# Synthetic Distribution Grid Generation Using Power System Planning: Case Study of Singapore

**Conference Paper** · August 2018

DOI: 10.1109/UPEC.2018.8542054



Andrej Trpovski
TUM CREATE

# Synthetic Distribution Grid Generation Using Power System Planning: Case Study of Singapore

Andrej Trpovski[1‡], Dante Recalde[1§], Thomas Hamacher[2¶]
[1]TUMCREATE, 1 CREATE Way, #10-02 CREATE Tower, Singapore 138602
[2]Technical University of Munich (TUM), Chair of Renewable and Sustainable Energy Systems, Garching 85748, Germany
[‡]andrej.trpovski@tum-create.edu.sg [§]dante.melo@tum-create.edu.sg [¶]thomas.hamacher@tum.de

***Abstract*—Having grid models of existing power distribution systems can be essential for analysis, research and development of innovative smart grid methods and concepts. However, such models are often not available due to security concerns and confidentiality restrictions. This paper presents a novel approach for generating synthetic distribution grids that correspond to existing power distribution systems. The synthetic distribution grid generation model is defined using a power system planning optimization. The model is developed and structured to fit and utilize publicly available data of consumer demand and location, number of substations and their placement and specifications and estimate pricing of installed equipment. An equally weighted multi-objective cost minimization function of investment and operational costs is defined. To represent the physical constraints of the grid, a convex AC power flow is presented. A case study of the Singaporean power distribution system is shown.**

***Index Terms*—Synthetic Power System, Power Distribution, Power System Planning, Smart Grids, Optimization, Quadratic Programming.**

## NOMENCLATURE

*Sets*

| | | |
|---|---|---|
| $\Omega^l$ | Set of lines | |
| $\Omega^n$ | Set of nodes | |
| $\Omega^s$ | Set of substations | |

*Constants*

| | | |
|---|---|---|
| $n^n$ | number of nodes | |
| $n^s$ | number of substations | |
| $\Psi_{i,j}^{line}$ | Fixed investment cost for lines | $[S\$/km]$ |
| $\Psi_{i,j}^{loss}$ | Cost of power losses in lines | $[S\$/MWh]$ |
| $\Psi_i^{subs}$ | Cost of switchgear equipment | $[S\$/substation]$ |
| $\Psi_i^{trafo}$ | Cost of transformers in substations | $[S\$/pcs]$ |
| $L_{i,j}$ | Line length | $[km]$ |
| $R_{i,j}$ | Line resistance | $[\Omega/km]$ |
| $X_{i,j}$ | Line reactance | $[\Omega/km]$ |
| $Z_{i,j}$ | Line impedance | $[\Omega/km]$ |
| $P_i^d$ | Active power demand | $[MW]$ |
| $Q_i^d$ | Reactive power demand | $[MVar]$ |
| $S^{base}$ | Base apparent power for p.u. system | $[MVA]$ |
| $I_{ij}^{max}$ | Maximum current magnitude | $[kA]$ |
| $V^{max}$ | Maximum voltage magnitude | $[kV]$ |
| $P_i^{gmax}$ | Maximum generation of active power | $[MW]$ |
| $Q_i^{gmax}$ | Maximum generation of reactive power | $[MW]$ |
| $S_{i,j}^{max}$ | Line maximum apparent power | $[MVA]$ |
| $V^{min}$ | Minimum voltage magnitude | $[kV]$ |
| $P_i^{gmin}$ | Minimum generation of active power | $[MW]$ |
| $Q_i^{gmin}$ | Minimum generation of reactive power | $[MW]$ |
| $P_{tot}^d$ | Total active power demand | $[MW]$ |
| $\Xi^{line}$ | Annuity factor for lines | |
| $\Xi^{subs}$ | Annuity factor for substations | |
| $t$ | Total lifespan of investment | $[years]$ |
| $u$ | Annual discount rate | |
| $EPLT$ | Equivalent peak loss time | $[hours]$ |
| $N^{Tmax}$ | Maximum number of transformers per node | |

*Continuous variables*

| | | |
|---|---|---|
| $\Lambda$ | Total cost of objective function | $[S\$]$ |
| $\Lambda^*$ | Total cost of objective function | $[S\$/year]$ |
| $\lambda_1$ | Fixed investment cost of lines | $[S\$]$ |
| $\lambda_2$ | Variable cost of lines | $[S\$]$ |
| $\lambda_3$ | Investment cost of substations | $[S\$]$ |
| $\lambda_1^*$ | Annual fixed investment cost of lines | $[S\$/year]$ |
| $\lambda_2^*$ | Annual variable cost of lines | $[S\$/year]$ |
| $\lambda_3^*$ | Annual investment cost of substations | $[S\$/year]$ |
| $I_{i,j}$ | Current magnitude through lines | $[kA]$ |
| $V_i$ | Voltage magnitude | $[kV]$ |
| $P_{i,j}$ | Line active power flow | $[MW]$ |
| $Q_{i,j}$ | Line reactive power flow | $[MVar]$ |
| $P_i^g$ | Active power generated at node i | $[MW]$ |
| $Q_i^g$ | Reactive power generated at node i | $[MVar]$ |
| $\omega_{i,j}$ | Auxiliary variable to model the state of the lines | |

*Integer variables*

| | | |
|---|---|---|
| $\tau_i$ | Number of transformers at node i | |

*Binary variables*

| | | |
|---|---|---|
| $\alpha_{i,j}$ | Decision variable for line $ij$ | |

## I. INTRODUCTION

Electrical power systems are facing progressive challenges imposed by revolutionary technological breakthroughs. Distributed renewable energy generation combined with advanced information and communication technologies and rapid development of storage systems are just an introduction to the new Smart Grid.

Distribution grid models can be essential for the advancement of the power distribution system analysis and research on novel methods and concepts. The introduction of many new agents in the existing distribution system sets a transition in which the distribution system is becoming an active segment of the power system [1]. Using real distribution grid models can

978-1-5386-2910-9/18/$31.00 ©2018 IEEE

be essential to understand the performance of the distribution grid and help to identify the challenges for new methods and concepts in the future smart grid. However, a detailed power distribution system models of existing grids are not publicly available because of its importance as a high priority critical infrastructure and confidentiality restrictions.

Several different methods for generating synthetic distribution grids are studied and known in the literature [2]–[4]. Using large data set from a particular distribution system operator (DSO), the authors in [2] identify statistical patterns and propose an algorithm for automated generation of random distribution grids that statistically resemble the real grids. In [3], an algorithm that generates statistically correct random topologies for testing of the future smart grid systems is proposed. The algorithm is developed by studying both topological and electrical characteristics, based on a number of synthetic and real-world power systems. Using correlated assignment of generation, load and connection buses, an improved synthetic power grid modeling approach is proposed in [4]. Similarly, the characterization of the correlation of the bus type assignment is done by data obtained from number of different realistic grids.

The proposed methods generate random synthetic grids based on the statistical properties derived from analysis of limited samples of real and synthetic power systems. However, many of the distribution systems can not be described using statistical properties due to the lack of publicly available data. In [5], the authors show that results for many research topics in modern distribution systems are very often case specific and have limited applicability to other systems. The topological properties of a distribution system describe how different the configurations are, which can be critical when the performance of the system is being evaluated.

For many distribution grids, a partial set of data of the distribution system is publicly disclosed in the forms of annual and statistical reports from governmental institutions and DSOs. In this paper, we take the Singaporean distribution system as a case study for which data of consumer demand and location, substations, transformer sizing and electricity price are publicly available in publications and statistics provided by Energy Market Authority of Singapore (EMA) and Singapore Power as a DSO of Singapore [6], [7].

This paper proposes a novel approach for generating synthetic power distribution grid that corresponds to an existing power distribution system utilizing publicly available data. The approach is defined using a power distribution system planning optimization. Distribution system planning models are more frequently introduced as there is a need for restructuring the grid towards a more efficient, reliable and cost effective power systems. Many different distribution system planning techniques and algorithms are being developed and improved over the last four decades [8]. With the rapid advancement in information technologies, more accurate, more detailed and more sophisticated models are defined [9]. However, when generating a synthetic grid, most of the information typically known to the planning engineer are not publicly available, such as the routes and lengths of potential lines. Hence, potential line routes and line lengths can be determined based on the geographical and spatial evaluation such that obstacles like ponds, lakes, rivers and already existing buildings and infrastructure are considered. Having defined a set of potential lines, the distribution system planning approach proposed in this paper is used to generate a synthetic distribution grid topology and allocate the loads to adequate substations by utilizing the available system data.

In the proposed method, a radially operated distribution network is considered. A balanced distribution system is assumed, represented by its single phase equivalent as a steady state of the system. It is a simplified AC power flow model based on an electric equivalent radial power flow with eliminated voltage phase angle [10]. The AC definition of the power flow model makes sure that the electrical constraints of the grid such as voltage and current limits are considered and satisfied. Similar definitions are being discussed in [11], [12]. The nonlinear part of the power flow definition in Section III has less binary variables and constraints compared to [11], and it is a simplified version with less continuous variable and constraints compared to [12]. The simplifications made due to the different purpose and different objectives of the models can be beneficial when considering a large scale combinatorial problem.

While much effort in Complex Network Science (CNS) is allocated for statistical analysis and validation of transmission grids [13], there is a limited number of studies in the literature that investigate distribution grids in a statistical manner. The synthetic distribution grid generated by the method presented in this paper is validated using statistical data from a DSO in The Netherlands, detailed in [2], [14].

## II. Synthetic Grid Generation Model

### A. Cost Minimization Objective Function

Power distribution systems are planned and engineered in a cost efficient manner with respect to the overall investment and operational costs [15]. The synthetic power distribution grid model is generated employing a similar cost efficient grid planning approach that is used by the planning engineers. The synthetic grid generation model is defined using power system planning optimization that utilizes the publicly available data of the existing grid, which makes the synthetic grid to have high resemblance to the actual grid.

The cost minimization objective function is defined as follows

$$\min \qquad \Lambda = \lambda_1 + \lambda_2 + \lambda_3 \tag{1}$$

subject to

$$\lambda_1 = \sum_{i,j \in \Omega^l} L_{i,j} \cdot 2 \cdot \Psi_{i,j}^{line} \cdot \alpha_{i,j} \qquad \forall i,j \in \Omega^l \tag{2}$$

$$\lambda_2 = \sum_{i,j \in \Omega^l} R_{i,j} \cdot I_{i,j}^2 \cdot S^{base} \cdot \Psi_{i,j}^{loss} \qquad \forall i,j \in \Omega^l \tag{3}$$

$$\lambda_3 = \sum_{i \in \Omega^s} (\Psi_i^{trafo} \cdot \tau_i + \Psi_i^{subs}) \qquad \forall i \in \Omega^s \tag{4}$$

Equation (1) is a multi-objective function with a goal to minimize the total cost of the power distribution system considering the investment cost of lines, operational cost of the system and the investment cost of substations.

Equation (2) defines the investment cost of lines $\lambda_1$. It is calculated by multiplication of the length of the lines $L_{i,j}$ and the cost of the chosen cable $\Psi_{i,j}^{line}$. The cost minimization for $\lambda_1$ as a variable is achieved by selecting the most optimal line configuration considering the line investment costs. A binary variable $\alpha_{i,j}$ is used to assure that $I_{i,j}$ is considered only for lines selected in the solution, and set $I_{i,j} = 0$ when a line is not selected as part of the solution. Since the installation cost for underground cable installations can amount to 50 % of the total cost of the line, the cost of the cable is multiplied by a factor of two [16]. This way the optimization model considers both cable and installation costs.

The operational cost of the system $\lambda_2$ is defined with equation (3). Similarly to (2), the cost minimization is achieved by selecting the most optimal line configuration. This parameter is the operational cost of the system and is a direct representation of the cost of losses. The cost of the losses $\Psi_{i,j}^{loss}$ is estimated to equal the market cost of electricity taken from EMA.

In (4), the investment cost for the substations is calculated. The parameter $\Psi_i^{trafo}$ is the cost of a single transformer. A single type of transformer and sizing is chosen according to reports of the Singaporean DSO, matching the transformers installed in Singaporean power system [6]. An integer variable $\tau_i$ is defined to represent the number of transformers at each substation node.

The cost minimization of $\lambda_3$ is achieved by optimizing the allocation of each load to an adequate substation by analyzing different line configurations with respect to the installed transformer capacity. The obtained solution provides a minimum number of transformers $\tau_i$ for each substation node $i \in \Omega^s$. In this way the installed transformer capacity will be matched to the total demand of the system in a cost efficient manner. $\Psi_i^{subs}$ is another cost parameter which is fixed and includes the estimated cost of the switchgear equipment per substation. Although $\Psi_i^{subs}$ is a fixed parameter, it is included in the substation investment cost and the overall system cost so that the objective function and all of the objectives are weighted closer to the actual cost of the existing distribution system.

### *B. Annual Multi-Objective Weighting*

The cost minimization objective given in (1) is a multi-objective function represented as a sum of three different cost parameters (2)-(4). When the total sum of the investment costs of the lines and the substations are considered in $S\$$, the ratio between these two parameters and the variable operational cost of the system will be incomparable due to the time span at which the cost in per $S\$$ is calculated. The investment cost of lines and substations is calculated in $S\$$ for a period of 40 and 35 years respectively, while the operational cost of the system is calculated in $S\$$ for a period of one hour. In such case, the cost of the losses would be so small, that it can be easily neglected as irrelevant. For successful implementation of the proposed optimization, a certain methodology to equally weight the three cost parameters is implemented. Annual time period and a mutual cost unit of $S\$/year$ for all cost parameters are introduced. In this way, all three cost parameters will be equally considered when obtaining the grid topology and the loads allocation.

*1) Annual Investment Cost of Lines:*

$$\lambda_1^* = \Xi^{line} \cdot \lambda_1 = \frac{u \cdot (1+u)^{t^{line}}}{(1+u)^{t^{line}} - 1} \cdot \lambda_1 \tag{5}$$

The total annual investment cost of lines is calculated using the *annual equivalent-worth criterion* [17]. The investment cost parameter of the lines $\lambda_1$ is replaced with the *equivalent uniform annual worth* parameter $\lambda_1^*$ by using the *capital-recovery factor* $\Xi^{line}$ defined in (5).

*2) Annual Operational Cost of the System:*

$$\lambda_2^* = EPLT \cdot \lambda_2 = \frac{\frac{\sum_1^{365}(P_{tot}^d)^2}{365}}{max\{(P_{tot}^d)^2\}} \cdot 8760 \cdot \lambda_2 \tag{6}$$

The power distribution system is designed such that it meets the peak demand of the system. However, in normal daily operation of the system, peak loads are seldom reached. The cost of the losses calculated when only peak demand is considered is not an accurate cost consideration. According to statistics for half-hourly system demand data in Singapore, the system demand is within an interval of $5\%$ of the peak demand for only $9.22\%$ of the time. For a more accurate definition of the annual cost of the losses, an *equivalent peak loss time* factor (EPLT) is introduced [15]. The cost calculation of the annual variable cost of the distribution system is defined in equation (6).

*3) Annual Cost of Substations:*

$$\lambda_3^* = \Xi^{subs} \cdot \lambda_3 = \frac{u \cdot (1+u)^{t^{subs}}}{(1+u)^{t^{subs}} - 1} \cdot \lambda_3 \tag{7}$$

Similarly to (5), the cost of the substations $\lambda_3$ is reformulated to an annual cost of substations $\lambda_3^*$ using the *capital-recovery factors* $\Xi^{subs}$ in (7).

*4) Weighted Multi-objective Function:*

$$\min \qquad \Lambda^* = \lambda_1^* + \lambda_2^* + \lambda_3^* \tag{8}$$

Having the multiple objective parameters proportionally weighted, the objective function (1) is redefined in $S\$/year$ as shown in equation (8).

## III. Mixed-Integer Conic Load Flow Constraints

It is necessary that the synthetic grid generated with the proposed model satisfies the electrical constraints of the system. An AC power flow that constrains the node voltage and limits the line capacity is defined.

For simplified and more convenient notation, the following squared variables are introduced:

$$\upsilon_i = V_i^2 \qquad \text{and} \qquad \ell_{i,j} = I_{i,j}^2.$$

$$\sum_{j,i\in\Omega^l} P_{j,i} - \sum_{i,j\in\Omega^l} P_{i,j} - \sum_{i,j\in\Omega^l} [R_{i,j}\cdot \ell_{i,j}] + P_i^g = P_i^d \quad \forall i \in \Omega^n \quad (9)$$

$$\sum_{j,i\in\Omega^l} Q_{j,i} - \sum_{i,j\in\Omega^l} Q_{i,j} - \sum_{i,j\in\Omega^l} [X_{i,j}\cdot \ell_{i,j}] + Q_i^g = Q_i^d \quad \forall i \in \Omega^n \quad (10)$$

$$\upsilon_i - \upsilon_j = \omega_{i,j} + 2\cdot R_{i,j}\cdot P_{i,j} + 2\cdot X_{i,j}\cdot Q_{i,j} + Z_{i,j}^2\cdot \ell_{i,j} \quad \forall i,j \in \Omega^l \quad (11)$$

$$|\omega_{i,j}| \le \big((V^{max})^2 - (V^{min})^2\big)\cdot(1-\alpha_{i,j}) \quad \forall i,j \in \Omega^l \quad (12)$$

$$\upsilon_j \cdot \ell_{i,j} = P_{i,j}^2 + Q_{i,j}^2 \quad \forall i,j \in \Omega^l \quad (13)$$

$$|P_{i,j}| \le V^{max}\cdot I_{i,j}^{max}\cdot \alpha_{i,j} \quad \forall i,j \in \Omega^l \quad (14)$$

$$|Q_{i,j}| \le V^{max}\cdot I_{i,j}^{max}\cdot \alpha_{i,j} \quad \forall i,j \in \Omega^l \quad (15)$$

$$0 \le \ell_{i,j} \le (I_{i,j}^{max})^2\cdot \alpha_{i,j} \quad \forall i,j \in \Omega^l \quad (16)$$

$$(V^{min})^2 \le \upsilon_i \le (V^{max})^2 \quad \forall i \in \Omega^n \quad (17)$$

Equations (9) and (10) define a balanced active and reactive power flow. These constraints assure that the power demand is supplied to each load node. Equation (11) is a function of the current magnitude and branch parameters along with the active and reactive power flow, which determines the voltage drop across connected lines. In equation (12), an auxiliary variable $\omega$ is defined and assigned to zero when the line *ij* is connected. When line *ij* is not connected, $\omega$ can get any other value within the defined limits, to satisfy (11).

Equation (13) is used to calculate the power flow parameters of the lines. Constraints (14) and (15) set the boundaries of the active and reactive power flow. These are considered and calculated only when a line is connected, otherwise the active and reactive power flow parameters are assigned to zero. Similarly, equation (16) sets the limits of the current flowing in the connected lines. Constraint (17) defines voltage limits for each node respectively.

$$\ell_{i,j}\cdot \upsilon_i \ge P_{i,j}^2 + Q_{i,j}^2, \quad P_{i,j} \ge 0 \quad \forall i,j \in \Omega^l \quad (18)$$

Equation (13) is a non-linear quadratic equation. In order to make the optimization model convex, the quadratic equation (13) is replaced with a second order conic relaxation equation (18), assuring convergence to optimality.

$$\sum_{i,j\in\Omega^l} \alpha_{i,j} = n^n - n^s \quad \forall i,j \in \Omega^l \quad (19)$$

Equation (19) is defined to address the radial operation of the network. However, this condition alone is not enough to guarantee the network's radiality. Together with power flow balance constraints (9) and (10) both conditions are met and network radiality is guaranteed [18].

$$P_i^{gmin} \le P_i^g \le P_i^{gmax}\cdot \tau_i \quad \forall i \in \Omega^n \quad (20)$$

$$Q_i^{gmin} \le Q_i^g \le Q_i^{gmax}\cdot \tau_i \quad \forall i \in \Omega^n \quad (21)$$

$$1 \le \tau_i \le N^{Tmax} \quad \forall i \in \Omega^s \quad (22)$$

Equations (20) - (21) define the limit of power input into the grid with respect to the installed transformer capacity defined by the number of transformers selected in each substation node $i \in \Omega^s$. Equation (22) defines the bounds of the integer variable $\tau_i$ and sets the maximum possible number of transformers at a single substation node.

The model for synthetic grid generation is defined by the objective function (8) subject to the constraints in equations (2) - (7), (9) - (12) and (14) - (22).

## IV. Case Study and Results

A case study for the Singaporean power distribution system at $22kV$ is defined. In Singapore, consumer data for the end user at low voltage per zip code is available, with each zip code corresponding to a single building. The consumer data at $0.4kV$ and $6.6kV$ is then clustered and allocated to substations at power distribution level of $22kV$. The bottom up approach used for clustering of the end users data to a $22kV$ distribution level is elaborated and provided by the authors in [19]. Each $22/0.4$ $kV/kV$ and $22/6.6$ $kV/kV$ substation represents a load in the synthetic grid generation model.

Based on reports from the Singaporean DSO and TSO, a transformer capacity of 75 $MVA$ for the $66/22$ $kV/kV$ substations at the source nodes is considered [6].

The power system in Singapore includes only underground cable installations. In the synthetic grid model, XLPE copper cable with a cross section of $500mm^2$ is considered. Characteristic of the cable are given in Table I.

TABLE I
500 $mm^2$ XLPE Copper Cable Data

| Cross Section | Resistance ($R$) | Reactance ($X$) | Current limit ($I_{max}$) |
|---|---|---|---|
| $mm^2$ | $\Omega/km$ | $\Omega/km$ | $A$ |
| 500 | 0.0468 | 0.0990 | 600 |

The synthetic distribution grid is generated by considering a reserve in capacity of lines and substations, in order to considered the future growth of the peak demand of the system. The reserve is considered for a period of 10 years, considering annual peak demand growth of 1.80% [20]. The total reserve is calculated as follows

$$reserve = (1.018^{10} - 1)\cdot 100 = 19.53\% \quad (23)$$

The current limit of the cable and the transformer capacity are reduced by the amount of the calculated capacity reserve. The cable current limit and the transformer $MVA$ capacity are reduced to $483A$ and $60.35MVA$ respectively. The voltage limits are set to $\pm 6\%$. A lifespan of forty and thirty-five years are considered for lines and substations respectively.

When power distribution systems are designed, not all possible connections between all nodes are taken into account. Due to economic and technical reasons, it is most common that the potential set of line connections considers routes between neighboring nodes, while distant nodes with respect to the observed node are eliminated [15]. This approach is used to define the set of potential lines in the synthetic grid generation model.

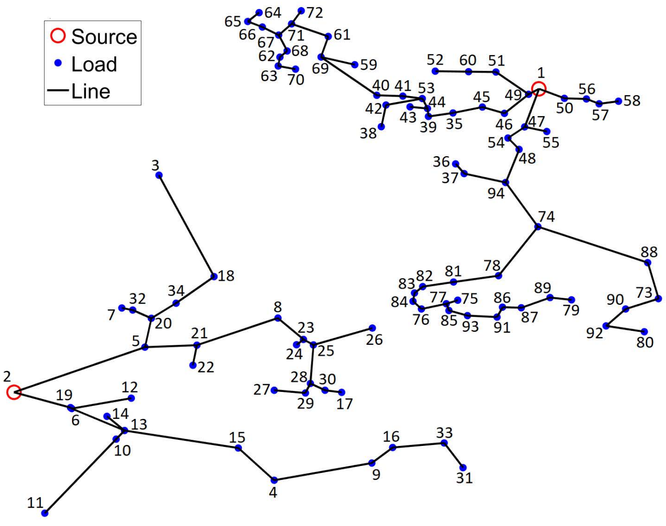


Fig. 1. Line configuration for a 94 bus sample of case study.

TABLE II
LINE DATA FOR SOURCE NODE 2 LINE CONFIGURATION

| Lines | | R | X | Load at End Nodes | |
|---|---|---|---|---|---|
| Initial Node | End Node | (Ω) | (Ω) | (MW) | (MVAR) |
| 2 | 5 | 0.082 | 0.173 | 4.250 | 1.397 |
| 2 | 19 | 0.033 | 0.070 | 0.076 | 0.025 |
| 6 | 12 | 0.033 | 0.069 | 0.097 | 0.032 |
| 20 | 32 | 0.013 | 0.027 | 0.075 | 0.025 |
| 5 | 20 | 0.030 | 0.063 | 0.065 | 0.021 |
| 20 | 34 | 0.020 | 0.042 | 0.064 | 0.021 |
| 13 | 10 | 0.010 | 0.021 | 0.064 | 0.021 |
| 6 | 13 | 0.035 | 0.075 | 0.075 | 0.024 |
| 13 | 14 | 0.017 | 0.036 | 0.077 | 0.025 |
| 13 | 15 | 0.061 | 0.130 | 0.100 | 0.033 |
| 13 | 11 | 0.094 | 0.199 | 0.065 | 0.021 |
| 5 | 21 | 0.027 | 0.057 | 0.064 | 0.021 |
| 21 | 22 | 0.021 | 0.044 | 0.065 | 0.021 |
| 21 | 8 | 0.050 | 0.106 | 2.169 | 0.713 |
| 32 | 7 | 0.006 | 0.013 | 1.697 | 0.558 |
| 19 | 6 | 0.001 | 0.003 | 0.731 | 0.240 |
| 34 | 18 | 0.034 | 0.071 | 0.055 | 0.018 |
| 23 | 24 | 0.006 | 0.014 | 0.072 | 0.024 |
| 8 | 23 | 0.026 | 0.054 | 0.063 | 0.021 |
| 23 | 25 | 0.008 | 0.016 | 0.069 | 0.023 |
| 18 | 3 | 0.107 | 0.227 | 0.069 | 0.023 |
| 25 | 28 | 0.039 | 0.083 | 0.063 | 0.021 |
| 25 | 26 | 0.035 | 0.074 | 0.077 | 0.025 |
| 28 | 29 | 0.010 | 0.021 | 0.066 | 0.022 |
| 28 | 30 | 0.010 | 0.022 | 0.075 | 0.025 |
| 29 | 27 | 0.016 | 0.035 | 0.056 | 0.018 |
| 30 | 17 | 0.009 | 0.019 | 0.075 | 0.025 |
| 15 | 4 | 0.038 | 0.080 | 0.078 | 0.025 |
| 9 | 16 | 0.019 | 0.041 | 0.120 | 0.039 |
| 16 | 33 | 0.027 | 0.057 | 0.065 | 0.021 |
| 4 | 9 | 0.053 | 0.113 | 3.394 | 1.115 |
| 33 | 31 | 0.027 | 0.057 | 0.056 | 0.018 |
| | | | **Total Load:** | **14.184** | **4.662** |

*Node 2 is a source node with no load
Note: Single phase load data.

TABLE III
LINE DATA FOR SOURCE NODE 1 LINE CONFIGURATION

| Lines | | R | X | Load at End Nodes | |
|---|---|---|---|---|---|
| Initial Node | End Node | (Ω) | (Ω) | (MW) | (MVAR) |
| 1 | 49 | 0.008 | 0.016 | 0.074 | 0.024 |
| 1 | 50 | 0.016 | 0.035 | 0.077 | 0.025 |
| 1 | 47 | 0.040 | 0.084 | 0.069 | 0.023 |
| 49 | 51 | 0.028 | 0.060 | 0.069 | 0.023 |
| 49 | 46 | 0.023 | 0.049 | 0.074 | 0.024 |
| 50 | 56 | 0.011 | 0.024 | 0.069 | 0.023 |
| 51 | 60 | 0.014 | 0.030 | 0.069 | 0.023 |
| 56 | 57 | 0.008 | 0.017 | 0.068 | 0.023 |
| 60 | 52 | 0.017 | 0.037 | 0.069 | 0.023 |
| 57 | 58 | 0.010 | 0.022 | 4.250 | 1.397 |
| 94 | 37 | 0.023 | 0.049 | 0.100 | 0.033 |
| 48 | 94 | 0.035 | 0.073 | 0.062 | 0.020 |
| 94 | 74 | 0.048 | 0.102 | 2.057 | 0.676 |
| 39 | 44 | 0.008 | 0.017 | 0.068 | 0.023 |
| 44 | 43 | 0.009 | 0.020 | 0.085 | 0.028 |
| 44 | 53 | 0.011 | 0.022 | 0.069 | 0.023 |
| 37 | 36 | 0.011 | 0.023 | 0.056 | 0.018 |
| 88 | 73 | 0.037 | 0.079 | 4.250 | 1.397 |
| 74 | 88 | 0.068 | 0.143 | 0.067 | 0.022 |
| 35 | 39 | 0.013 | 0.028 | 0.092 | 0.030 |
| 63 | 70 | 0.009 | 0.020 | 0.077 | 0.025 |
| 86 | 87 | 0.010 | 0.020 | 0.080 | 0.026 |
| 87 | 89 | 0.018 | 0.038 | 0.069 | 0.023 |
| 45 | 35 | 0.016 | 0.035 | 1.264 | 0.415 |
| 53 | 42 | 0.020 | 0.042 | 0.068 | 0.022 |
| 42 | 38 | 0.022 | 0.047 | 0.099 | 0.033 |
| 53 | 41 | 0.010 | 0.022 | 0.092 | 0.030 |
| 62 | 63 | 0.009 | 0.019 | 0.092 | 0.030 |
| 91 | 86 | 0.011 | 0.022 | 0.069 | 0.023 |
| 89 | 79 | 0.011 | 0.024 | 0.056 | 0.018 |
| 46 | 45 | 0.013 | 0.028 | 0.097 | 0.032 |
| 41 | 40 | 0.013 | 0.028 | 0.067 | 0.022 |
| 40 | 69 | 0.049 | 0.103 | 0.062 | 0.020 |
| 68 | 62 | 0.007 | 0.016 | 0.091 | 0.030 |
| 74 | 78 | 0.054 | 0.114 | 0.075 | 0.025 |
| 78 | 81 | 0.024 | 0.051 | 0.077 | 0.025 |
| 93 | 91 | 0.015 | 0.032 | 0.069 | 0.023 |
| 67 | 68 | 0.017 | 0.036 | 0.069 | 0.023 |
| 85 | 93 | 0.011 | 0.023 | 0.069 | 0.023 |
| 90 | 92 | 0.020 | 0.042 | 0.074 | 0.024 |
| 73 | 90 | 0.020 | 0.042 | 0.068 | 0.022 |
| 47 | 54 | 0.014 | 0.030 | 0.069 | 0.023 |
| 54 | 48 | 0.013 | 0.027 | 0.090 | 0.029 |
| 69 | 61 | 0.021 | 0.045 | 2.263 | 0.744 |
| 61 | 71 | 0.023 | 0.049 | 0.069 | 0.023 |
| 67 | 66 | 0.012 | 0.025 | 0.068 | 0.022 |
| 71 | 67 | 0.013 | 0.028 | 0.092 | 0.030 |
| 77 | 85 | 0.007 | 0.015 | 0.069 | 0.023 |
| 92 | 80 | 0.021 | 0.044 | 0.069 | 0.023 |
| 47 | 55 | 0.012 | 0.026 | 0.069 | 0.023 |
| 69 | 59 | 0.019 | 0.040 | 0.060 | 0.020 |
| 66 | 65 | 0.009 | 0.020 | 0.080 | 0.026 |
| 71 | 72 | 0.014 | 0.030 | 0.069 | 0.023 |
| 77 | 75 | 0.007 | 0.015 | 0.069 | 0.023 |
| 76 | 77 | 0.014 | 0.030 | 0.077 | 0.025 |
| 65 | 64 | 0.011 | 0.023 | 0.057 | 0.019 |
| 84 | 76 | 0.009 | 0.019 | 0.069 | 0.023 |
| 83 | 84 | 0.009 | 0.018 | 0.074 | 0.024 |
| 82 | 83 | 0.008 | 0.016 | 0.077 | 0.025 |
| 81 | 82 | 0.017 | 0.035 | 0.069 | 0.023 |
| | | | **Total Load:** | **18.130** | **5.959** |

*Node 1 is a source node with no load
Note: Single phase load data.

The grid topology and load allocation of the synthetic grid generated for a 94 bus sample from the case study data is shown on Fig. 1. The sample data includes two substations and represents one of the forty-two planning areas in Singapore.

The loads are assigned to the two substations, defining two distribution feeder configurations detailed in Tables II and III. Using MATPOWER, an AC power flow is run to confirm and assure that the voltage and the line current limits are within the permissible boundaries.

The synthetic grid generation model is programmed in GAMS using CPLEX as a solver.

## V. Validation

TABLE IV
Samples of the Singaporean Synthetic Distribution System

| ID | Number of nodes | Number of lines | Average node degree |
|---|---|---|---|
| 1* | 61 | 60 | 1.967 |
| 2** | 33 | 32 | 1.939 |
| 3 | 113 | 112 | 1.982 |
| 4 | 279 | 278 | 1.993 |
| 5 | 24 | 23 | 1.917 |

* source node 1 shown in Table III
** source node 2 shown in Table II

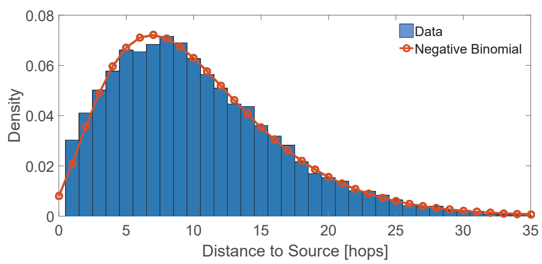


Fig. 2. Distribution of hop distances and the Negative Binomial fit.

TABLE V
Length Analysis of the Samples of the Singaporean Synthetic Distribution System

| ID | Characteristic Path Length | Average Line Length | Normalized Characteristic Path Length |
|---|---|---|---|
| 1* | 5.354 | 0.382 | 14.019 |
| 2** | 4.748 | 0.666 | 7.131 |
| 3 | 4.624 | 0.271 | 17.044 |
| 4 | 10.333 | 0.436 | 23.695 |
| 5 | 2.780 | 0.561 | 4.957 |

CNS is used to describe the generated synthetic grid using statistical parameters which are then compared to a real distribution system in The Netherlands [2], [14]. Table IV shows the calculated average node degree for selected samples of the synthetic grid, which correlate to the average degree for radial configurations given in [14]. The total average node degree for the Singaporean synthetic grid is $1.954$. As described in [2], the hop distance to the source node shown in Figure 2 follows a Negative Binomial distribution. Considering the normalized characteristic path length given in Table V, the Singaporean synthetic grid corresponds to the data in [14]. Even though the statistical parameters for the Singaporean synthetic grid are not directly compared to the Singaporean real power system due to lack of data, the synthetic grid satisfies the general statistical patterns for power distribution systems.

## VI. Conclusion

A novel approach for generating synthetic distribution grid that corresponds to a particular existing power distribution system is proposed. An power system planning optimization with annually weighted cost multi-objective function is introduced. Using publicly available data, the generated synthetic distribution grid has a high resemblance to the actual grid. A convex AC radial power flow assures that all power system constraints are satisfied. Future peak demand growth of the power system is adequately introduced. A case study of the Singaporean power distribution system is used and validated using general statistical patterns derived from a real distribution system in The Netherlands.

Further research on synthetic grids with different distribution grid configurations such as ring configuration is to be conveyed. The definition of the set of potential lines including multiple geographical constraints and more realistic line routes can be further improved and investigated.


## Acknowledgment

This work was financially supported by the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme.


## References


[1] M. C. Falvo, L. Martirano, D. Sbordone, and E. Bocci, "Technologies for smart grids: A brief review," in *2013 12th Int. Conf. on Envir. and Elect. Eng.*, May 2013, pp. 369–375.

[2] E. Schweitzer, A. Scaglione, A. Monti, and G. A. Pagani, "Automated generation algorithm for synthetic medium voltage radial distribution systems," *IEEE J. on Emer. and Select. Topics in Circ. and Syst.*, vol. 7, no. 2, pp. 271–284, June 2017.

[3] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating statistically correct random topologies for testing smart grid communication and control networks," *IEEE Trans. on Smart Grid*, vol. 1, no. 1, June 2010.

[4] S. H. Elyas and Z. Wang, "Improved synthetic power grid modeling with correlated bus type assignments," *IEEE Trans. on Power Syst.*, vol. 32, no. 5, pp. 3391–3402, Sept 2017.

[5] S. Abeysinghe, J. Wu, M. Sooriyabandara, M. Abeysekera, T. Xu, and C. Wang, "Topological properties of medium voltage electricity distribution networks," *Applied Energy*, vol. 210, pp. 1101 – 1112, 2018.

[6] Singapore Power Limited, "Singapore Power Annual Report 2008/09," Tech. Rep., 2008.

[7] Energy Market Authority, "Singapore Energy Statistics 2016," 2016.

[8] S. Ganguly, N. C. Sahoo, and D. Das, "Recent advances on power distribution system planning: a state-of-the-art survey," *Energy Systems*, vol. 4, no. 2, pp. 165–193, Jun 2013.

[9] P. S. Georgilakis and N. D. Hatziargyriou, "A review of power distribution planning in the modern power systems era: Models, methods and future research," *Elect. Power Syst. Res.h*, vol. 121, pp. 89 – 100, 2015.

[10] R. G. Cespedes, "New method for the analysis of distribution networks," *IEEE Trans. on Power Deliv.*, vol. 5, no. 1, pp. 391–396, Jan 1990.

[11] M. C. O. Borges, J. F. Franco, and M. J. Rider, "Optimal reconfiguration of electrical distribution systems using mathematical programming," *J. of Cont., Autom. and Elect. Syst.*, vol. 25, no. 1, pp. 103–111, 2014.

[12] J. C. López, M. Lavorato, and M. J. Rider, "Optimal reconfiguration of electrical distribution systems considering reliability indices improvement," *Int. J. of Elect. Power and Energy Syst.*, vol. 78, 2016.

[13] A. B. Birchfield, E. Schweitzer, M. H. Athari, T. Xu, T. J. Overbye, A. Scaglione, and Z. Wang, "A metric-based validation process to assess the realism of synthetic power grids," *Energies*, vol. 10, no. 8, 2017.

[14] G. A. Pagani and M. Aiello, "Towards decentralization: A topological investigation of the medium and low voltage grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 538–547, Sept 2011.

[15] C. R. Bayliss and B. J. Hardy, *Transmission and Distribution Electrical Engineering*, 4th ed. Newnes, Elsevier, 2012.

[16] K. Hall and H. E. Consulting, "Out of Sight, Out of Mind," Edison Electric Institute (EEI), Tech. Rep. January, 2013.

[17] Chan S. Park, *Contemporary Engineering Economics*, 5th ed. Prentice Hal, 2010.

[18] M. Lavorato, S. Member, J. F. Franco, S. Member, and M. J. Rider, "Imposing Radiality Constraints in Distribution System Optimization Problems," *IEEE Trans. on Power Syst.*, vol. 27, no. 1, 2012.

[19] D. Ciechanowicz, D. Pelzer, B. Bartenschlager, and A. Knoll, "A modular power system planning and power flow simulation framework for generating and evaluating power network models," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2214–2224, May 2017.

[20] EMA, "Singapore Electricity Market Outlook," 2016.