterns and variations present in the data. Furthermore, a constant offset in the predictions is observed, suggesting a possible systematic bias. This deviation can be manually corrected to further improve the model accuracy.



(a) Test signal *(b) Predictions VS True signal* *(c) Zoom on warm-up zone*

(d) Zoom after warm-up zone *(e) MAE profile* *(f) Zoom after warm-up zone*

Fig. 3. Ultra Short-term Frequency Prediction for a realistic case study: Second Approach

B. *Result Discussion*

The analysis of the obtained results confirmed that after an initial adaptation phase, the CNN model exhibits a high accuracy in predicting the actual signal. The strong alignment between the predicted and the actual signals in the post-warm-up phase highlights the model ability to effectively learn from the input data and generalize the acquired information. The average error observed after the warm-up phase is tiny, equal to 0.0054044, suggesting a potential margin for further optimization of the model accuracy. The model demonstrates robustness and reliable performance, confirming its effectiveness in analyzing and predicting frequency variations in a realistic application context. It should be pointed out that the choice of using CNN is related to their high capacity to capture patterns from the signal, unlike recurrent neural networks, which are also used in frequency estimation and allow better understanding of the time dynamics of the signal. There are also hybrid approaches that tend to improve the extraction capability of both spatial and temporal dynamics.

C. *Short-term frequency predictions*

Further experimental analysis have been performed to assess the performance of the CNN-based model for frequency prediction over longer time frames. For this purpose, we chose to limit the predictions to frequency $k+100$ since this interval already allows us to identify a significant trend. Indeed, the prediction error tends to increase as the temporal distance of the predicted frequencies increases. For a realistic For this analysis, the mean squared error (MSE) on the test signal is calculated using two distinct approaches:

1. **The Persistence Method:** This method relies on a simple temporal shift to estimate future frequency values. In particular, the current value of the signal is used as a prediction for a future time step. The implementation involves shifting the test signal by T steps and then calculating the MSE between the original and the shifted signal. This calculation is performed for different values of T, ranging from 1 to 100.
2. **CNN predictions:** To evaluate the CNN ability to predict frequencies up to $k+100$, a pre-trained model is used. For each time step k, it generates frequency predictions for the following k steps, with k varying from 1 to 100.

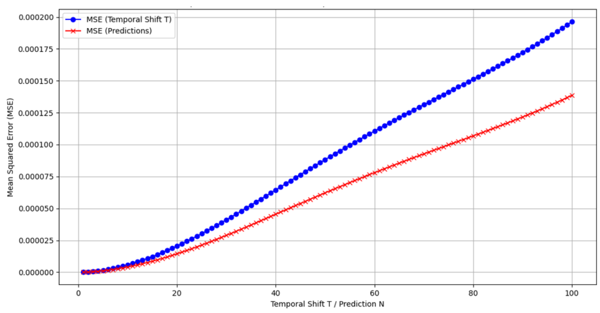


Fig. 4. Comparison of MSE based on temporal shift T and MSE for predictions

The results obtained using the two approaches are reported in Fig. 4. Specifically, the blue curve represents the MSE trend as a function of the temporal shift T for the persistence method, while the red curve illustrates the MSE associated with the CNN predictions up to $N+100$. The graph analysis shows that the MSE of the CNN predictions is lower than that obtained with the persistence method, suggesting that the neural network can capture signal trends more accurately than a simple model based on the temporal shift. This comparison confirms that, although the error increases with the temporal distance of the predictions, the CNN model provides more accurate results than the persistence method.

IV. CONCLUDING REMARKS

The planned CNN has proved highly effective in predicting frequency changes in power systems. Indeed, the results show that the model can accurately estimate frequency in a variable time scenario. Validation tests performed on several signals have further confirmed the accuracy of the CNN, providing results in line with theoretical expectations. In particular, the approach based on splitting the data into training, validation, and test sets has proven particularly effective, enabling the model optimization by hyperparameter tuning. This resulted in improved generalization ability and predictive accuracy. Overall, from the analysis carried out, the implementation of identification models based on CNN could represent a reliable tool for the estimation of inertia in modern energy systems, providing a strategic tool in the context of increasing integration of renewable energies. The model's ability to predict frequency variations with high precision could help improve situational awareness, which is essential for improving the resilience of the grid to dynamic disturbances. Future developments will focus on implementing and comparing recursive and hybrid neural networks, which combine convolutional and recursive. The goal is to exploit the strengths of both architectures: the ability of CNNs to identify patterns in the data, and the ability of recursive networks to extract information, particularly in the case of dynamic series, to improve predictive accuracy.

V. REFERENCES

[1] I. Jayawardene and G. K. Venayagamoorthy, "Cellular computational extreme learning machine network based frequency predictions in a power system," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 3377–3384.

[2] Y. Zhang, X. Wang, and L. Ding, "Lstm-based dynamic frequency prediction," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.

[3] N. Tomin, "Intelligent monitoring and forecasting of the expected operating conditions of electric power system," in *Proceedings of the 2011 3rd International Youth Conference on Energetics (IYCE)*. IEEE, 2011, pp. 1–8.

[4] M. P. Garcia and D. S. Kirschen, "Forecasting system imbalance volumes in competitive electricity markets," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 240–248, 2006.

[5] O. Yurdakul, F. Eser, F. Sivrikaya, and S. Albayrak, "Very short-term power system frequency forecasting," *IEEE Access*, vol. 8, pp. 141 234–141 245, 2020.

[6] J. Dong, X. Ma, S. M. Djouadi, H. Li, and Y. Liu, "Frequency prediction of power systems in fnet based on state-space approach and uncertain basis functions," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2602–2612, 2014.

[7] V. Negri, A. Mingotti, R. Tinarelli, and L. Peretto, "Comparison between the machine learning and the statistical approach to the forecasting of voltage, current, and frequency," in *2023 IEEE 13th International Workshop on Applied Measurements for Power Systems (AMPS)*. IEEE, 2023, pp. 01–06.

[8] A. Routray, A. K. Pradhan, and K. P. Rao, "A novel kalman filter for frequency estimation of distorted signals in power systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 3, pp. 469–479, 2002.

[9] V. V. Terzija, M. B. Djuric, and B. D. Kovacevic, "Voltage phasor and local system frequency estimation using newton type algorithm," *IEEE Transactions on Power Delivery*, vol. 9, no. 3, pp. 1368–1374, 1994.

[10] T. Tyagi and P. Sumathi, "Comprehensive performance evaluation of computationally efficient discrete fourier transforms for frequency estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 2155–2163, 2019.

[11] A. Pradhan, A. Routray, and A. Basak, "Power system frequency estimation using least mean square technique," *IEEE transactions on power delivery*, vol. 20, no. 3, pp. 1812–1816, 2005.

[12] L. Lai, W. Chan, C. Tse, and A. So, "Real-time frequency and harmonic evaluation using artificial neural networks," *IEEE Transactions on power delivery*, vol. 14, no. 1, pp. 52–59, 1999.

[13] D. V. Coury, A. C. Delbem, J. R. De Carvalho, M. Oleskovicz, E. V. Simoes, D. Barbosa, and T. V. da Silva, "Frequency estimation using a genetic algorithm with regularization implemented in fpgas," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1353–1361, 2012.

[14] Q. Bo, X. Wang, and K. Liu, "Minimum frequency prediction of power system after disturbance based on the v-support vector regression," in *2014 International Conference on Power System Technology*. IEEE, 2014, pp. 614–619.