

# An Observability Index for Distribution Networks using Information Entropy

Bernd Brinkmann, *Student Member, IEEE*, Michael Negnevitsky,  
*Senior Member, IEEE*, and Timothy Yee  
 School of Engineering and ICT  
 University of Tasmania, Australia  
 Michael.Negnevitsky@utas.edu.au

Thanh Nguyen, *Member, IEEE*  
 TasNetworks  
 Moonah TAS 7009  
 Australia  
 Thanh.Nguyen@tasnetworks.com.au

**Abstract**—Information about the accuracy of state estimation results in distribution networks is crucial for an effective and safe network management. This paper explores a new method of quantifying the uncertainty of a state estimation result by using information entropy as an index for observability. This index has the ability to represent the network observability as a single continuous number, making it possible for the network operator to objectively evaluate the reliability of a state estimation result. This paper also investigates the changes in observability due to load variations and demonstrates how an optimal meter placement method based on the proposed observability index can be implemented. The proposed method has been tested via simulations on a modified IEEE 34 Bus Test Feeder and was compared with an existing method.

**Index Terms**— state estimation, observability, information entropy, optimal meter placement

## I. INTRODUCTION

The increasing amount of distributed generation such as wind and solar in distribution networks can lead to significant voltage and safety concerns. One solution to this problem is to actively manage distribution networks [1-3]. For an effective and safe network management accurate information about the network state is crucial. The state of an electrical network is defined by the voltage magnitude and angle at every bus throughout the network. State estimation (SE) is used to calculate the network state from an available set of measurements and their positions in the network [4, 5]. Whether the network state can be estimated from a given set of measurements is expressed by the network observability. Traditionally a network is classified as either observable if the SE can provide an estimate of the network state or unobservable if it cannot [6-8].

This binary classification works well in transmission networks where generally numerous real time measurements are available. Distribution networks usually do not have enough measurements for the SE to be performed and therefore are often classified as unobservable [9]. Thus, pseudo-measurements are used in the absence of real time measurements to make a network observable [8]. Pseudo-

measurements are load forecasts generated from historical data. They typically have large margins of error, which reflect the inability to accurately predict loads from historical data [10]. It is important to understand the effect of these margins of error on the SE results. If the margins of error are large it is possible that the estimated network state is significantly different from the actual network state. Therefore, even if the network is found to be observable by adding pseudo measurements, the calculated state may contain a significant amount of uncertainty and can therefore be unreliable [11, 12]. The uncertainty of a SE result is the inability of the SE to calculate the actual network state due to unavoidable measurement errors. The traditional observability classification is based on the available set of measurements and their positions in the network [13]. This approach does not provide any information about the accuracy of the state estimation result.

To overcome this limitation an observability index that quantifies the network observability from the probability distribution of the state variables by employing information entropy is proposed in this paper. This makes it possible to quantify the observability as a continuous number. The proposed method provides more information about the estimated state than previous methods. It is also shown that the observability can change due to load variations. This is illustrated by a simulation over a 24 hour period using residential and commercial load profiles. Furthermore this paper demonstrates how an optimal meter placement method based on the proposed observability index can be implemented. The proposed methods are verified by simulations on a modified IEEE 34 Bus Test Feeder [14].

The rest of the paper is organized as follows. Section II defines entropy as an index for observability. It also points out important characteristics of entropy and gives a simple example. Section III explains how the new observability index is calculated. In Section IV, a case study is provided to test the proposed methods and Section V concludes the paper.

## II. PROPOSED ALGORITHM

### A. Observability Index

In this section an observability index based on information entropy which quantifies the uncertainty of an estimated network state as a continuous number is introduced.

A continuous observability index should have two extremes, where one extreme represents full observability, corresponding to a single network state that is known to be true. The other extreme should imply that all possible network states have an equal probability of being the true network state, leaving the network entirely undetermined. One such index is information entropy which is a measure of the uncertainty of a discrete random variable [15, 16]. Therefore, the proposed network observability index is defined as the sum of the entropy of the state variables in a network.

The entropy of the  $i$ -th state variable  $x^i$  is calculated by

$$H(x^i) = - \sum_{n=1}^{N_i} P(x_n^i) \log_2 P(x_n^i) \quad (1)$$

Where  $x$  is the vector of state variables;  $x^i$  denotes a specific states variable for  $i = 1, 2, \dots, I$ ;  $I$  is the number of all state variables in the network;  $n$  is a possible state variable value for  $n = 1, 2, \dots, N_i$ ;  $N_i$  is the number of possible values of the state variable  $x^i$ ;  $P(x_n^i)$  is the probability of the possible value  $n$  of state variable  $x^i$  being true. Hence,  $H(x^i)$  quantifies the uncertainty of the state variable  $x^i$  by taking every possible state variable value and its probability of being true into account.

$H(x^i)$  is a monotonic increasing function of  $N_i$

$$H(x^i)_{N_i} \leq H(x^i)_{N_i+1} \quad (2)$$

The network observability index  $K$  is then calculated by

$$K = \sum_{i=1}^I H(x^i). \quad (3)$$

The range of  $K$  is given by

$$0 \leq K \leq \sum_{i=1}^I \log_2(N_i). \quad (4)$$

If  $K = 0$ , the network is fully observable. On the other hand,  $K = \sum_{i=1}^I \log_2(N_i)$  implies no observability at all. In general, a low value of  $K$  corresponds to a low uncertainty of the network state and therefore, a highly observable network and vice versa.

$K$  is a relative number that can only be compared with the observability value of different load or metering configurations of the same network.

A state variable is a continuous number and therefore, can take an infinite number of possible values. In order to reduce the infinite number of possible values to a finite number, the state variables have to be "binned" before the observability index can be calculated. Each "bin" represents a range of possible values within the minimum and maximum values defined beforehand. The bin size has to be small enough to accurately represent the probability distribution of the state variables. However, if the bin size is too small more Monte Carlo simulations have to be performed in order to deliver accurate results. Therefore, the bin size has to be chosen as a tradeoff between computation time and accuracy.

### B. Simple example

To show how the observability value of a network is calculated, let us consider a three bus network with buses A, B and C. The network state is defined by the voltage amplitudes at all buses. Let us assume that each bus has two possible state values ("bins") namely V1 and V2. Let us also assume that the probability of a state being equal to V1 or V2 has been determined by an appropriate method as discussed in the section III. The probability values are shown in Table I.

TABLE I: Example data

State	Bus	Probability	Entropy (H)
V <sub>1</sub>	A	Pr(V <sub>A</sub> =V <sub>1</sub> )=0.8	H(V <sub>A</sub> =V <sub>1</sub> )=0.26
	B	Pr(V <sub>B</sub> =V <sub>1</sub> )=0.7	H(V <sub>B</sub> =V <sub>1</sub> )=0.36
	C	Pr(V <sub>C</sub> =V <sub>1</sub> )=0.4	H(V <sub>C</sub> =V <sub>1</sub> )=0.53
V <sub>2</sub>	A	Pr(V <sub>A</sub> =V <sub>2</sub> )=0.2	H(V <sub>A</sub> =V <sub>2</sub> )=0.46
	B	Pr(V <sub>B</sub> =V <sub>2</sub> )=0.3	H(V <sub>B</sub> =V <sub>2</sub> )=0.52
	C	Pr(V <sub>C</sub> =V <sub>2</sub> )=0.6	H(V <sub>C</sub> =V <sub>2</sub> )=0.44
			$K = 2.57$

According to (1), bus A has a total entropy value of  $H_A = 0.26+0.46 = 0.72$ ; bus B a total value of  $H_B = 0.36+0.52 = 0.88$ ; and bus C a total value of  $H_C = 0.53+0.44=0.97$ . The network observability is then equal to  $K = H_A+H_B+H_C = 0.72+0.88+0.97 = 2.57$  according to (3). Since the state of the network is defined by three state variables and each state variable has two possible values the observability value  $K$  can range from 0 to 3, according to (4).

### C. Probability Distribution of the State Variables

In order to calculate the proposed observability index, the probability distribution of the state variables needs to be obtained.

In this paper the probabilities of the network state variables is calculated by employing a Monte Carlo simulation. The Monte Carlo simulation was chosen because it is a well-known method and delivers reliable results with a high degree of accuracy. This accuracy comes at the expense of a

high computation time. The Monte Carlo simulation used is described in detail in section III. For practical applications a Monte Carlo simulation might not be the best method to determine the probability distribution of the state variables. An appropriate method has to be chosen depending on the network topology, measurement configuration, the desired accuracy and computational time constrains. It should however be known, that the proposed observability index is *independent* of the method chosen to calculate the probability distribution of the state variables.

### III. MONTE CARLO SIMULATION

#### A. Calculating a possible network state

Although the state of the network  $\mathbf{x}$  is defined by the voltage magnitude and angle at every bus, the principal concern of a distribution network operator is the voltage magnitude [17]. For this reason only voltage magnitudes are used to calculate the proposed observability index.

In the Monte Carlo simulation, a possible measurement vector  $\mathbf{z}^k$  has to be calculated first. The possible measurement vector  $\mathbf{z}^k$  deviates from the actual measurements in  $\mathbf{z}$  by

$$\mathbf{z}^k = \mathbf{z} - \mathbf{e}^k. \quad (5)$$

where  $\mathbf{e}^k$  is an independent random Gaussian zero-mean error vector  $\mathbf{e}^k = \mathbf{e}_1^k, \mathbf{e}_2^k, \dots, \mathbf{e}_m^k$ ; and  $m$  is the total number of measurements in the network;  $\mathbf{e}^k$  is generated as a normally distributed random number vector according to the mean values and standard deviations of the corresponding network measurements.

The dimension of  $\mathbf{z}^k$  can be *greater* than the dimension of  $\mathbf{x}^k$  due to the use of redundant measurements. Therefore, the dimension of  $\mathbf{z}^k$  needs to be reduced to match the dimension of  $\mathbf{x}^k$  to calculate the possible network state using a power flow analysis.

Reducing the dimension of  $\mathbf{z}^k$  is done by extracting  $\mathbf{z}_{pf}^k$  from  $\mathbf{z}^k$ ; where  $\mathbf{z}_{pf}^k$  is a subset of  $\mathbf{z}^k$

$$(\mathbf{z}_{npf}^k \cup \mathbf{z}_{pf}^k = \mathbf{z}^k). \quad (6)$$

$\mathbf{z}_{pf}^k$  only contains the following measurements:

- Real and reactive power at load busses;
- Voltage magnitude and angle measurements at the slack bus;
- Voltage magnitude and real power at generation busses.

$\mathbf{z}_{npf}^k$  is made up of all other measurements that are not used in the power flow solution.

The possible state  $\mathbf{x}^k$  can now be determined by the Newton–Raphson iterative method.

$$\mathbf{x}_{i+1}^k = \mathbf{x}_i^k + \mathbf{J}^{-1}(\mathbf{x}_i^k)[\mathbf{z}_{pf}^k - \mathbf{h}_{pf}(\mathbf{x}_i^k)] \quad (7)$$

$$\mathbf{h}_{pf}(\mathbf{x}_i^k) = \mathbf{z}_{pf}^k - \mathbf{r}_i^k \quad (8)$$

where  $\mathbf{h}_{pf}(\mathbf{x})$  is the non-linear measurement function that relates the state variables to the measurement in  $\mathbf{z}_{pf}^k$  (8);  $\mathbf{r}^k$  is the vector of residuals;  $\mathbf{J}(\mathbf{x})$  is the Jacobian of the non-linear measurement function  $\mathbf{h}_{pf}(\mathbf{x})$  [7].

#### B. Correcting possible network states

After a possible state  $\mathbf{x}^k$  has been calculated, any mismatch between its values and  $\mathbf{z}_{npf}^k$  needs to be reduced

$$\mathbf{e}_{mismatch} = \mathbf{h}_{npf}(\mathbf{x}^k) - \mathbf{z}_{npf}^k \quad (9)$$

Where  $\mathbf{h}_{npf}(\mathbf{x})$  is the non-linear measurement function that relates the state variables to the measurements in  $\mathbf{z}_{npf}^k$ . If for instance, in a radial network, the power delivered through a branch is measured. The sum of all power injections and line losses that are hierarchically lower than the flow measurement must be equal to the amount of power delivered through the measured branch. Therefore, a mismatch between the measurements can occur and has to be reduced. This can be done by the following steps:

**Step 1:** Select the hierarchically lowest measurement from  $\mathbf{z}_{npf}^k$  that has not been corrected yet. If no measurements have to be corrected stop. Else go to step 2.

**Step 2:** Determine the mismatch  $\mathbf{e}_{mismatch}$  between the calculated state and the selected measurement by (9). Go to step 3.

**Step 3:** If  $\mathbf{e}_{mismatch}$  is smaller than a convergence criteria go to step 1, else go to step 4.

**Step 4:** Determine all measurements from  $\mathbf{z}_{pf}^k$  that are hierarchically lower than the selected measurement and have not been corrected in the previous iteration. Go to step 5.

**Step 5:** Calculate the scaling factor (SF) by

$$SF = \frac{\mathbf{e}_{mismatch}}{\sum_{i=1}^Y STD_i} \quad (10)$$

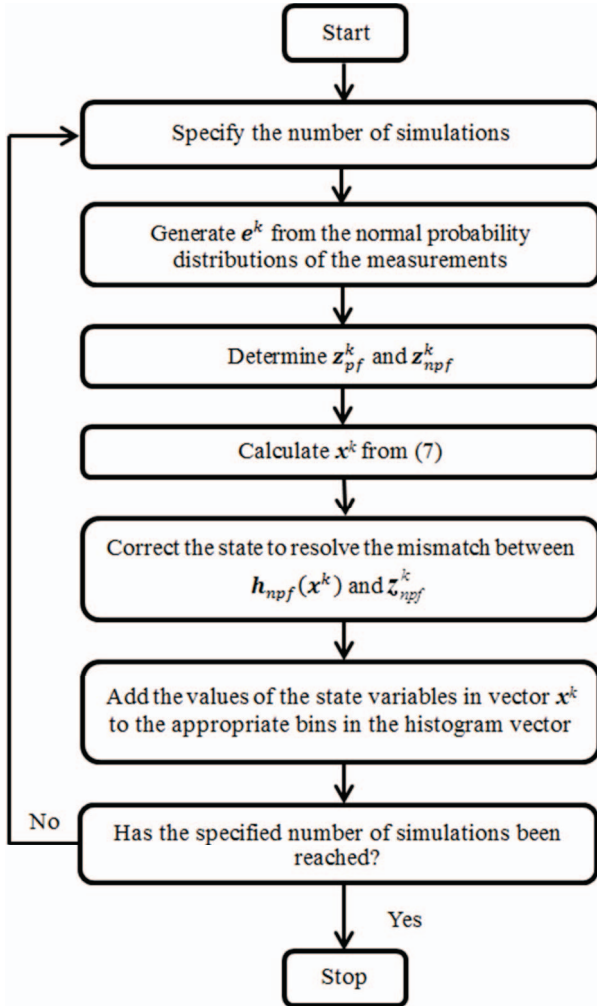
where  $Y$  is the number of downstream measurements determined in step 4, plus the measurement selected in step 1;  $STD$  is the vector of standard deviations of the selected measurements. Go to step 6.

**Step 6:** Correct all selected measurements by multiplying their standard deviation with  $SF$  and subtracting the resulting values from the corresponding measurement values. Go to step 7.

**Step 7:** Recalculate  $\mathbf{x}^k$  for the corrected state by (7) and go back to step 2.

### C. Calculating the Histogram

After a sufficiently large number of possible network states  $x^k$  have been calculated and corrected, the resulting state variable values are binned. This way a histogram vector of the same size as  $x$  is obtained. The following steps are applied to calculate the histogram vector:



The probability of a state variable taking a certain value is equal to the number of results in its corresponding bin, divided by the total number of simulations conducted. The observability index of the network can then be calculated as illustrated in the example given in Section II C.

## IV. CASE STUDY

### A. Test Network

This case study is conducted on a modified IEEE 34-node test network to demonstrate the applicability of the proposed methods. The test network is shown in Fig. 1 whilst original data can be found in [14].

For simplicity, all voltage control devices have been removed and three-phase balanced loads were assumed. The

network loading has been reduced by factor four to account for the removed voltage control devices.

The following monitoring scheme is applied throughout the case study.

- 1) Power Flow measurements (illustrated by dotted arrows), have a standard deviation of 1% of their expected value.
- 2) Pseudo-measurements for power injections (illustrated by solid arrows), have a standard deviation of 10% of their expected value.
- 3) Busses without an arrow have very little information about connected loads. They are represented by pseudo-measurements, with a standard deviation of 100% of their expected value.

Note that the assumed three-phase balanced loads are used for demonstration purposes only and that for practical applications unbalanced three phase loads should be assumed [18, 19].

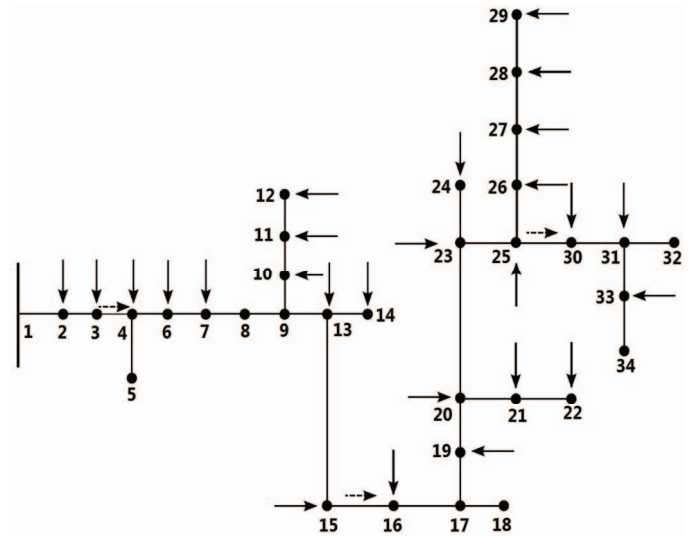


Fig. 1. The modified IEEE 34-node test network.

We chose the minimum and maximum state variable value to be 0 and 2 p.u and a bin size of  $2e^{-4}$  p.u. The small bin size ensures that the probability distribution of the state variables is represented correctly by the histogram. The maximum value of the observability index was calculated to be 451.8 according to (4). For the test network, 5000 Monte Carlo simulations were found to be sufficient to produce consistent results. Therefore, to calculate the observability index a number of 5000 simulations was used.

### B. Influence of the Network Loading on Observability

To illustrate the influence of the network loading on the observability, the proposed observability index was calculated in 15 min intervals over a 24 hour period.

The nodes 29, 19, 30, 26, 20, 16 and 14 were assigned commercial load profiles. All other nodes were assigned residential load profiles.

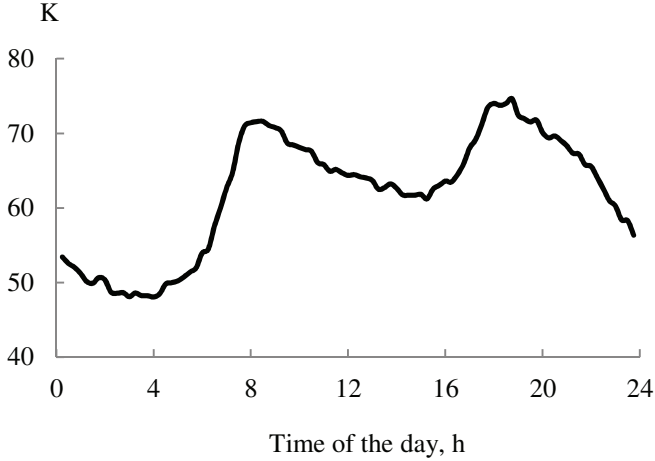


Fig. 2. Observability index K over a period of 24 h.

As shown in Fig. 2, the observability index K changed during the 24 hour period due to load variations. The reason for this behavior is that the measurement error depends on the measurement accuracy given as a percentage of the measured value. Therefore, if the measured value increases the measurement error also increases proportionally. As a result the observability changes with the network loading. This implies that the same network with the same meter configuration can be found observable at one point in time and less/more observable at another.

### C. Optimal Measurement Placement

In order to increase the network observability to a desired level, additional meters have to be placed in the network. But due to financial restrictions, only a limited number of metering devices can be placed in a distribution network. This leads to the problem of optimal meter placement [9].

Since the proposed observability index value changes, depending on the measurement configuration of the network, it is possible to compare different meter placements with each other and determine the optimal location for additional metering devices.

In order to determine the optimal location for an additional power flow meter, the placement of a power flow meter at every possible location in the test network was simulated.

The resulting reductions in SE uncertainty expressed by the observability index are then compared with each other in order to determine the optimal meter placement location. Table II shows the reductions achieved by the meter placements. Before the meter placement, the observability index was equal to 74.0.

It can be seen the metering position between busses 7 and 8 results in the highest reduction of the state estimation uncertainty. Therefore, this position should be chosen for the placement of an additional power flow meter.

The placement of a metering device at a particular bus can significantly affect the observability of the entire network. Therefore, to find the optimal position for multiple metering

devices every possible combination of metering positions needs to be evaluated.

TABLE II: Results of the optimal meter placement

From Bus	To Bus	Obs. Index Reduction	From Bus	To Bus	Obs. Index Reduction
1	2	0.9	17	19	3.8
2	3	0.8	19	20	3.6
4	5	1.6	20	21	0.9
4	6	6.2	21	22	1.4
6	7	7.8	20	23	0.8
7	8	8.4	23	24	0.6
8	9	0.5	23	25	0.7
9	10	0.3	25	26	1.4
10	11	0.2	26	27	1.2
11	12	0.8	27	28	0.5
9	13	1.0	28	29	0.6
13	14	0.7	30	31	0.5
13	15	0.1	31	32	0.7
16	17	0.8	31	33	0.4
17	18	0.5	33	34	0.5

### D. Comparison with Existing Method

To demonstrate the advantages of the proposed observability index it was compared with the method presented in [20]. This method calculates the network observability based on the inverse function theory and the jacobian matrix of the network. Thus, the observability classification only depends on the set of available measurement and their positions in the network.

Since the busses without an arrow only have very little information about the connected loads they are considered as unmeasured busses. Fig. 3 shows the result of the observability analysis using the method in [20].

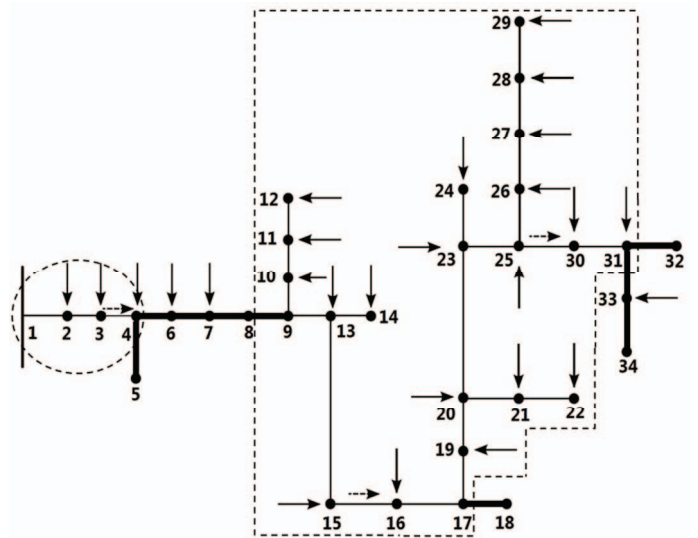


Fig. 3. The network used for the comparison observability analysis.

It can be seen that the network is found to be unobservable and is subdivided into two observable islands (indicated by the dotted lines). Unobservable branches are shown as bold lines. The state of an unobservable island can be calculated if at least one voltage measurement is present in the unobservable island as a voltage reference. However, due to the generally limited number of real time measurements this is rarely the case in distribution networks.

If the pseudo-measurements at the busses 5, 8, 9, 17, 18, 32 and 34 would be used as additional measurements, the network would be classified as observable without the use of additional expensive metering devices. However, in the traditional approach the SE uncertainty introduced by the large margins of error associated with these pseudo-measurements is not known and could be significant.

In contrast to the traditional method, the proposed method provides a continuous observability index value of 74.0 for the same network. This value takes the measurement uncertainties into account. An experienced network operator would be able to interpret this value and could define a threshold for the network observability index value, which represent an acceptable level of uncertainty in the state estimation result.

This illustrates the advantages of the proposed observability index which provides a continuous number that quantifies the uncertainty of the system state, and therefore, gives more reliable information about the network observability.

## V. CONCLUSION

A new observability index for distribution networks based on information entropy has been proposed and described. It was shown that this index can represent the network observability as a single continuous number, by quantifying the uncertainty of a distribution network state estimation. This method provides valuable information about the accuracy of the state estimation result. Simulations over a 24 hour period have been used to illustrate the influence of the network loading on the network observability. An optimal meter placement method based on the proposed observability index was shown to be a useful and effective tool in determining the optimal meter placement in a distribution network. Finally the advantages of the proposed method were shown by comparing the results of an existing method with the proposed observability index.

## ACKNOWLEDGMENT

The authors would like to acknowledge the support of TasNetworks and ASTP (Australian Strategic Technology Program) for this project.

## REFERENCES

- [1] A. Abur, A. Keyhani, and H. Bakhtiari, "Autoregressive Filters For the Identification and Replacement of Bad Data in Power System State Estimation," *Power Engineering Review, IEEE*, vol. PER-7, pp. 33-34, Aug. 1987
- [2] A. P. Sakis Meliopoulos and Z. Fan, "Multiphase power flow and state estimation for power distribution systems," *IEEE Trans. Power Syst.*, vol. 11, pp. 939-946, May 1996
- [3] L. Whei-Min and T. Jen-Hao, "State estimation for distribution systems with zero-injection constraints," *IEEE Trans. Ind. Electron.*, vol. 11, pp. 518-524, Feb 1996.
- [4] F. C. Schweppe and J. Wildes, "Power System Static-State Estimation, Part I: Exact Model," *IEEE Trans. Power App. Syst.*, vol. PAS-89, pp. 120-125, Jan. 1970.
- [5] C. N. Lu, J. H. Teng, and W. H. E. Liu, "Distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 10, pp. 229-240, Feb 1995
- [6] A. Monticelli and F. F. Wu, "Network Observability Theory," *IEEE Trans. Power App. Syst.*, vol. Vol.PAS-104, pp. 1042-1048, May 1985
- [7] A. Abur and A. G. Expósito, *Power System State Estimation Theory and Implementation*. New York: Marcel Dekker, Inc., 2004.
- [8] G. R. Krumpolz, K. A. Clements, and P. W. Davis, "Power System Observability: A Practical Algorithm Using Network Topology," *IEEE Trans. Power App. Syst.*, vol. PAS-99, pp. 1534-1542, July 1980
- [9] N. Nusrat, M. Irving, and G. Taylor, "Novel meter placement algorithm for enhanced accuracy of distribution system state estimation," in *Power and Energy Society General Meeting*, 2012, pp. 1-8.
- [10] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling," *IEEE Trans. Power Syst.*, vol. 27, pp. 1888-1896, April 2012
- [11] L. Ke, "State estimation for power distribution system and measurement impacts," *IEEE Trans. Power Syst.*, vol. 11, pp. 911-916, May 1996
- [12] D. Echternacht and A. Moser, "Cost optimal meter placement in low and medium voltage grids considering stochastic dependencies," in *Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, 2013, pp. 1-5.
- [13] M. Powalko, A. G. Orths, H. Abildgaard, P. B. Eriksen, K. Rudion, I. I. Golub, *et al.*, "System observability indices for optimal placement of PMU measurements," in *Power and Energy Society General Meeting, IEEE*, 2012, pp. 1-6.
- [14] (05 August 2013). *IEEE PES Radial Distribution Test Feeders* [Online]. Available: <http://ewh.ieee.org/soc/pes/dsacom/testfeeders/>
- [15] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal, The*, vol. 27, pp. 379-423, Oct. 1948
- [16] M. Negnevitsky, J. Terry, and N. Thanh, "Using information entropy to quantify uncertainty in distribution networks," in *Power Engineering Conference (AUPEC), Australasian Universities*, 2014, pp. 1-6.
- [17] C. M. Hird, H. Leite, N. Jenkins, and H. Li, "Network voltage controller for distributed generation," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 151, pp. 150-156, March 2004
- [18] I. Dzafic, D. Ablakovic, and S. Henselmeyer, "Real-time three-phase state estimation for radial distribution networks," in *Power and Energy Society General Meeting, IEEE*, 2012, pp. 1-6.
- [19] E. Farantatos, R. Huang, G. J. Cokkinides, and A. P. Meliopoulos, "Implementation of a 3-phase state estimation tool suitable for advanced distribution management systems," in *Power Systems Conference and Exposition (PSCE), IEEE/PES*, 2011, pp. 1-8.
- [20] R. Gelagaev, P. Vermeyen, J. Vandewalle, and J. Driesen, "Numerical observability analysis of distribution systems," in *Harmonics and Quality of Power (ICHQP)*, 2010, pp. 1-6.