Data-Driven Methods for Optimal Setting of Legacy Control Devices in Distribution Grids

Alper Savasci[⊕], Oguzhan Ceylan[◊], and Sumit Paudyal[†]
[⊕]Abdullah Gul University; [◊]Kadir Has University; [†]Florida International University
Emails: alper.savasci@agu.edu.tr, oguzhan.ceylan@khas.edu.tr, spaudyal@fiu.edu

***Abstract*—This study presents machine learning-based dispatch strategies for legacy voltage regulation devices, i.e., on-load tap changers (OLTCs), step-voltage regulators (SVRs), and switched-capacitors (SCs) in modern distribution networks. The proposed approach utilizes k-nearest neighbor (KNN), random forest (RF), and neural networks (NN) to map nodal net active and reactive injections to the optimal legacy controls and resulting voltage magnitudes. To implement these strategies, first, an efficient optimal power flow (OPF) is formulated as a mixed-integer linear program that obtains optimal decisions of tap positions for OLTCs, SVRs, and on/off status of SCs. Then, training and testing datasets are generated by solving the OPF model for daily horizons with 1-hr resolution for varying loading and photovoltaic (PV) generation profile. Case studies on the 33-node feeder demonstrate high-accuracy mapping between the input feature and the output vector, which is promising for integrated Volt/VAr control schemes.**

***Index Terms*—Distribution Grid, Optimal Power Flow, Voltage Control, Machine Learning.**

I. Introduction

Voltage regulation in distribution networks primarily relies on utility-owned regulation devices, commonly known as legacy devices, such as on-load tap changers (OLTCs), step voltage regulators (SVRs), and fixed/switched capacitors (SCs) [1]. Legacy devices typically operate through mechanical switching, involving on/off status in SCs or discrete tap positions in OLTCs and SVRs. The need for coordination among different types of legacy devices has been emphasized [2] and has a long history in distribution automation. Due to the integration of renewable-based distributed energy resources (DERs), i.e. photovoltaics (PV), and new load types such as electric vehicles induce fast dynamic voltage fluctuations along the distribution feeders, making voltage regulation a more challenging task.

In modern distribution systems, voltage regulation assets can be coordinated by an integrated Volt/VAR control (IVVC), which is a function of advanced distribution management systems (ADMS) [3]. The IVVC framework is typically interfaced with a telemetry system, collecting network data from strategically located sensors and sending controls to the field devices by directly commanding the remote regulation devices [4] or setting their controller parameters for autonomous operation [5] based on the voltage and asset control strategy. Among the conventional strategies, setting gradual time delays for multiple SVRs cascaded along the feeder or setting time clock switches for capacitor banks to connect and disconnect to the distribution system during predetermined times. In this context, [6] proposes a time-delay strategy coordinating the simultaneous operation of substation OLTC, feeder SVRs, and a synchronous machine-based unit for voltage support. In modern applications, the voltage regulation task is widely considered as an optimization problem in which the operation of the legacy devices is coordinated to optimize a certain network objective, such as minimizing the total system losses or total voltage violations. In [7], a mixed-integer second-order cone program (MISOCP)-based optimal power flow (OPF) is solved to dispatch the legacy devices as well as the power output of PV-based DERs to optimize the network-wide voltage profile. Since legacy devices have discrete controls, solving mixed-integer type OPF problems might be computationally costly for large-scale networks, even with a few dispatchable legacy devices. For example, solving a model-based OPF to dispatch 3 LTCs each of which has 33-tap for 15 minutes horizon with 1 minute resolution incurs 1485 discrete variables, which challenges the general mixed-integer linear programming solvers. Thus, model-based optimization naturally face computational challenges. Apart from model-based optimization, learning-based approaches are being employed to improve the OPF solution procedure as proposed in [8], [9], [10]. In [8], a neural network model is trained to solve AC-OPF cases with significant speedups. In [9], a deep neural network (DNN) is trained to predict all bus voltages; then, other variables are computed through power flow equations for the feasibility of power balance and bus voltages. The authors in [10] propose a DNN-based model incorporating topology switching, which has a discrete nature. In [11], a graph neural network model is proposed to approximate the AC-OPF solution and a dataset is generated by solving AC-OPF instances by varying loading conditions. In [12], a DNN-based AC-OPF solution approach is presented, in which the DNN model is trained with artificially created network loading samples and corresponding optimal generation set points and generator bus voltages. In [13], a neural network (NN) model is proposed to coordinate an OLTC transformer and a STATCOM, assuming both are placed at the substation. The NN model is utilized as an optimal tap position selector, making the STATCOM output minimum. A training dataset is created by solving AC power flow instances for each 33 OLTC tap positions under a given loading condition and taking the instance that results in minimum STATCOM output.

This study contributes to the existing works by showing the discrete controls in voltage regulation applications can

be learned via basic machine learning (ML) methods such as k-nearest neighbor (KNN), random forest (RF) and neural networks with high accuracy by training the data-driven methods using the optimal solutions of an efficient mixed-integer linear OPF model. The rest of this paper is organized as follows. Section II describes the mathematical models of the distribution network and major component models. Section III presents the proposed approach with ML methods. Section IV describes the test feeder setup and discusses the results based on the case studies. Section V concludes the paper.

II. Mathematical Models

A. Distribution Network and Power Flow Models

A radial distribution network can be represented as a directed graph, $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{E}$ is the set of branches. Notation-wise, the electrical quantities and variables associated with a node are indexed by a single subscript, while double subscripts are for the branch variables. A linear relationship between branch power flows and node voltage can be represented by *LinDistFlow* [14] model as,

$$p_j^{g,t} - p_j^{d,t} = \sum_{k:(j,k)\in\mathcal{E}} P_{jk}^t - \sum_{i:(i,j)\in\mathcal{E}} P_{ij}^t, \tag{1a}$$

$$q_j^{g,t} + q_j^{c,t} - q_j^{d,t} = \sum_{k:(j,k)\in\mathcal{E}} Q_{jk}^t - \sum_{i:(i,j)\in\mathcal{E}} Q_{ij}^t, \tag{1b}$$

$$V_j^t = V_i^t - \frac{r_{ij}P_{ij}^t + x_{ij}Q_{ij}^t}{V_{ref}^t}, \tag{1c}$$

where t denotes the time interval index, i, j are nodes on the feeder. The total active (reactive) power generation is denoted by $p_j^{g,t}$ ($q_j^{g,t}$), and the active (reactive) power demand is represented by $p_j^{d,t}$ ($q_j^{d,t}$). The sending-end active and reactive power flowing on the line (i, j) are denoted by P_{ij}^t and Q_{ij}^t, respectively. The voltage of node i at time interval t is denoted by V_i^t.

The mathematical relationship of the terminal voltages of an OLTCs and SVRs can be defined as,

$$V_j^t = V_i^t(1 + \delta V \cdot T_{ij}^t) \tag{2a}$$

where T_{ij} denotes the tap position, δV is the step voltage per tap change. Note that (2a) is a nonlinear equation as the product of voltage and tap variables, which can be replaced with a new variable $W_{ij}^t := V_i^t T_{ij}^t$. Then McCormick relaxation can be employed to linearize the nonlinear product term as,

$$\begin{aligned} W_{ij}^t &\geq V_i^l\, T_{ij}^t + V_i^t\, T_{ij}^l - V_i^l\, T_{ij}^l \\ W_{ij}^t &\geq V_i^u\, T_{ij}^t + V_i^t\, T_{ij}^u - V_i^u\, T_{ij}^u \\ W_{ij}^t &\leq V_i^u\, T_{ij}^t + V_i^t\, T_{ij}^l - V_i^u\, T_{ij}^l \\ W_{ij}^t &\leq V_i^l\, T_{ij}^t + V_i^t\, T_{ij}^u - V_i^l\, T_{ij}^u \end{aligned} \tag{2b}$$

where V_i^l, V_i^u and T_{ij}^l, T_{ij}^u denote the upper and lower limits of the corresponding variables. The OLTCs and SVRs can be collected by a set $\mathcal{E}_r = \{(i, j)|i, j \in \mathcal{N}\}$.

SCs are another type of legacy devices that can improve the feeder voltages and operation by injecting reactive power. SCs can be modeled in mixed-integer linear form as,

$$u_{j,t}\, Q_j^C = q_j^{c,t} \tag{3}$$

where Q_j^C represents VAr rating of each capacitor, whose switching operation is modeled by binary variable u_j. SCs can be collected by the set $\mathcal{N}_c = \{i|i \in \mathcal{N}\}$.

B. Operational Model of Distributed Generation

Fig. 1 shows a typical capability curve of a distributed generator, which defines the operational mathematical model with the corresponding power output limits.

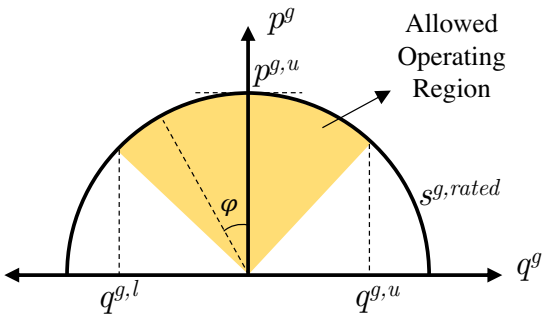

Fig. 1. Capability curve of a distributed generator

where $p_i^{\text{PV},t}$ is the produced PV power by the PV system, $p_i^{\text{curt},t}$ is the curtailed power by the PV inverter, active (p^g) and reactive (q^g) power outputs can be adjusted within the allowed operating region or with respect to a specified power factor angle (φ). The operating region can be defined by a set of constraints as,

$$p_i^{g,t} = p_i^{\text{PV},t} - p_i^{\text{curt},t} \tag{4a}$$

$$0 \leq p_i^{\text{curt},t} \leq p_i^{g,u} \tag{4b}$$

$$q_i^{g,l} \leq q_i^{g,t} \leq q_i^{g,u} \tag{4c}$$

$$-tan(\varphi_{min})\, p_i^{g,t} \leq q_i^{g,t} \leq p_i^{g,t}\, tan(\varphi_{max}) \tag{4d}$$

$$(p_i^{g,t})^2 + (q_i^{g,t})^2 \leq (s_i^{g,rated})^2 \tag{4e}$$

where $s_i^{g,rated}$ denotes the apparent power limit of the generating unit. To obtain a fully linearized model, the quadratic inequality (4e) can be linearly approximated, as proposed [15], by making use of a 32-vertex polygon (k = 16) and defining a polyhedral norm as,

$$-s_i^{g,rated} \leq \cos(h\gamma)\; p_i^{g,t} + \sin(h\gamma)\; q_i^{g,t} \leq s_i^{g,rated} \tag{4f}$$
$$\gamma = \frac{\pi}{k}, \quad h = 1, \ldots, k.$$

The index of DG units can be collected by the set $\mathcal{N}_g = \{i|i \in \mathcal{N}\}$. The operational model of a DG unit can be linearly modeled by (4a) - (4d) and (4f).

C. Optimal Power Flow as A Network Controller

The operation of a distribution system can be considered within a time horizon that is decomposed into multiple quasi-static time intervals, which are collected by the set

$\mathcal{T} = \{1, \ldots, |\mathcal{T}|\}$. During each interval, all state and control variables are assumed constant.

A specific OPF model can be devised and then solved periodically to dispatch the controllable network equipment with optimal set points such that a certain objective function is minimized. It is important that an OPF has tractable formulation, that is, efficiently solvable within a reasonable time and computational resources. In this study, such an OPF formulation given compactly as in (5)

$$\underset{\substack{p_i^g, q_i^g, P_{ij}, Q_{ij} \\ V_i, T_{ij}^t, W_{ij}^t}}{\text{argmin}} \quad J \tag{5a}$$

$$\text{Subject to :}$$

$$\text{Power balance : (1a)} - \text{(1b)}, \qquad \forall j \in \mathcal{N}, \forall t \in \mathcal{T} \tag{5b}$$

$$\text{Voltage drop : (1c)}, \qquad \forall (i,j) \in \mathcal{E} \setminus \mathcal{E}_r, \forall t \in \mathcal{T} \tag{5c}$$

$$\text{LTC model : (2a)} - \text{(2b)}, \qquad \forall (i,j) \in \mathcal{E}_r, \forall t \in \mathcal{T} \tag{5d}$$

$$\text{Capacitor model : (3)}, \qquad \forall j \in \mathcal{N}_c, \forall t \in \mathcal{T} \tag{5e}$$

$$\text{DG model : (4a)} - \text{(4d)}, \text{(4f)} \qquad \forall i \in \mathcal{N}_g, \forall t \in \mathcal{T} \tag{5f}$$

$$p_j^{g,t} \in [p_j^{g,l}, p_j^{g,u}], \qquad \forall j \in \mathcal{N}' \setminus \mathcal{N}_g, \forall t \in \mathcal{T} \tag{5g}$$

$$q_j^{g,t} \in [q_j^{g,l}, q_j^{g,u}], \qquad \forall j \in \mathcal{N}' \setminus \mathcal{N}_g, \forall t \in \mathcal{T} \tag{5h}$$

$$T_{ij}^t \in Z,\ T_{ij}^t \in [T_{ij}^l, T_{ij}^u], \qquad \forall (i,j) \in \mathcal{E}_r, \forall t \in \mathcal{T} \tag{5i}$$

where l and u denote the lower and upper limit of the variables, $\mathcal{N}'$ is the set of nonslack nodes. J denotes the objective function. We adopt an objective function as in (5j), where ω_1 and ω_2 denote the weights on the individual objective term.

$$J = \omega_1 \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}} p_i^{curt,t} + \omega_2 \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}} \max(0, V_i^l - V_i^t) + \max(0, V_i^t - V_i^u). \tag{5j}$$

III. Proposed Approach

Fig. 2 shows the input and output relations of a general framework that we follow in each ML method. For NN, we selected Bayesian Regularization with 72 inputs, 5 hidden and 41 outputs layers. We used training rate as %70, test rate and validation rate are selected as %15. In KNN-based approach, we selected K as 5. In RF model, sample size is assumed as 100. In both KNN, and NN our models are multi-input, multi output models. In RF-based approach, each single output was calculated independently.

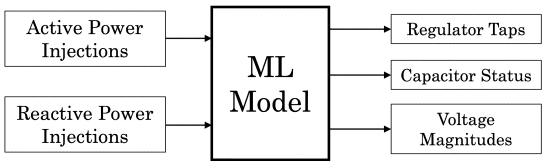

Fig. 2. ML framework.

A. K-nearest neighbors

We first use a non-parametric supervised learning method, K-nearest neighbors (KNN) [16]. The method aims to find the relation between the input and outputs of a given dataset. In other words, the output variables may be determined using the historical values of the inputs and outputs. For this a function $h : X \rightarrow Y$ is used to find the output value of $h(x)$ where, X represents the inputs, and Y represents the outputs. Since the approach is based on the historical values, it uses similarity measures such as Euclidean, Mannhattan, and Chebyshev based on proximity between data points for making predictions or classifications. The distances are computed for all data points in the entire dataset. The K nearest points to the desired output value are identified, and in classification problems, the new point is assigned to the class with the majority of nearby values among these K nearest points. In the case of value estimation, if the algorithm predicts a numerical value, the initial approach often involves taking the arithmetic mean of the values associated with the K nearest points.

B. Neural Networks

The neural networks excel particularly when there are non-linear relationships between inputs and outputs by simulating the emulation of biological neural networks [17]. Neurons simulating biological counterparts receive signals from neighboring neurons and can process them using predefined simple functions. Based on the outcome of this process, a neuron can generate an output signal. If the artificial neurons in question can take real-number values, the output value can be any number between 0 and 1; otherwise, it is either 0 or 1. The function that computes the output using M-dimensional input values has two main components: one processes input values, and the other transfers input values to output values nonlinearly. Then, this value is inputted into the transfer function known as the activation function [17], which takes the input to a neuron and produces the output.

C. Random Forests

Random forest structures classification trees based on numerous randomly distributed vectors. Breiman introduced the method [18], who defined the Random Forest as a collection of independently distributed random vectors comprising a structure of classifiers in a tree format. The vector representing the new data is placed in each tree within the forest to classify new data. Each tree produces a classification result. In summary, the results produced by the decision trees created in the forest are examined, and it is assumed that the new data belongs to the class that yields the most frequent results.

IV. Numerical Results

A. Test Network

To test the proposed OPF formulation and generate a dataset for training and testing purposes, the well-known 33-node feeder is modified as shown in Fig. 3, by placing two SC units and SVRs and an OLTC at the substation bus. The network line parameters and base loading are used as provided in the reference [19]. In addition, identical 5 utility-scale PV units, each of which is rated as 550 kVA with $|\cos^{-1}\varphi| \leq 0.95$, and two 500 kVAr capacitor units are connected to the feeder.

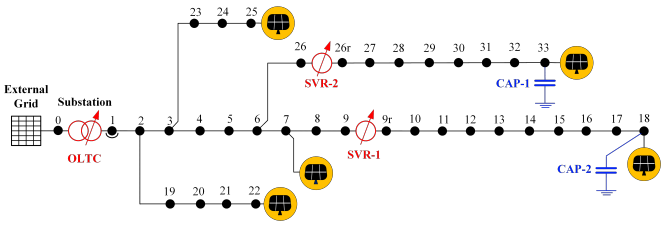

Fig. 3. Modified 33-node test network.

B. Dataset Generation Procedure

Fig. 4 illustrates the data generation procedure. The devised OPF model is solved for each day by using annual loading and PV capacity profiles. Then, input and output data vectors are obtained for training and testing stages for the employed ML methods. The input vector is comprised of the net active and reactive power injections from each node. The output vector consists of the node voltages, the tap positions of OLTCs and SVRs, and the status of SCs.

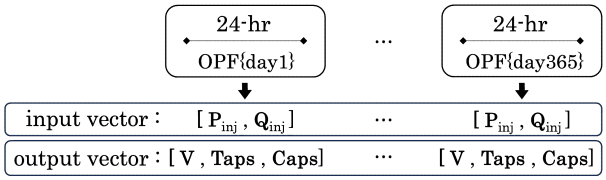

Fig. 4. Dataset generation procedure.

Each OPF case is solved by using Gurobi solver with default optimality settings. Hence, the solutions are considered optimal.

C. Simulation Results

We performed all the trainings using 300 days data, after creating the models we inputted the remaining 65 days. The simulations are performed using Matlab on a laptop with Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz.

Figures 5, 6, 7 show the differences between the actual voltage magnitudes and those obtained using RF, NN, and KNN based approaches for a sample daily simulation, consecutively. From these three figures, one can easily see that the best voltage magnitudes are obtained by using KNN based approach. NN based approach simulation results are second best, RF based simulation results are the worst for this sample day. Fig. 8 illustrates the calculated positions of the SVRs, OLTCs and capacitors using KNN, RF and NN based approaches for all simulated days. It is seen from the figure that the simulation results are similar for all three methods.

We used two metrics to measure the performance of the applied methods. The first one is the mean square error (MSE) given by $\sum_{i=1}^{D}(x_i - y_i)^2$, where x_i and y_i represent the real and forecasted values, and D shows the dimension. Table I shows the mean square errors for all 3 methods for the whole simulation time. From the results, it is observed that KNN based approach outperforms the other methods.

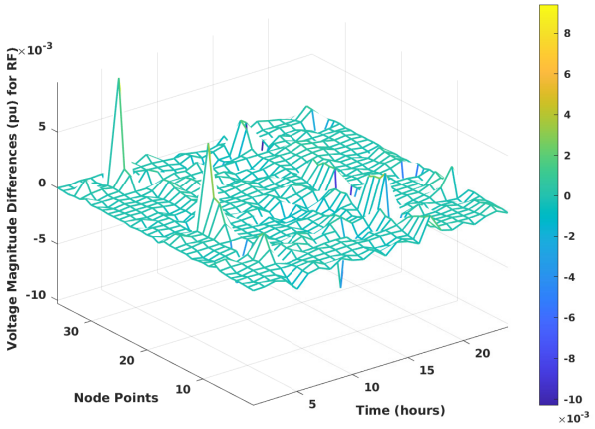

Fig. 5. Hourly voltage differences for RF based approach for a sample day.

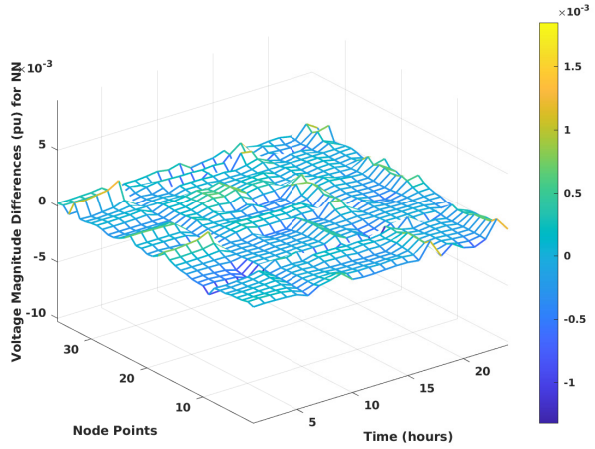

Fig. 6. Hourly voltage differences for NN based approach for a sample day.

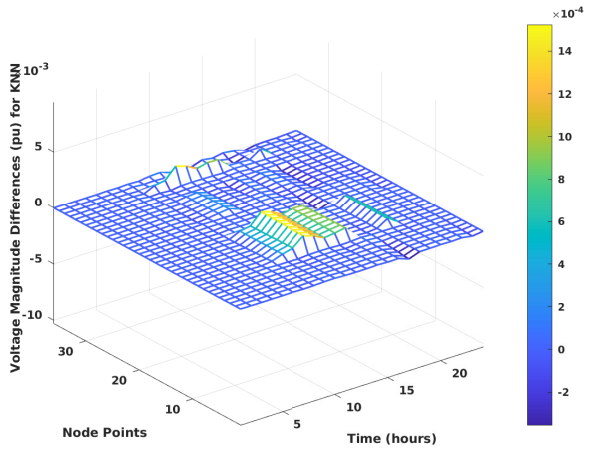

Fig. 7. Hourly voltage differences for KNN based approach for a sample day.

The second metric is the root mean square error (RMSE)

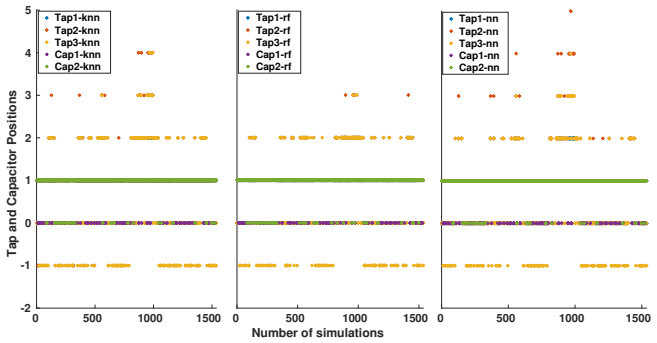

Fig. 8. Hourly tap and on/off positions for KNN, RF, and NN based approaches for all simulated days.

TABLE I
MSE FOR VOLTAGE MAGNITUDES, AND TAP AND SWITCH POSITIONS

MSE	Voltage Magnitudes	Tap and Switch Positions
RF	0.0086	2.6781
NN	0.0056	1.2875
KNN	0.0017	0.7312

and is calculated as,

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

where, n is the number of observations or data points, y_i is the actual value of the dependent variable for the i^{th} observation, $\hat{y}_i$ is the predicted value of the dependent variable for the i^{th} observation. Table II shows the RMSE values for all the methods. RMSE values comply with the MSE values; KNN-based approach gives the best RMSE values, RF-based approach is the worst one. The effectiveness of KNN in our context can be attributed to its ability to capture local patterns and relationships in the dataset.

TABLE II
RMSE FOR VOLTAGE MAGNITUDES, AND TAP AND SWITCH POSITIONS

MRSE	Voltage Magnitudes	Tap and Switch Positions
RF	0.0264	0.5732
NN	0.0187	0.4676
KNN	0.0134	0.4087

V. CONCLUSION

This study aims at integrating efficient machine learning(ML)-based methods for modernizing voltage regulation applications. To this end, we focused on fast and accurate dispatching the legacy voltage regulation devices such as on-load tap changers, step-voltage regulators (SVRs), and switched-capacitors (SCs). Due to the computational hardship of dispatching these devices by frequently solving model-based optimization routines, our approach employs k-nearest neighbor (KNN), random forest (RF), and neural networks (NN). To be able to obtain a trustworthy ML model, a training dataset is generated by solving an efficient optimal power flow (OPF) to find optimal tap positions for OLTCs, SVRs and on/off status decisions. Based on the numerical tests conducted on a modified 33-node test feeder, it is found that the trained ML models have achieved high-accuracy mappings between the input feature vector, which comprises the net active and reactive power injections from each feeder node, and the output vector involving the legacy controls and resulting voltage magnitudes.

REFERENCES

[1] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2006.

[2] S. Civanlar and J. Grainger, "Volt/var control on distribution systems with lateral branches using shunt capacitors and voltage regulators part iii: The numerical results," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 11, pp. 3291–3297, 1985.

[3] B. Palmintier, J. Giraldez, K. Gruchalla, P. Gotseff, A. Nagarajan, T. Harris, B. Bugbee, M. Baggu, J. Gantz, and E. Boardman, "Feeder voltage regulation with high-penetration pv using advanced inverters and a distribution management system: A duke energy case study," *Tech. Report, National Renewable Energy Lab. (NREL), Golden, CO*, 2016.

[4] A. Savasci, A. Inaolaji, and S. Paudyal, "Two-stage volt-var optimization of distribution grids with smart inverters and legacy devices," *IEEE Transactions on Industry Applications*, vol. 58, no. 5, pp. 5711–5723, 2022.

[5] A. Savasci, A. Inaolaji, and S. Paudyal, "Optimal tuning of local voltage control rule of load tap changers for dynamic operation of unbalanced distribution networks," *IEEE Transactions on Industry Applications*, pp. 1–10, 2023.

[6] D. Ranamuka, A. P. Agalgaonkar, and K. M. Muttaqi, "Online voltage control in distribution systems with multiple voltage regulating devices," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 617–628, 2014.

[7] A. Savasci, A. Inaolaji, S. Paudyal, and S. Kamalasadan, "Efficient distribution grid optimal power flow with discrete control of legacy grid devices," in *Proc. 2021 IEEE Power Energy Society General Meeting*, 2021, pp. 1–5.

[8] K. Baker, "Emulating ac opf solvers with neural networks," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4950–4953, 2022.

[9] W. Huang, X. Pan, M. Chen, and S. H. Low, "Deepopf-v: Solving ac-opf problems efficiently," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 800–803, 2022.

[10] M. Zhou, M. Chen, and S. H. Low, "Deepopf-ft: One deep neural network for multiple ac-opf problems with flexible topology," in *Proc. 2023 IEEE Power Energy Society General Meeting*, 2023, pp. 1–1.

[11] D. Owerko, F. Gama, and A. Ribeiro, "Optimal power flow using graph neural networks," in *Proc. 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 5930–5934.

[12] A. S. Zamzam and K. Baker, "Learning optimal solutions for extremely fast ac optimal power flow," in *Proc. 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, 2020, pp. 1–6.

[13] G. W. Kim and K. Lee, "Coordination control of ultc transformer and statcom based on an artificial neural network," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 580–586, 2005.

[14] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Local control of reactive power by distributed photovoltaic generators," in *Proc. first IEEE international conference on smart grid communications*. IEEE, 2010, pp. 79–84.

[15] R. A. Jabr, "Linear decision rules for control of reactive power by distributed photovoltaic generators," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2165–2174, 2018.

[16] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475879

[17] R. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4–22, 1987.

[18] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: http://link.springer.com/10.1023/A:1010933404324

[19] M. Baran and F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.