Vol. 05, No. S1 (2023) 86-92, doi: 10.24874/PES.SI.01.011

Proceedings on Engineering Sciences

www.pesjournal.net

MACHINE LEARNING BASED GRID SAFETY ASSESSMENT THROUGH SIMULATION OF UNEXPECTED CONTINGENCIES DURING MAINTENANCE

Kamalraj R.[1]
Vikas Verma
Manish Joshi

Received 28.04.2023.
Accepted 29.06.2023.

Keywords:

Renewable energy, machine learning (ML), lightning search optimised random forest (LSORF), contingency.

A B S T R A C T

Effective maintenance coordination has become essential to ensuring a reliable electricity supply in power systems primarily powered by renewable sources. On the other hand, the computational complexity of the operational security standards presents difficulties for the existing planning tools. To solve this problem, a research paper suggests applying the machine learning (ML) method known as lightning search optimised random forest (LSORF) to anticipate the results of contingency analyses rapidly and effectively. The entire regional transmission system of Belgium (BE), which includes voltage ranges of 200 kV to 50 kV, is the subject of the study. Results show that LSORF regularly outperforms other benchmarks. The results demonstrate that LSORF consistently outperforms other benchmark methods. Furthermore, the study highlights the impact of projected growth in renewable energy on maintenance feasibility. This strategy provides useful insights for improving maintenance planning in renewable energy systems.

1. INTRODUCTION

Ensuring the security and reliability of power supply is a paramount concern in modern power systems. Regular maintenance activities are crucial in ensuring the grid infrastructure's smooth functioning (Peyghami et al. (2019)). Maintenance activities are essential for ensuring power systems' reliability and smooth operation. However, unexpected contingencies during scheduled maintenance pose significant challenges to system security (Wu et al. (2019)). These contingencies can range from equipment failures and unplanned outages to unpredictable changes in demand or the intermittent nature of renewable energy sources. When such contingencies arise, they can jeopardize the ability of the grid to safely accommodate these unexpected events (Duman et al. (2023)).

The management of unexpected contingencies during maintenance requires careful planning and consideration. Maintaining operational security standards and ensuring the grid can swiftly and effectively respond to unforeseen events (Dudurych (2021)). However, current planning tools used for maintenance coordination in power systems often struggle with tractability issues when incorporating operational security standards. Considering the impact of

[1] Corresponding author: Kamalraj R.
Email: r.kamalraj@jainuniversity.ac.in

unexpected contingencies significantly increases the computational burden. To ensure the safety and reliability of the grid, planners must simulate numerous scenarios to evaluate the system's ability to withstand contingencies during maintenance. This requires extensive computational resources and poses challenges in time and efficiency (Medina et al. (2022)).

One of the key challenges in maintenance planning for renewable-dominated power systems lies in addressing unexpected contingencies. These contingencies can arise from various factors, such as extreme weather events, equipment failures, or grid disturbances. When a contingency occurs during a scheduled maintenance period, it can disrupt the system's normal functioning and lead to potential reliability and security issues. Therefore, ensuring that the grid can safely accommodate any unexpected contingencies that may arise during maintenance activities becomes crucial.

Innovative approaches are being explored to alleviate these computational burdens and enhance maintenance coordination. Using ML models to anticipate contingency analysis results quickly and accurately is one relevant system. By incorporating ML into maintenance planning, it is possible to expedite the decision-making process and streamline the identification of suitable maintenance periods. In this context, we provide a methodology for predicting the effects of unanticipated events during maintenance operations in power networks with a high proportion of renewable energy sources. The process is tested especially on BE entire regional transmission grid, which covers a wide variety of voltage ranges from 200 kV to 50 kV. The goal is to create a quick and accurate model to pinpoint when maintenance can be conducted securely.

Overall, this research aims to contribute to developing efficient and reliable maintenance coordination strategies for renewable-dominated power systems. By leveraging ML and understanding the implications of unexpected contingencies during maintenance, we can optimize the planning process and improve the overall security of power supply in these evolving energy landscapes.

The remainder of this paper is as follows: part 2 describes the related works, part 3 explains the methodology, part 4 discusses the result of our proposed method, and Part 5 concludes the paper.

2. RELATED WORKS

For effective operation and maintenance of the power grid in an uncertain environment, a Reinforcement Learning framework was proposed by Bellani et al. (2019). They demonstrated a method utilising an ensemble of "Artificial Neural Networks" and the "Q-learning algorithm" that can be advantageous for large systems with massive state-action fields. The proposed technique yields solutions as precise as the true optimal, and an analytic (Bellman's) one was offered for the miniaturised power grid. It gives the system operator helpful guidance despite the fact that approximation errors are unavoidable and the processing time is still an issue of dispute.

Nakabi and Toivanen (2021) evaluate deep reinforcement learning methods for microgrid energy management. Priority assets, immediate demand signals for control, and power prices integrate flexible sources in the suggested energy management system. They evaluated seven deep reinforcement learning methods. Numerical studies reveal that "deep reinforcement learning algorithms" varied greatly in their capacity to converge to ideal rules.

The dual-stream CNN approach was used by Tian et al. (2022). It receives the voltage of each of the nodes and lines as inputs and outputs the key eigenvalue. It quickly recognises the primary oscillation patterns of the power system (PS) and provides a qualitative evaluation. The dual-stream CNN method enables dispatching procedures and improves PS safety and stability.

Using "reinforcement learning and a deep neural network," Lu and Hong (2019) proposed a novel incentive-based real-time demand-side response method for smart grid systems, with the final goal of assisting the provider of services in purchasing renewable energy sources via its subscribed consumers with the goal to equalised energy fluctuations and improve grid reliability.

Menke et al. (2018) used artificial neural networks for contingency analysis in order to improve accuracy and forecast more PS characteristics. Twenty percent of the AC power flow obtained from an entire year of time series simulation is used to train deep feedforward network topologies. Next, it generates predictions about the remainder of line loadings and bus voltages.

Varbella et al. (2022) developed a data-driven approach for online cascading failure risk computation. They train "Feedforward Neural Networks (FNN) and Graph Neural Networks (GNN)" on synthetic data. GNNs can generalised to graphs of various sizes and improve graph-structured data performance. FNN and GNN are compared, and test grids indicate GNN's inductive capabilities. Transfer learning improves GNN model performance on power grids not used during training. The GNN model can detect if several failures produce a critical grid state under defined grid operating parameters.

Gargari et al. (2021) suggested a "sequential maintenance scheduling" of a multi-energy micro grid to improve system resilience. To achieve accurate and reliable results, the "maintenance scheduling problem" accounts for energy carrier interactions. Three alternative schedule scenarios are presented for each failure. Three instances show the approach's effectiveness in traditional multi-energy micro grids.

Yang et al. (2021) analysed a PS cascading vulnerability using higher security standards and renewable energy integration. To reflect cascade propagation, they employ a “graph-based thermal inertia-based cascades model with an N-k contingency sampling approach ."The method shows cascade propagation and helps visualise and analyse system vulnerability.
Dorile et al. (2021) aimed to help transmission network planners and operators integrate “large-scale wind farms into the transmission grid" with substantial wind power penetration. “Q-V, P-V, and N-1 contingencies with Remedial Action Schemes (RAS)” are used to examine wind-dominated PS stability. The most serious situations and voltage collapse during the maximum wind penetration level are ranked and predicted. The transmission system operator can utilise the results to predict PS instability in voltage or collapse during high wind penetration.

3. METHODS

3.1 Dataset creation

We created a database $Y^{orig} \in R^{|E|\times|V|}$ that contains data on variables $d \in E$ gathered on an hourly basis $t \in V$. Both international and regional information can be found in the $|E|$ variables. The overall load is available to us at the system scale. In addition, we rely on the aggregate production from several technologies. Table 1 shows the generational distribution (in 2020)

Table 1. Generational distribution in 2020.

Generation mix	Percentage (%)
Biogas	0%
onshore	5%
nuclear	39.3%
Wind	5.1%
coal	-0.6%
others	5.7%
offshore	8.5%
Photovoltaic	2.8%
gas	34.1%

Due to transparency requirements, Elia provides all of these market data analyses accessible through the website. This work also makes use of non-public weather data at a single site in the geographic middle of BE. Measurements of generation and consumption at the various transmission system nodes are the local variables. The dimensionality of the problem is further increased by the fact that there are, in practise, more than a thousand locations on the BE transmission grid. Confidential information about Elia's nodal energy exchanges. It is also important to note that Elia provides the structure of the BE grid in addition to all of its fundamental assets, making it possible to conduct useful research on the BE transmission system in the context of the study.

ML models are often trained with samples. Market simulations, on the other hand, can construct these input possibilities by simulating a variety of novel grid circumstances that haven't been included in the historical database. One can build a model that can correctly generalise to unanticipated scenarios in the future by replicating the present state of the market. Additionally, there are no gaps in the historical data because it is required by statute that the market data be correctly evaluated (due to the repercussions for the economy). Therefore, in this study, there was no need to implement a strategy for data aggregation. The many grid assets that require maintenance are listed as an addition to this database.

The objective is to determine whether it is feasible to maintain the database for each system state ($t \in V$) for each asset ($d \in D$). This is accomplished through a "quasi N-2" contingency analysis, simulating all pertinent unexpected problems occurring concurrently with the scheduled maintenance. The following four conditions must all be met in order for asset d to be maintained ($y_{d,t} = 1$).

Several factors must be taken into account in the grid maintenance strategy to guarantee the system's dependability and stability. The overloading criterion is crucial as it aims to prevent congestion and cascading effects that can lead to severe negative impacts such as load loss. The voltage criterion is equally important to maintain power quality standards and ensure that the connection points of grid users stay within acceptable voltage limits. Additionally, the load at risk criterion plays a significant role in preventing excessive power loss following a contingency event, whereas the resources at risk criterion restrict the amount of energy that cannot be transmitted to grid users. These criteria collectively guide the decision-making process and help in maintaining a robust and reliable power system.

Maintenance cannot be performed ($y_{d,t} = 0$) if any of these requirements are broken. The viability of a maintenance operation can be defined using a multicriteria approach that incorporates both economic and reliability perspectives.

3.2 Selecting suitable input variables

Principal component analysis (PCA) can be used to reduce the dimensionality of the training data. This process involves normalising the output data as well as producing the training data covariance matrix, eigenvalues, and eigenvectors. For a specific set z of training dataset input data, the covariance matrix can be defined as follows.

$$E = \frac{1}{N}\sum_{k=1}^{N} \left(z_k - \underline{z}\right)\left(z_k - \underline{z}\right)^{V} \quad (1)$$

Where N is the entire number of samples, z_k is each data in the training set, and z_k is the average of the samples.

$$\underline{z} = \frac{1}{N}\sum_{k=1}^{N} z_k \quad (2)$$

The eigenvalues and eigenvector pairs that were calculated are subsequently ordered from lowest to highest. The PCA subspace is then encompassed by the L biggest eigenvectors $V = [\underline{v_1}, \underline{v_2}, \dots, \underline{v_l},]$ As an instance of the process used to transform high-dimensional data into the low-dimensional PCA subspace, examine the steps that follow:

$$a = W^v z \; (3)$$

The average of every set of data is then determined by using the formulae below:

$$\underline{w_n} = \frac{1}{N} \sum_{a \epsilon \omega_k} \quad A \tag{4}$$

Where N_k is the total number of data used for training samples, and ω_k and a are the coordinates of the simulated data. Similarly, the average of all shown data can be found by using the formula:

$$\underline{z} = \frac{1}{N_k} \sum_{for\ all\ a} \quad A \tag{5}$$

Where N the total number of data is points, and a represents the set of standard data used for training.

3.3 Random forest

In describing RF, the term "contingency analysis" refers to the practise of utilising the algorithm to examine the association between categorical variables. Because of its flexibility, RF can be used for both quantitative and qualitative analyses of contingencies. An RF is a collection of decision trees, each of which has been trained using a different sampling of the data and the characteristics. Overfitting is mitigated, and the model's generalisation performance is enhanced by this randomness.

Step 1-Bootstrap Sampling: To generate a bootstrap sample, a fraction of the original data is selected at random and then replaced. Bootstrapping is a method for creating various groups of data used in the training of each tree.

Step 2-Feature Randomization: Create a feature set by picking features at random. Standard practise is that a square root of the total number of features is used to determine how many features to select at each node. Overfitting is avoided with the help of this random selection's introduction of variation.

Step 3-Construct Decision Trees: Select features to incorporate to construct a decision tree using the bootstrap instance. The decision tree is constructed by dividing the data into subsets using the features that were chosen in a recursive manner. When deciding where to make the splits, we use metrics like Gini impurity and information gain to ensure the cleanest possible data at each stage.

Step 4-Repeat for Forest: Make a forest of decision trees by repeating steps 1 through 3. The user can specify the hyperparameter "number of trees in the forest."

Step 5-Prediction: According to the input features, each tree in the forest independently categorises or forecasts the target variable. For classification problems, the class with the greatest number of votes across all trees is chosen as the final forecast, which is decided by majority voting. The average of the projected values from all the trees often serves as the final prediction for regression tasks.

The Random Forest algorithm's equation appears as follows:

$$h(x) = \Sigma \left(\frac{1}{N}\right) * h_i(x)$$

- $h(x)$ - predicted output for a given input x.
- Σ - sum over all trees in the forest.
- N - total number of trees in the forest.

$h_j(x)$ - predicted output of the j^{th}tree for input x.

3.4 Lightning Search Optimization (LSO)

A modern metaheuristic algorithm called the Lightning Search optimization (LSO) algorithm draws inspiration from lightning as a natural occurrence. In order to construct a binary tree-like structure resembling a step leader, it makes use of an assortment of rapid particles called projectiles that move around the search space. Transition projectiles (TP), space projectiles (SP), and lead projectiles (LP) are the three types of projectiles used in LSO.

TP: The initial group of step leaders is comprised of these projectiles. They are produced using arbitrary numbers taken from the regular standard distribution for probabilities. These projectiles explore the search space in an initial exploration phase.

SP: These projectiles are updated and evolved over time, aiming to find the optimal solution. Through iterative steps, these projectiles move in the search space, exploring different regions and refining their positions. The goal is to converge toward a high-quality solution by gradually improving the positions of these projectiles.

LP: The lead projectile represents the best solution found so far. It keeps track of the best objective value obtained during the search process. The lead projectile guides the movement of the other projectiles, influencing their exploration and exploitation of the search space.

LSO conducts a search procedure that resembles the propagation of a step leader in lightning by using a combination of TP, SP, and LP. The algorithm seeks to efficiently investigate the search space, converge to the best solution, and iteratively modify the LP to reflect the best option identified. The below pseudocode describe the process of LSO.

Pseudocode 1: Process of LSO

Initialize the TP with random positions in the search space
Evaluate the objective function for each TP
Set the LP as the projectile with the best objective value
Repeat until a termination condition is met:
Update the SP based on the position of the LP
Evaluate the objective function for each SP
Update the LP if an SP has a better objective value
Generate new TP using arbitrary numbers drawn from the standard distribution for probabilities
Evaluate the objective function for each new TP
Update the lead projectile if a new TP has a better objective value
Choose a leader among the SP
Update the positions of the SP based on the LP
Evaluate the objective function for each updated SP
Update the LP if an updated SP has a better objective value
End Repeat
Return the LP as the best solution found.

4. RESULT AND DISCUSSION

Utilising actual data from BE, the suggested methodology is used. The R programming language, which is open-source, has been used to implement all of the categorization tools. The "Power Factory" programme was used to simulate the network and perform contingency analysis, or load-flow calculations that estimate the state of the PS in various contingency (grid asset outage) situations.

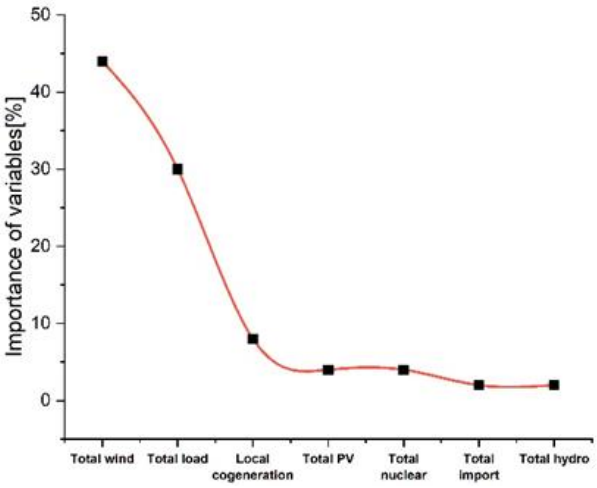

Figure 1. Maintenance feasibility ranking of key factors

The main high-voltage line for transmission of the BE system is the focus of our initial investigation on maintainability. To that purpose, by training an RF on the entire database, we first choose the most significant explanatory factors (EF) to anticipate the probability of maintenance tasks under various scenarios. The significance of the trained model's variables can be measured. As a reminder, the most important factors are those that achieve a significance level higher than an entirely random variable. Figure 1 depicts the ranking of key factors. In particular, we find the state of the global grid greatly influences the maintainability of the investigated asset. Since the investigated asset is a critical part of the transmission system, it stands to reason that it would be impacted most by large-scale power trades, as shown by the results. These findings show that constructing the line's maintenance plan is difficult since the important EF cannot be reliably anticipated over long periods.

Our proposed model can be trained on the basis of the chosen attributes. The number of decision trees used to create an average and lower the model's variance is a crucial hyper-parameter in RF. The training duration of the final model increases linearly with the number of trees. Thus it's important to strike an appropriate equilibrium between the two. Figure 2 shows the out-of-bag error of the RF, obtained by providing each tree with instances that were not used in the learning technique and combining the classification error.

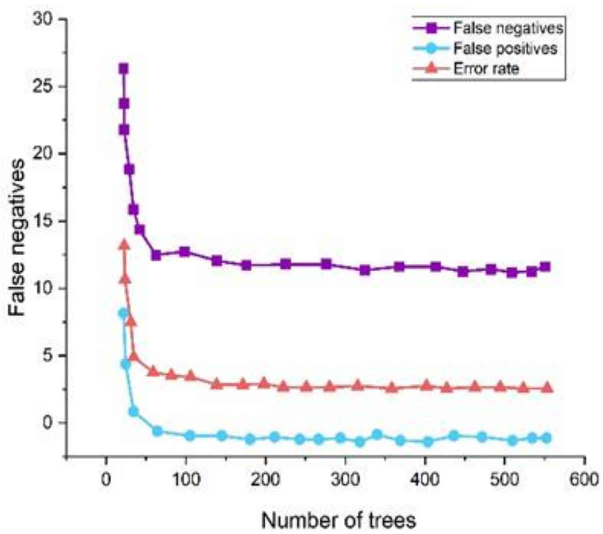

Figure 2. Out of bag result

In order to evaluate a classifier's efficacy in a binary classification task, a graph called an Area under the Curve (AUC) graph is typically employed. The “true positive rate (TPR) and false positive rate (FPR)” at different categorization levels are discussed. TPR is determined by dividing the number of correctly categorised positive instances by the total number of actual positive instances for a certain threshold. FPR can be determined by comparing the number of false negatives to the true negatives. The results for the AUC curve are shown in Figure 3. AUC of 0.98.5 is reached when using RF with all variables. However, training is made much more difficult by the requirement to deal with the enormous dimensionality of the input space, which is greater than 1000. In addition, this approach limits the identification of crucial characteristics, which is crucial data for specialised organisations to possess.

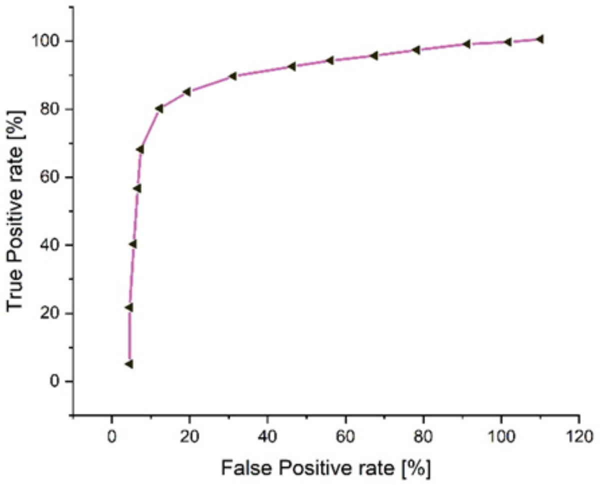

Figure 3. Outcome of AUC curve

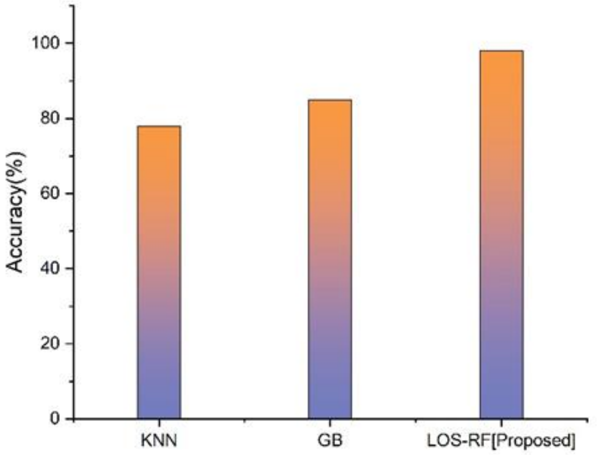

Figure 4. Accuracy outcome

When performing classification tasks, accuracy is a frequently used metric that assesses how accurate a classifier's predictions are overall. It calculates the percentage of correctly identified examples among all the instances in a dataset. The level of accuracy is apparent in Figure 4. It demonstrates that, when compared to the current methods (KNN, GB), our suggested method (LSO-RF) is effective.

5. CONCLUSION

This research paper proposes the application of the lightning search-optimized random forest (LSO-RF) algorithm for rapidly and effectively predicting the outcomes of contingency analyses in the entire regional transmission system of BE. The LSORF bridges the gap among ML and field experience with its ease of use, interpretability and outstanding performance. The results demonstrate that LSO-RF consistently outperforms other benchmark methods. Furthermore, the study highlights the impact of projected growth in renewable energy on maintenance feasibility. These findings offer valuable insights for enhancing maintenance planning in renewable energy systems. By leveraging LSO-RF, decision-makers and operators can make more informed and efficient maintenance decisions, thus optimising renewable energy integration and transmission system reliability. Additionally, incorporating these models into interruption scheduling instruments is a crucial next step in raising the grid maintenance's cost-effectiveness.

References:

Sar, K. and Ghadimi, P., 2023. A Systematic Literature Review of the Vehicle Routing Problem in Reverse Logistics Operations. Computers & Industrial Engineering, p.109011.

Peyghami, S., Davari, P., Fotuhi-Firuzabad, M. and Blaabjerg, F., 2019. Standard test systems for modern power system analysis: An overview. IEEE Industrial Electronics Magazine, 13(4), pp.86-105.

Wu, Y.K., Chang, S.M. and Mandal, P., 2019. Grid-connected wind power plants: A survey on the integration requirements in modern grid codes. IEEE Transactions on Industry Applications, 55(6), pp.5584-5593.

Duman, S., Kahraman, H.T. and Kati, M., 2023. Economical operation of modern power grids incorporating uncertainties of renewable energy sources and load demand using the adaptive fitness-distance balance-based stochastic fractal search algorithm. Engineering Applications of Artificial Intelligence, 117, p.105501.

Dudurych, I.M., 2021. The impact of renewables on operational security: Operating power systems that have extremely high penetrations of nonsynchronous renewable sources. IEEE Power and Energy Magazine, 19(2), pp.37-45.

Medina, C., Ana, C.R.M. and González, G., 2022. Transmission Grids to Foster High Penetration of Large-Scale Variable Renewable Energy Sources–A Review of Challenges, Problems, and Solutions. International Journal of Renewable Energy Research (IJRER), 12(1), pp.146-169.

Rocchetta, R., Bellani, L., Compare, M., Zio, E. and Patelli, E., 2019. A reinforcement learning framework for optimal operation and maintenance of power grids. Applied energy, 241, pp.291-301.

Nakabi, T.A. and Toivanen, P., 2021. Deep reinforcement learning for energy management in a microgrid with flexible demand. Sustainable Energy, Grids and Networks, 25, p.100413.

Tian, Z., Shao, Y., Sun, M., Zhang, Q., Ye, P. and Zhang, H., 2022. Dynamic stability analysis of power grid in high proportion new energy access scenario based on deep learning. Energy Reports, 8, pp.172-182.

Lu, R. and Hong, S.H., 2019. Incentive-based demand response for smart grid with reinforcement learning and deep neural network. Applied energy, 236, pp.937-949.

Schäfer, F., Menke, J.H. and Braun, M., 2018, October. Contingency analysis of power systems with artificial neural networks. In 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm) (pp. 1-6). IEEE.
Varbella, A., Gjorgiev, B. and Sansavini, G., 2023. Geometric deep learning for online prediction of cascading failures in power grids. Reliability Engineering & System Safety, 237, p.109341.
Gargari, M.Z., Hagh, M.T. and Zadeh, S.G., 2021. Preventive maintenance scheduling of multi energy microgrid to enhance the resiliency of system. Energy, 221, p.119782.
Yang, S., Chen, W., Zhang, X. and Yang, W., 2021. A graph-based method for vulnerability analysis of renewable energy integrated power systems to cascading failures. Reliability Engineering & System Safety, 207, p.107354.
Dorile, P.O., Jagessar, D.R. and McCann, R.A., 2021, November. Security of a Power System Under High Penetration of Wind Energy Considering Contingencies and Stability Margins. In 2021 IEEE 19th Student Conference on Research and Development (SCOReD) (pp. 450-455). IEEE.

R. Kamalraj
Jain (deemed to be) University,
Bangalore, India
r.kamalraj@jainuniversity.ac.in
ORCID 0000-0003-0489-6062

Vikas Verma
Vivekananda Global University,
Jaipur, India
vikash.verma@vgu.ac.in
ORCID 0000-0003-4139-4854

Manish Joshi
Teerthanker Mahaveer University,
Moradabad,
Uttar Pradesh, India
gothroughmanish@gmail.com
ORCID 0000-0002-3543-5915