Optimal Allocation of Soft Open Point Devices in Renewable Energy Integrated Distribution Systems

Ali Farzamnia1, Sassan Marjani1, Sadjad Galvani1, Kenneth Teo Tze Kin2
1Department of Power Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran
2Faculty of Engineering, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
Corresponding author: Sadjad Galvani (e-mail: s.galvani@urmia.ac.ir).
This work was supported by Research Management Center (PPP) and Faculty of Engineering, Universiti Malaysia Sabah(UMS)

ABSTRACT Renewable Energy Sources (RESs) are increasingly integrated into distribution networks due to their undeniable technical and environmental advantages. Despite all the economical and environmental benefits of RESs, they can negatively affect the distribution network. For instance, their output generations are not predictable, and these uncertainties will lead to some operational challenges. Also, RESs are affected by the climate situation, and there are high correlations between them, generally. The correlations between RESs intensify the operational challenges of energy systems. Soft open points (SOPs) are flexible power electronic devices that can effectively increase the efficiency of energy systems. They realize accurate active power control and reactive power compensation to reduce power losses and adjust three-phase voltages. This paper focused on the optimal determination of the location and setting points of SOPs in unbalanced distribution networks in the presence of correlated uncertain sources. The genetic algorithm (GA) was used to solve the main optimization problem, and the correlation between uncertain sources was managed by the Nataf transformation technique. An IEEE 37 bus test system was utilized to illustrate the effectiveness of the proposed method.

INDEX TERMS Data Clustering, Correlation, Genetic Algorithm, Soft Open Point, Uncertainty.

I. INTRODUCTION

A. Motivation

The use of RESs such as wind turbines and photovoltaics (PVs) panels is expanding in distribution networks due to their technical, economic, and environmental benefits such as improved reliability, increased efficiency, reduced rates, and diminished emissions [1].

However, they create some new problems in the operation of distribution systems. For instance, they may reverse power flows and disrupt the protection scheme of distribution networks; they may worsen power quality, and they may increase three-phase unbalancing factors in unbalanced distribution networks [2].

Also, the RESs lead to the non-dispatch ability of these resources and impose other planning and operating challenges [3]. The correlation between random input variables makes the operation of distribution networks even more complicated [4]. Therefore, there is a higher requirement for operational stability and flexibility in distribution networks. SOP are power electronic devices installed in normally open points that can provide effective control of active and reactive power flow among feeders [5]. In this way, the application of fast probabilistic assessment methods with a high level of accuracy is considered for correlation management [4].

B. Literature Review

Over the past years, several studies have been conducted on the use of SOP for improving distribution network performance.

Reference [6] has investigated the effect of the number of SOPs installed in the distribution network on ameliorating the voltage profile as well as, reducing active power losses and load balancing by using the GA and improved Powell’s direct set (PDS) method. In reference [7], the optimal use of energy storage devices in the presence of SOP, distributed generation sources, and network reconfiguration have been investigated to reduce the cost of network utilization, decrease losses, and improve the voltage profile.

In the active distribution networks with a high penetration of PV integration, to eliminate the voltage violations and reduce the power losses, reference [8] proposed a two-stage
adjustable robust optimization model with the column and constraint generation (C & CG) algorithm formulated as a second-order cone program (SOCP) to tackle the uncertainties of PV system outputs.

In [9], in a medium voltage distribution network in the presence of distributed generation sources, the efficiency of SOP was investigated for improving the network voltage profile, lines’ load balancing, and reducing the losses of the lines. Sensitivity analysis based on a Jacobian matrix on load and generation variations was performed to regulate the active and reactive power of the SOP. In reference [10], power supply restoration in distribution networks by SOP was investigated with the primal-dual interior-point algorithm. In [11], the distribution network hosting capacity, for distributed generation sources, was improved by the presence of an SOP to prevent operational malfunctions such as overvoltage and exceeding in the thermal capacity of lines. In reference [12], SOP performance was investigated in the case of single-phase-to-ground, two-phase, and three-phase fault for identification and isolation.

C. Paper Contribution

The main contribution of this paper is considering the voltage unbalancing indices that have not been used in SOP-based studies in a probabilistic environment, recently. Also, other efficient characteristics of this paper can be summarized as follows.

1) Proposing the optimal location and set points of an unbalanced structured SOP in an unbalanced distribution network.
2) Decreasing active power loss as well as improving the unbalancing of the voltages.
3) Considering the probabilistic nature of the distribution system due to loads and renewable generations uncertainties.
4) Considering the correlations among random input variables.
5) The K-means-based data clustering technique was conducted for the probabilistic assessment.
6) The Cholesky decomposition technique and the Nataf transformation were utilized for the management of the correlations between random input variables.

D. Paper Structure

The paper is organized as follows. Section 2 presents the mathematical models for uncertain variables. Section 3 presents the principles of correlation management by Nataf transformation. The application of data clustering in a probabilistic assessment of the distribution network is presented in Section 4. The SOP is modeled in Section 5. Section 6 formulates the problem. The GA is briefly described in Section 7. Simulation results are discussed in Section 8. Finally, the conclusion is presented in Section 9.

II. UNCERTAINTY MODELLING

A. Load

The uncertainty of active loads is usually described as (1) by the normal distribution.

\[ f(P_D) = \frac{1}{\sqrt{2\pi \sigma D^2}} e^{-\frac{(P_D - \mu_D)^2}{2\sigma D^2}} \]  

(1)

Where \( P_D \) is the active load demand; \( f(P_D) \) is the probability density function of \( P_D \); and \( E[P_D] \) and \( \sigma[P_D] \) are the mean and standard deviation of \( P_D \), respectively.

B. Wind generation

Wind speed uncertainty is usually expressed as (2) using the Weibull distribution.

\[ f(v) = \frac{k}{\lambda v^{k-1}} e^{-\left(\frac{v}{\lambda}\right)^k} \]  

(2)

Where \( v \) is the wind velocity; \( f(v) \) is the probability density function of \( v \); and \( k \) and \( \lambda \) are the parameters of Weibull distribution i.e., shape and the scale factors, respectively.

The wind turbine’s active power is dependent on the wind speed, as in (3).

\[ P_W = \begin{cases} v - v_{cut.in} & v_{cut.in} < v \leq v_{rated} \\ v_{rated} & v_{rated} < v \leq v_{cut.out} \\ 0 & \text{other} \end{cases} \]  

(3)

Where \( P_W \) is the active power generation of the wind turbine; \( v_{cut.in} \), \( v_{rated} \), and \( v_{cut.out} \) are the cut-in, rated, and cut-out speeds of the wind turbine, respectively.

The wind generation is connected to the network by converters. As one of the main advantages of converters, their reactive power can be controlled, independently. They are usually set to generate some reactive power to compensate for the reactive power requirements of the network. Reactive power is considered as a fraction of active power to keep the power factor constant, conventionally. This value is normally 0.8-0.9 [4].

The reactive power of the wind turbine was considered with a power factor of 0.85.

III. GENERATING CORRELATED RANDOM VARIABLES

Each variable’s variation affects the other variables when they are interdependent. The covariance and/or correlation coefficient matrix represents the correlation among variables and is calculated through [13].
Consider \( X \) as an \( N \)-dimensional random vector that includes input random variables, i.e., \( X = [X_1 \ X_2 \ \cdots \ X_N]^T \). The correlation matrix of \( X \) is as (4).

\[
R_X = \begin{bmatrix}
R(X_1,X_1) & R(X_1,X_2) & \cdots & R(X_1,X_N) \\
R(X_2,X_1) & R(X_2,X_2) & \cdots & R(X_2,X_N) \\
\vdots & \vdots & \ddots & \vdots \\
R(X_N,X_1) & R(X_N,X_2) & \cdots & R(X_N,X_N)
\end{bmatrix}
\] (4)

where,

\[
R(X_i,X_j) = \frac{\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]}{\sigma_{X_i}\sigma_{X_j}}
\] (5)

The correlation coefficient can be used to express the degree of dependency between two or more variables that range from -1 to 1. Positive values indicate that the variables are simultaneously increased or decreased, whereas negative values indicate that by increasing one variable, the other variable decreases and vice versa. A correlation coefficient closer to 1, suggests a higher degree of dependence. When this coefficient is closer to -1, it indicates a higher degree of non-dependence between the two variables. When the correlation coefficient is zero, the two variables are not correlated [4].

The symmetric matrix \( R_X \) is decomposed by the Cholesky decomposition method as (6).

\[
R_X = LL^T
\] (6)

Where \( L \) is the Cholesky factor of \( R_X \).

Now suppose \( Z = [Z_1 \ Z_2 \ \cdots \ Z_N]^T \) is an independent standard normal vector, i.e. \( \mathbb{E}[Z_i] = 0 \) and \( \sigma[Z_i] = 1 \). Now, the vector \( Y = [Y_1 \ Y_2 \ \cdots \ Y_N]^T \) which is obtained from (7) has a correlation matrix of \( R_X \).

\[
Y = LZ
\] (7)

Now, \( X \) can be generated from the dependent standard normal vector \( Y = [Y_1 \ Y_2 \ \cdots \ Y_N]^T \) as (8).

\[
X_i = F_{X_i}^{-1}(F_{Y_i}(Y_i))
\] (8)

Where \( F_{X_i} \) and \( F_{Y_i} \) denote the cumulative distribution function (CDF) of \( X_i \) and \( Y_i \), respectively [14]. Also, \( F_{X_i}^{-1} \) is the inverse of CDF of \( X_i \).

IV. K-MEANS CLUSTERING ALGORITHM

The probabilistic power flow (PPF) is a useful tool for analyzing the uncertainties existing in a power system with the emergence of a large number of wind power generations. Several methods have been presented for solving the PPF. In general, these methods are categorized into numerical, analytical, and approximation categories [15]. In numerical Monte Carlo simulation-based techniques, samples are produced from stochastic input variables, and then the deterministic load flow problem is solved for each sample [16]. While this method can provide results of high accuracy, it suffers from the shortage of a heavy computation burden and the high storage required for a large number of repeated calculations [17]. Therefore, approximation methods such as the clustering method have been introduced to determine the PPF output random variables with a short computational time and appropriate accuracy.

Clustering is a process of grouping data objects into disjointed clusters. The data in the same cluster are similar, while the data belonging to different clusters differ. Instead of examining a large amount of data, clustered data can be used for the effective probabilistic assessment of distribution networks. Various methods have so far been introduced for data clustering. Mac Queen proposed the K-means clustering algorithm in 1967 [18], which is a simple, efficient, scalable, and fast algorithm when dealing with large datasets.

\( M \) random variables with \( N \) observation for each can be represented by \( N \) observations \((X_1,X_2,...,X_N)\) where each observation is an \( M \)-dimensional real vector. The K-means clustering method tries to partition \( N \) observations into \( K \) sets \((S_1,S_2,...,S_K)\) with an agent for each set generally represented by \( A_k \). The within-cluster sum of squares is minimized as (9) [19].

\[
\min \sum_{k=1}^{K} \sum_{x \in S_k} \|X - A_k\|^2
\] (9)

The K-means method steps are explained as follows [20].

**Step 1:** Select the number of clusters \( K \).

**Step 2:** Select the initial cluster centroids, randomly.

**Step 3:** Assign each observation to the nearest cluster based on the Euclidean distance between the \( i \)th observation and the \( k \)th cluster agent.

**Step 4:** Calculate the mean of objects in each cluster as the new cluster centroids, according to (10).

\[
A_k = \frac{\sum_{x \in S_k} x}{N_{S_k}}
\] (10)

\( N_{S_k} \) shows the number of members in the \( k \)th cluster.

**Step 5:** Repeat from 3 until no change occurs in the agents.

**Step 6:** Assign a probability to each agent as (11).

\[
p(a_k) = \frac{N_{S_k}}{N}
\] (11)

If \( f(X) \) will be a function of \( X \), \( f(X) = f([X_1 \ X_2 \ \cdots \ X_N]^T) \), and the \( E[f^i] \), \( i \)th statistical moment of \( f(X) \) can be calculated according to (12).

\[
E[f^i] = \sum_{k=1}^{K} p(A_k) \cdot f(A_k)^i
\] (12)

Figure 1 schematically represents these steps.
The operating power of SOP is equal to the total power of its VSCs. To evaluate the ability of the SOP in order to improve the network status, converters are modeled by two voltage sources that are linked to each other. This model is considered a mathematical power injection model without considering the details of its structure and control methods hence it enables straightforward incorporation of SOPs into existing power flow analysis algorithms without considering the detailed controller design.

Figure 3 shows a representative one-line diagram of an SOP model with real and reactive power, injecting into buses $m$ and $n$ through both terminals. There are several physical limitations in operating SOPs as follows.

\[
P_{\text{loss}} = P_{\text{loss,m}} + P_{\text{loss,n}}
\]

The injectable active power on both sides of the SOP is limited as (13).

\[
\sum_{\varphi=\text{a}}^{\text{c}} P_{\text{vsc,m}}^\varphi + \sum_{\varphi=\text{a}}^{\text{c}} P_{\text{vsc,n}}^\varphi = P_{\text{loss,m}} + P_{\text{loss,n}}
\]

Where $P_{\text{vsc,m}}^\varphi$ and $P_{\text{vsc,n}}^\varphi$ represent the active power flowing from each phase of both VSCs, and $P_{\text{loss,m}}$ and $P_{\text{loss,n}}$ are the internal power losses of each VSC.

Internal losses of each VSC of SOP are calculated as (14) and (15).

\[
P_{\text{loss,m}} = U_m \cdot P_{\text{vsc,m}}
\]

\[
P_{\text{loss,n}} = U_n \cdot P_{\text{vsc,n}}
\]

$p_{\text{loss,m}}$ and $p_{\text{loss,n}}$ are internal losses of each VSC. $U_m$ and $U_n$ are the loss coefficients of two converters at nodes $m$ and $n$. according to (14) and (15) converters losses are proportional to their internal power flow [10]. The associated complex loss is modeled by multiplying the injected active power of each VSC to a constant coefficient, conventionally. this procedure was used in references [10], [23].

The operational boundaries of the SOP are according to (16) to (18).
The main component of the objective function is power losses. Besides, the network voltage unbalance factor (VUF) is considered as the second component of the objective function. The objective functions are expressed as follows.

\[
P_{\text{loss}} = \sum_{i=1}^{N_{\text{branch}}} \sum_{\phi} r_{i\phi} \times I_{i\phi}^2
\]  

(24)

Where \( r_{i\phi} \) and \( I_{i\phi} \) are resistance and current of \( i^{th} \) branch and phase \( \phi \), respectively.

In standard publication number MG 1-1993, the National Electrical Manufacturers Association (NEMA) defines the voltage unbalance factor as “maximum deviation from the average” over “the average of three-phase voltages” multiplied by 100 [24]. This index is formulated as (25).

\[
VUF = \frac{\max(V_{\text{avg}} - V_{\text{phase}})}{V_{\text{avg}}} \times 100
\]  

(25)

Where \( V_{\text{phase}} \) is the voltage of each phase, and \( V_{\text{avg}} \) is the average of three-phase voltages.

However, the system has a delta configuration but the phase voltages can be still calculated by measuring them with respect to the ground. The voltage unbalancing index can be obtained using the phase to phase voltage as you said but the final results will be the same. If the line voltages unbalancing are reduced, the phase voltages unbalancing are reduced, consequently.

By defining the voltage unbalance index in the form of the sum of voltage unbalance factor indices as expressed in (26), the balancing of bus voltages is achieved by minimizing this index.

\[
voltage unbalance index (VUI) = \sum_{i=1}^{N_{\text{bus}}} VUF_i
\]  

(26)

\( VUF_i \) is the voltage unbalance factor for each bus; \( N_{\text{bus}} \) is the number of total buses.

Unbalanced voltages in unbalanced networks cause inefficient operation and increased losses in motor loads. In this study, SOP was utilized to reduce VUF which, according to the NEMA standard, should be \(<1\%\).

Problem formulation by a linear weighted combination of the objective functions is expressed below.

\[
\text{minimize } F_{\text{obj}}(S) = P_{\text{loss}} + VUI + \alpha \times \sum_{x=1}^{k} f_{\text{sop}}^x + \beta \times \sum_{x=1}^{\ell} f_{\text{vsc}}^x
\]  

(27)

Objective functions are normalized by dividing them into their base value. Also, penalty functions are presented for avoiding violation of the constraints. \( \alpha \) and \( \beta \) are weight coefficients of the penalty functions, and they are usually set at big values.

The \( S \) vector contains the problem control variables, according to (28).

\[
S = [L_{\text{vsc,n}}^1, L_{\text{vsc,n}}^1, \ldots, L_{\text{vsc,n}}^k, t_{\text{vsc,m}}^k, p_{\text{vsc,n}}^a, p_{\text{vsc,n}}^b, p_{\text{vsc,n}}^c, Q_{\text{vsc,n}}^a, Q_{\text{vsc,n}}^b, Q_{\text{vsc,n}}^c, \ldots, p_{\text{vsc,m}}^a, p_{\text{vsc,m}}^b, p_{\text{vsc,m}}^c, Q_{\text{vsc,m}}^a, Q_{\text{vsc,m}}^b, Q_{\text{vsc,m}}^c, p_{\text{vsc,m}}^k, p_{\text{vsc,m}}^c, p_{\text{vsc,m}}^k, Q_{\text{vsc,m}}^c, Q_{\text{vsc,m}}^c, Q_{\text{vsc,m}}^c]
\]  

(28)

Where,

\( L_{\text{vsc,n}} \) and \( t_{\text{vsc,m}} \) represent the number of buses connected to both VSCs of the \( k^{th} \) SOP; \( p_{\text{vsc,n}}^a, p_{\text{vsc,n}}^b \) and \( p_{\text{vsc,n}}^c \) represent the active power of \( VSC_n \) of the \( k^{th} \) SOP from phase a to c;
where, $V_{i,q}$ is the voltage of bus $i$ and phase $q$; $V^\text{min}$ and $V^\text{max}$ are the minimum and maximum voltage limits; $VUF_i$ is the voltage unbalance factor for bus $i$; $I_{i,q}$ is the current of branch $i$ and phase $q$; and $I^\text{max}_i$ is the maximum current capacity of branch $i$.

### VII. GENETIC ALGORITHM

The optimization method in this study was based on a GA. The GA is an optimization algorithm that was invented to mimic some of the processes observed in natural evolution. The GA is a stochastic search technique based on the mechanism of natural selection and natural genetics and presents only some practical solutions to the problem [25]. Also, mathematical-based optimization methods can not be applied to this problem without some simplifications. In order to take the complete model of the problem, the use of evolutionary-based techniques is required. However, they may reach different solutions in each execution. The steps to implement this algorithm are as follows [26].

**Step 1:** Coding chromosomes, where each chromosome contains a number of the particle (called gene) and every particle is a control variable. Figure 4 shows the structure of a chromosome, schematically.

**Step 2:** Creating a primitive population of a certain number of chromosomes, where every particle is initialized randomly within its bounds.

The initial population has a size of $N_{pop} \times N_S$ where $N_{pop}$ is the population size and $N_S$ is the number of control variables or genes in each chromosome. The $i^{th}$ chromosome is formed according to (39) and (40).

$$P_i = [p_{i,1}, p_{i,2}, \ldots, p_{i,j}, \ldots, p_{i,N_S}]$$

(39)

$$p_{i,j} = (p_{j}^{\text{max}} - p_{j}^{\text{min}})U + p_{j}^{\text{min}}$$

(40)

where, $p_{i,j}$ is the $i^{th}$ chromosome; $P_{j}$ is the $j^{th}$ control variable for each chromosome; $p_{j}^{\text{max}}$ is the maximum limit of the $j^{th}$ control variables; $p_{j}^{\text{min}}$ is the maximum limit of the $j^{th}$ control variables; $U$ is a random number between 0 and 1 generated uniformly.

**Step 3:** Evaluating the objective function for all population chromosomes. According to (27) This equation is generally shown by (41).

$$f_i = f(P_i)$$

(41)

The $j^{th}$ gene of the new chromosome is formed according to (42).
Where, 

\[ p_{q,i} \] 

is the \( i^{th} \) gene of \( q^{th} \) chromosome as the first parent. 

\[ p_{r,j} \] 

is the \( j^{th} \) gene of \( r^{th} \) chromosome as the second parent. 

\[ \gamma \] 

is a random number between 0 and 1 selected uniformly. 

\[ new \] 

denotes the new chromosome as a child. 

After generating a new child from the crossover operator, the mutation operator is applied to the new population according to (43), 

\[ p_{i,j} = (p_{j}^{\text{max}} - p_{j}^{\text{min}})U + p_{j}^{\text{min}} \] (43)

In this equation \( j^{th} \) gene of \( i^{th} \) chromosome is replaced with a random number in its bound. 

The number of chromosomes and the number of genes to apply the mutation operator are dependent on the mutation rate. If this rate is high, the number of selected chromosomes for mutation and the number of genes for selected chromosomes are increasing. 

**Step 5:** Evaluating the objective function for a new population and examining the stopping condition to terminate or continue the algorithm. 

**Step 6:** Choosing the best chromosome in terms of the objective function. 

In the probabilistic environment, samples are clustered into agents, and the objective function is calculated using (27). 

Figure 5 depicts an overview of the proposed method and its solving procedure by GA. The GA (as well as other evolutionary-based optimization techniques) can not guarantee to find the global optimum point, as they have stochastic evolutionary procedures, they may reach various results in each execution. For this means, they need to be repeated in large numbers to get assurance from the good convergence of the algorithm. A large number of generations can guarantee the same results in the next execution, to some extent (not certainly).

**VIII. Simulation and Results**

**A. Case study**

The proposed optimization method was implemented on an IEEE standard 37-bus test feeder [27]. The network one-line diagram is displayed in Fig. 6. This feeder is an actual feeder located in California, which is rated at 4.8 kV and has a total demand of 2461 kW and 1202 kVAR. This network is heavily unbalanced loaded.

The network used in this study did not specify any open network switches. Therefore, here, the SOP could be located between two different buses, according to constraint (35). Also, the number of SOP allocations is continued until there was no further improvement in the objective functions simultaneously.
B. Assumptions

Some simulation assumptions are given below. The GA parameters are assumed as follows.

- The population size is considered to be 100.
- The number of iterations is set to 1000.
- The crossover operator is based on the blending method with a rate of 0.8.
- The mutation operator is based on the uniform changes and its rate is 0.1.
- The roulette wheel parent selection technique is employed for the selection of the parents.
- Weight coefficients $W_1$ and $W_2$ are set at 0.5.
- Penalty coefficient $\alpha$ and $\beta$ are set at 10, experimentally.

Also, the assumptions for uncertainty variables are made as follows.

- Two wind turbines are installed at buses 28 and 32, and three random loads are considered at buses 3, 7, and 8.
- The wind speed is modeled with the Weibull distribution with a scale factor of 7.28 and a shape factor of 2.01 [4].
- Wind turbine parameters are assumed to follow values in Table I [4].
- Random loads are modeled with a normal distribution whose mean value is 100 kW and the standard deviation value is 2 kW.
- The wind speed at bus 28 is correlated to the wind speed at bus 32 by 0.7.
- The correlation coefficient between random loads is considered equal to 0.1.
- The loads at buses 3, 7, and 8 are correlated with the wind speed at buses 28 and 32 by 0.2.
- The apparent power of each VSC is assumed to approximately be 20% of the total network active load level. The total network load is 2461 kW and 20% of it is 492.2 kW, so approximately, it is assumed that the rated power of each VSC is 500 kW.
- Also, related to the probabilistic assessment by the K-means data clustering method, the number of clusters, K is equal to 5.
- It is assumed that per length parameters of the line connecting two ends of the SOP are identical with parameters of line 1-2.
- Also, it is assumed that $U_m$ and $U_n$ in equations (14) and (15) are equal to 0.005.

| TABLE I |
| WIND TURBINES PARAMETERS |
| Parameter (wind turbine) | Value |
|--------------------------|-------|
| Rating capacity          | 100 (kW) |
| Cut-in speed             | 3 (m/s) |
| Rating speed             | 12.5 (m/s) |
| Cut-out speed            | 25 (m/s) |

Figures 7 and 8 show a scatter plot with two marginal histograms to visualize the relationship between the samples of wind speed at buses 28 and 32, and the samples of loads at buses 7 and 8, respectively. Also, Fig. 9 illustrates the scatter plot with two marginal histograms to visualize the relationship between samples of loads at bus 8 and wind speed samples at bus 32. These types of plots can visualize the dependency of samples.
In reference [23], to optimally operate a three-phase unbalanced network using SOP, SOP is assumed to be three single-phase SOPs, each of which is used in a phase. However, in this study, we considered each of the VSCs as a three-phase power electronic device, each of which produced its active and reactive power with SOP constraints.

C. Discussions

To investigate the effect of SOP on an unbalanced network and to improve the objective functions mentioned in Section VI, two scenarios were considered.

Scenario I: Unbalanced network operation with the presence of wind turbines and random loads without the presence of SOP.

In the first scenario, the network is assessed with uncertainty sources and without the presence of SOP. The power losses in this scenario are 32.64 kW; the voltage unbalancing index, VUI, is 20.98 %. In this scenario, buses number 12, 31, 32, 33 and 34 have a VUF of more than 1 %.

Scenario II: Optimal operation of an unbalanced network with uncertainty sources and SOPs. The second scenario is studied once in the presence of one SOP and then with two SOPs.

Simulation results indicate that in Scenario II, to exploit one SOP in the network, the optimal location for allocating the SOP is between buses 2 and 28. Table II presents the details of the proposed solution obtained by the optimization algorithm. The active and reactive power flows from both VSCs of the SOP are seen in this table. Positive values indicate the injection of power into the network and negative values indicate the absorption of power from the network. By installing an SOP between buses 2 and 28 and setting the set-points according to table II, the power losses are reduced from 32.64 kW to 24.8 and the VUI is reduced from 20.98 % to 9.57 %. There is no bus with a VUF of more than 1 %, in this scenario. According to table II, the sum of active power flows in three phases of two voltage sources is equal to -2.1. This means the losses of the SOP are 2.1 kW. The negative sign shows the absorption of the power from the network by voltage sources of the SOP. However, the sum of the reactive powers doesn’t obey the previous rule. The voltage sources can inject or absorb reactive power freely and there is not necessary that their sum is equal to zero or near zero. The total rate of the SOP is 993 kVA.

Besides, by using two SOPs in the network, the optimal location of the first SOP is between buses 4 and 32, and the second one is between buses 2 and 18. Table III presents the details of the proposed solution obtained by the optimization algorithm. The active and reactive power flows from both VSCs of the SOP are seen in this table. By installing two SOPs in mentioned locations and setting the set-points according to table III, the power losses are reduced 22.78 kW and the VUI reaches 7.38 %. There is no bus with a VUF of more than 1 %, in this scenario.

As discussed before, the sum of active power flows is near zero and this represents the SOP converter’s losses. However, the sum of reactive power flows is not zero and they are dictated by the reactive power requirement of the network. These reactive powers are required for voltage balancing.

The comparison of the results of the two scenarios is given in Table IV.

| TABLE II | ACTIVE AND REACTIVE POWER OF SOP IN SCENARIO II WITH ONE SOP |
|---------|----------------------------------------------------------|
| SOP1    |                                                          |
| Active power (kW) | VSC-2 | 145.8 | -137.1 | -119.3 | 2.1 |
|          | VSC-28 | 209.4 | -30.6  | 221.3  |     |
| Reactive power (kVAr) | VSC-2 | 57.7  | 93.5   | 137    | 585.3 |
|          | VSC-28 | 150.8 | 61.5   | 84.8   |     |
| Rating (kVA) |         | 993   |        |        |     |

| TABLE III | ACTIVE AND REACTIVE POWER OF SOP IN SCENARIO II WITH TWO SOP |
|-----------|---------------------------------------------------------------|
| SOP1      | Active power (kW) | VSC-4 | 157.3 | -193.6 | -153 | -1.2 |
|          | VSC-32 | 130.8 | -149  | 206.3  |     |
| Reactive power (kVAr) | VSC-4 | 87.5  | 100.8  | 180  | 553.5 |
|          | VSC-32 | 126.7 | -17.5 | 76    |     |
| Rating (kVA) |         | 678.3 |        |       |     |
| SOP2      | Active power (kW) | VSC-2 | -182.5 | -160.3 | -27.1 | -1.5 |
|          | VSC-18 | 33    | 212.8  | 122.6  |     |
| Reactive power (kVAr) | VSC-2 | 73.3  | 98.7   | 148.5  | 489.1 |
|          | VSC-18 | 31.6  | 98.3   | 38.7   |     |
| Rating (kVA) |         | 894.5 |        |       |     |

| TABLE IV | SIMULATION RESULTS |
|----------|---------------------|
| Scenario I | Scenario II |
| (1 SOP) | (2 SOP) |
| Power losses (kW) | 32.64 | 24.8 | 22.78 |
| Power loss reduction % | - | 24.02 | 30.21 |
| Voltage unbalance index (VUI) | 20.98 | 9.57 | 7.38 |
| VUI reduction % | - | 54.38 | 64.82 |
| Bus with VUF beyond 1% | 12-31-32-33-34 | - | - |
| Minimum voltage (p.u.) | A | 0.9803 | 0.9893 | 0.9929 |
|          | B | 0.9861 | 0.9857 | 0.9888 |
|          | C | 0.9807 | 0.9887 | 0.9910 |

In the first scenario, the active power losses equal 32.64 kW and the VUI equals 20.98. Also, the voltage unbalance factor for buses 12, 31, 32, 33, and 34 is >1%.

By operating an SOP in the network, the active power losses are reduced by 24.02 % and reach from 32.64 kW to 24.8 Kw. Also, the voltage unbalance index is decreased by 54.38 % and reaches from 20.98 to 9.57.
By allocation and regulation of two SOPs in the network, further improvements to the network are achieved, such that the active power losses and VUI are reduced by 30.21% and 64.82%, respectively. Figure 10 shows the voltage unbalance factor for all the buses in two scenarios. As shown in Fig. 10, the VUF is considerably low in the second scenario and is <1% for all the buses. The three-phase voltage profile for both scenarios is presented in Fig. 11. From figures 10 and 11, it can be found that the SOP-based unbalanced operation strategy, which is implemented in Scenario II considerably mitigates the voltage unbalanced condition and improves the voltage profile. Moreover, the minimum voltage of each phase of all buses in Scenario II is closer than Scenario I to the standard value of 1 p.u.

The CDF of state variables such as active power losses and buses voltage gives full probabilistic information about these variables values. In order to compare the accuracy of the K-means method with the basic Monte Carlo method in estimating the distribution function of network output variables, the CDF of the network active power losses and the voltage magnitude of bus 12 in phase A are drawn.

Figure 12 illustrates the CDF of the network active power losses, respectively during the absence and presence of an SOP in the network with both Monte Carlo and K-means evaluation methods. Based on Fig. 12, the probability that the active power losses without the presence of an SOP in the network are less than 32.64 kW is 0.4 in the K-means method and is equal to 0.409 in the Monte Carlo method. In addition, when the network operates with one SOP, the probability that the active power loss is lower than 24.8 kW is equal to 0.4 in the K-means method and is 0.414 in the Monte Carlo method. Therefore, the accuracy of the K-means method is significant in estimating the CDF.

Figure 13 displays the CDF of the voltage of bus 12 in Phase A, respectively in the absence and presence of an SOP in the network with both Monte Carlo and K-means evaluation methods. Based on Fig. 13, the probability that the voltage magnitude without the presence of an SOP in the network is less than 0.9894 is 0.6 in the K-means method and is 0.618 in the Monte Carlo method. Therefore, the accuracy of the K-means method is suitable for estimating the CDF of the voltage of the system.
Figure 12. CDF plot of network active power losses in scenario I and II. (a) CDF plot of network active power losses in scenario I. (b) CDF plot of network active power losses in scenario II with one SOP

Figure 13. CDF plot of the voltage of bus #12 at phase A in scenario I and II. (a) scenario I. (b) scenario II with one SOP

IX. Conclusions

In this paper, the SOP is used to improve the performance of the unbalanced distribution networks. Optimal location and set points of the SOP are found to decrease the losses and improve the unbalancing of the system.

In order to make the results more applicable, the uncertainty of the distribution network, as well as the correlations among various input random variables, was considered. A backward-forward load flow algorithm with a generic power injection model of SOP was developed for modeling the SOP. Also, the SOP losses are considered to obtain reasonable results.

The correlations among various input random variables were managed using the Cholesky decomposition technique and the Nataf transformation. Furthermore, the probabilistic assessment of the system was conducted by a popular scheme of data clustering, i.e., the K-means method.

The optimal location and active and reactive power setting points of the SOP were obtained using a GA while considering the active power losses and voltage unbalance index as objective functions.

Simulation results showed that an SOP with a rate of 993 kVA could decrease active power losses by 24.02% while improving the voltage unbalance index from 20.98 to 9.57.

Two SOPs with the rate of 678.3 and 894.5 can decrease the losses by 30.21% and improve the voltage unbalancing index from 20.98 to 7.38.

Also, the statistical behavior of the network’s active power losses and some selected buses’ voltage was obtained that can be very useful in operational decision making for distribution networks. This gives extensive information about all probable values that can be encountered during the system’s various operating conditions. For example, these figures answer some important questions, such as what is the probability of the voltage of the desired bus at a desired phase being greater/less than a specific amount. Answers to these questions affect the operational decisions, completely.
REFERENCES

[1] H. Liu, C. Tang, J. Han, T. Li, J. Li, and K. Zhang, “Probabilistic load flow analysis of active distribution network adopting improved sequence operation methodology,” IET Gener. Transm. Distrib., vol. 11, no. 9, pp. 2147–2153, 2017.

[2] P. Li, Z. Wu, Y. Wang, X. Dou, M. Hu, and J. Hu, “Adaptive robust optimal reactive power dispatch in unbalanced distribution networks with high penetration of distributed generation,” IET Gener. Transm. Distrib., vol. 12, no. 6, pp. 1382–1389, 2018.

[3] S. Galvani, M. T. Hagh, and M. B. B. Sharifian, “Unified power flow controller impact on power system predictability,” IET Gener. Transm. Distrib., vol. 8, no. 5, pp. 819–827, 2014.

[4] S. Galvani and M. Choogan, “Data clustering based probabilistic optimal power flow in power systems,” IET Gener. Transm. Distrib., vol. 13, no. 2, pp. 181–188, 2019.

[5] H. Ji, C. Wang, P. Li, J. Zhao, G. Song, and J. Wu, “Quantified flexibility evaluation of soft open points to improve distributed generator penetration in active distribution networks based on difference-of-convex programming,” Appl. Energy, vol. 218, pp. 338–348, 2018.

[6] W. Cao, J. Wu, N. Jenkins, C. Wang, and T. Green, “Benefits analysis of soft open points for electrical distribution network operation,” Appl. Energy, vol. 165, pp. 36–47, 2016.

[7] L. Bai, T. Jiang, F. Li, H. Chen, and X. Li, “Distributed energy storage planning in soft open point based active distribution networks incorporating network reconfiguration and DG reactive power capability,” Appl. Energy, vol. 210, pp. 1082–1091, 2018.

[8] H. Ji, C. Wang, P. Li, F. Ding, and J. Wu, “Robust operation of soft open points in active distribution networks with high penetration of photovoltaic integration,” IEEE Trans. Sustain. Energy, vol. 10, no. 1, pp. 280–289, 2018.

[9] C. Long, J. Wu, L. Thomas, and N. Jenkins, “Optimal operation of soft open points in medium voltage electrical distribution networks with distributed generation,” Appl. Energy, vol. 184, pp. 427–437, 2016.

[10] P. Li, G. Song, H. Ji, J. Zhao, C. Wang, and J. Wu, “A supply restoration method of distribution system based on Soft Open Point,” in 2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia), 2016, pp. 535–539.

[11] L. J. Thomas, A. Burchill, D. J. Rogers, M. Guest, and N. Jenkins, “Assessing distribution network hosting capacity with the addition of soft open points,” 2016.

[12] A. Aithal, G. Li, J. Wu, and J. Yu, “Performance of an electrical distribution network with Soft Open Point during a grid side AC fault,” Appl. Energy, vol. 227, pp. 262–272, 2018.

[13] C. S. Saunders, “Point estimate method addressing correlated wind power for probabilistic optimal power flow,” IEEE Trans. Power Syst., vol. 29, no. 3, pp. 1045–1054, 2013.

[14] M.-R. Andervazh and S. Javadi, “Emission-economic dispatch of thermal power generation units in the presence of hybrid electric vehicles and correlated wind power plants,” IET Gener. Transm. Distrib., vol. 11, no. 9, pp. 2232–2243, 2017.

[15] Y.-K. Tung and B. C. Yen, Hydro systems engineering uncertainty analysis. McGraw-Hill New York, 2005.

[16] R. Y. Rubinstein and D. P. Kroese, Simulation and the Monte Carlo method, vol. 10. John Wiley & Sons, 2016.

[17] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, “Probabilistic load flow evaluation with hybrid latin hypercube sampling and cholesky decomposition,” IEEE Trans. Power Syst., vol. 24, no. 2, pp. 661–667, 2009.

[18] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967, vol. 1, no. 14, pp. 281–297.

[19] S. Song, C. Han, S. Jung, M. Yoon, and G. Jang, “Probabilistic power flow analysis of bulk power system for practical grid planning application,” IEEE Access, vol. 7, pp. 45494–45503, 2019.

[20] M. Ramezani, C. Singh, and M.-R. Haghigham, “Role of clustering in the probabilistic evaluation of TTC in power systems including wind power generation,” IEEE Trans. Power Syst., vol. 24, no. 2, pp. 849–858, 2009.

[21] T. Friedli, J. W. Kolar, J. Rodriguez, and P. W. Wheeler, “Comparative evaluation of three-phase AC–AC matrix converter and voltage DC-link back-to-back converter systems,” IEEE Trans. Ind. Electron., vol. 59, no. 12, pp. 4487–4510, 2011.

[22] A. Rodríguez-Cabero, F. H. Sánchez, and M. Prodanovic, “A unified control of back-to-back converter,” in 2016 IEEE Energy Conversion Congress and Exposition (ECCE), 2016, pp. 1–8.

[23] P. Li et al., “Optimal operation of soft open points in active distribution networks under three-phase unbalanced conditions,” IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 380–391, 2017.

[24] A. Kharrazi, V. Sreeram, and Y. Mishra, “Assessment of voltage unbalance due to single phase rooftop photovoltaic panels in residential low voltage distribution network: A study on a real LV network in Western Australia,” in 2017 Australasian Universities Power Engineering Conference (AUPEC), 2017, pp. 1–6.

[25] K. Mahmoud and M. Lehtonen, “Comprehensive analytical expressions for assessing and maximizing technical benefits of photovoltaics to distribution systems,” IEEE Trans. Smart Grid, 2021.

[26] R. L. Haupt and S. E. Haupt, Practical genetic algorithms. John Wiley & Sons, 2004.

[27] M. Bazrafshan and N. Gatsis, “Comprehensive modeling of three-phase distribution systems via the bus admittance matrix,” IEEE Trans. Power Syst., vol. 33, no. 2, pp. 2015–2029, 2017.