

# POWER SYSTEM STATE ESTIMATION AND OBSERVABILITY ANALYSIS USING ARTIFICIAL NEURAL NETWORKS

Soham Chakraborty

SRM University, Kattankulathur, Chennai

**Abstract:** This work proposes a static state estimation method and network observability analysis using artificial neural networks trained using Bayesian Regularization. This method is aimed at reducing computational complexities as in case of conventional method. The proposed method is to be carried out and compared with the conventional weighted least square state estimation algorithm on a system.

By using ANN based State estimator, it will be a direct procedure unlike WLS method which is a iterative procedure comprising of complex calculations. Here the input and output are mapped in an ANN, so we get direct estimates of state variables given the measurements as input. Hence there is no need to store the huge complex YBUS admittance matrix and perform complex calculations. ANN is implemented to get better and faster results from the state estimator used in load dispatch centres and also smart grids.

Observability Analysis of power system is done prior to running state estimator. If a system is found to be unobservable, then additional meters may have to be placed at particular locations to make the system observable. Network observability analysis identifies the observable islands, unobservable branches and the locations for placement of meters to make the system observable. This work is an iterative procedure and can be done using ANN that will provide a direct procedure to solve the problem.

**Key words:** State Estimation, Weighted Least Square method, Artificial Neural Networks, Numerical Method based on Nodal Variable formulation, Observability Analysis, Measurement Placement to restore observability

## I. INTRODUCTION

State estimation is the process of assigning a value to an unknown system state variable based on measurements from that system according to some criteria. Usually, the process involves imperfect measurements that are redundant, and the process of estimating the system states is based on a statistical criterion that estimates the true value of the state variables to minimize or maximize the selected criterion.

In a power system, the state variables are the voltage magnitudes and relative phase angles at the system nodes. Measurements are required in order to estimate the system

performance in real time for both system security control and constraints on economic dispatch.

The inputs to an estimator are imperfect power system measurements of voltage magnitudes and power, VAR, or ampere-flow quantities. The estimator is designed to produce the “best estimate” of the system voltage and phase angles, recognizing that there are errors in the measured quantities and that there may be redundant measurements.

The output data are then used in System Control Centers or Load Dispatching Centre in the implementation of the security-constrained dispatch and control of the system. The best estimates help in maintaining power system monitoring, security, reliability, and proper control of the system.

## II. STATE ESTIMATION USING WEIGHTED LEAST SQUARE METHOD

### A. Maximum Likelihood Weighted Least-Squares Estimation

a) *The maximum likelihood criterion*, where the objective is to maximize the probability that the estimate of the state variable,  $\hat{x}$ , is the true value of the state variable vector,  $\hat{x}$ , (i.e., maximize  $P(\hat{x}=x)$  [8].

b) *The weighted least-squares criterion*, where the objective is to minimize the sum of the squares of the weighted deviations of the estimated measurements,  $\hat{z}$ , from the actual measurements,  $z$  [8].

c) *The minimum variance criterion*, where the object is to minimize the expected value of the sum of the squares of the deviations of the estimated components of the state variable vector from the corresponding components of the true state variable vector [8].

### B. Gaussian or Normal Probability Density Function

If the uncertainties in measured values are plotted as a function of their relative frequency of occurrence, a histogram is obtained to which a continuous curve can be fitted as the number of measurements increases (theoretically, to an infinite number). The continuous curve most commonly encountered is the bell shaped function called the Gaussian or Normal Probability Density Function [6].

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} \quad (1)$$

z – random variable,  $\mu$  – mean,  $\sigma$  – standard deviation

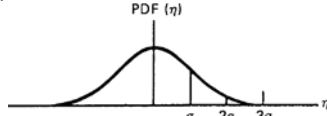


Figure 1: The Normal density function

### C. The Likelihood Function

The joint probability density function which represents the probability of measuring m independent measurements, each having the same Gaussian P.D.F can simply be expressed as the product of individual P.D. F's if each measurement is assumed to be independent of the rest.

$$f_m(z) = f(z_1) f(z_2) \dots f(z_m) \quad (2)$$

where  $z_i$  : ith measurement,  $z^T = [z_1, z_2, \dots, z_m]$ . The function  $f_m(z)$  is called the likelihood function for z. It is a measure of the probability of observing the particular set of measurements in the vector z. The objective of the maximum likelihood estimation is to maximize this likelihood function by varying the assumed parameters of the density function, namely its mean and standard deviation [6]. The modified function is called the Log-Likelihood function, L.

$$L = \log f_m(z) = \sum_{i=1}^m \log f(z_i) = -\frac{1}{2} \sum_{i=1}^m \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 - \frac{m}{2} \log 2\pi - \sum_{i=1}^m \log \sigma_i \quad (3)$$

MLE will maximize the likelihood or log-likelihood function for a given set of observations  $z_1, z_2, \dots, z_m$ . It can be obtained by solving the following optimization problem and the solution is called the *weighted least square* (WLS) estimator for x.

Maximize  $\log f_m(z)$

or

Minimize  $\sum_{i=1}^m \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 \quad (4)$

or

Minimize  $\sum_{i=1}^m W_{ii} r_i^2 \quad (5)$

subject to  $z_i = h_i(x) + r_i, i=1,2,\dots,m$

here  $W_{ii} = \sigma_i^{-2}, r_i = z_i - \mu_i$

### D. Measurement Model

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} h_1(x_1, x_2, \dots, x_m) \\ h_2(x_1, x_2, \dots, x_m) \\ \vdots \\ h_m(x_1, x_2, \dots, x_m) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = h(x) + e \quad (6)$$

here  $h_i(x)$  is the nonlinear function relating measurement  $i$  to the state vectors  $x_1, x_2, \dots, x_m$  and  $e_1, e_2, \dots, e_m$  are the measurement errors.

Assumptions:  $E(e_i) = 0$  and measurement errors are independent, i.e.  $E\{e_i e_j\} = 0$ , hence Covariance  $(e) = E\{e e^T\} = R = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2\}$ .

The WLS estimator will minimize the following objective function:

$$J(x) = \sum_{i=1}^m \frac{(z_i - h_i(x))^2}{R_{ii}} = [z - h(x)]^T R^{-1} [z - h(x)] \quad (7)$$

At the minimum, the first order optimality conditions will have to be satisfied.

$$g(x) = \frac{\partial J(x)}{\partial x} = -H^T R^{-1} [z - h(x)] = 0 \quad (8)$$

where  $H(x) = \left[\frac{\partial h(x)}{\partial x}\right]$  is the Jacobian matrix

Expanding the nonlinear function  $g(x)$  into its Taylor series around the state vector  $x^k$  yields:

$$g(x) = g(x^k) + G(x^k)(x - x^k) + \dots = 0$$

Neglecting the higher order terms leads to an iterative solution scheme known as the Gauss-Newton method as shown below (k – iteration index):

$$x^{k+1} = x^k - [G(x^k)]^{-1} \cdot g(x^k)$$

$$G(x^k) = \left[\frac{\partial g(x^k)}{\partial x}\right] = H^T(x^k) \cdot R^{-1} \cdot H(x^k) \quad (9)$$

$G(x)$  is called the Gain matrix. It is sparse, positive-definite, symmetric provided that the system is fully observable.

$$G(x^k) \Delta x^{k+1} = H^T(x^k) \cdot R^{-1} \cdot [z - h(x^k)] \quad (10)$$

### E. WLS Algorithm

The iterative solution algorithm for WLS state estimation problem can be outlined as follows [6]:

1. The iterations are started with iteration index  $k=0$ .
2. The state vector  $x^k$  is initialized as a flat start.
3. The gain matrix is calculated,  $G(x^k) = H^T(x^k) \cdot R^{-1} \cdot H(x^k)$ .
4. The inverse of gain matrix is determined by obtaining the adjoint matrix and dividing it by the determinant and so  $\Delta x^{k+1} = G(x^k)^{-1} H^T(x^k) \cdot R^{-1} \cdot [z - h(x^k)]$ . (Generally gain matrix is decomposed into Cholesky factors  $G = L \cdot L^T$  and forward/backward substitution is performed to obtain  $\Delta x^{k+1}$ ).
5. Convergence criterion is tested  $\max |\Delta x^{k+1}| < \epsilon$
6. If no, update  $x^{k+1} = x^k + \Delta x^k, k = k + 1$ , and go to step 3, Else stop.

### F. The Measurement Function

Most commonly used measurements are the line power flows, bus power injections, bus voltage magnitudes and line current flow magnitudes. These measurements can be expressed using state variables using rectangular or polar coordinates. When using polar coordinates for a system containing N buses, we have  $2N-1$  state variables, N bus voltages and  $N-1$  phase angles, with the phase angle of reference bus is set equal to an arbitrary value, such as 0. The state vector x will have the following form assuming bus 1 is chosen as the reference:

$$x^T = [\theta_2 \theta_3 \dots \theta_N V_1 V_2 \dots V_N]$$

Assuming a two-port  $\pi$  model for the network branches, the expressions for different types of measurements are given below.

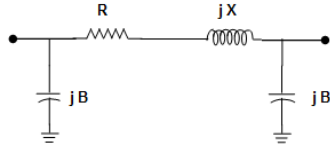


Figure 2: Two port pi model of a network branch

- Real and Reactive power injection at bus  $i$ :
 
$$P_i = V_i \sum_{j \in Ni} V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (11)$$

$$Q_i = V_i \sum_{j \in Ni} V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (12)$$

- Real and Reactive power flow from bus  $i$  to bus  $j$ :
 
$$P_{ij} = V_i^2 (g_{si} + g_{ij}) - V_i V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (13)$$

$$Q_{ij} = V_i^2 (b_{si} + b_{ij}) - V_i V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (14)$$

- Line current flow magnitude from bus  $i$  to  $j$ :

$$\frac{\sqrt{P_{ij}^2 + Q_{ij}^2}}{V_i}$$

or ignoring the shunt admittance ( $g_{si} + jb_{si}$ ):

$$\sqrt{(g_{ij}^2 + b_{ij}^2)(V_i^2 + V_j^2 - 2V_i V_j \cos \theta_{ij})} \quad (15)$$

where  $V_i, \theta_i$  are the voltage magnitude and phase angle at bus  $i$ .

$$\theta_{ij} = \theta_i - \theta_j$$

$G_{ij} + B_{ij}$  is the  $ij^{\text{th}}$  element of the complex bus admittance matrix.

$g_{ij} + b_{ij}$  is the admittance of the series branch connecting buses  $i$  and  $j$ .

$g_{sj} + b_{sj}$  is the admittance of the shunt branch connected at bus shown in figure 3.2.

$Ni$  is the set of buses connected directly to bus  $i$ .

### G. The Measurement Jacobian, $H$

The basic structure of the Jacobian matrix is as follows:

$$H = \begin{bmatrix} \frac{\partial P_{inj}}{\partial \theta} & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \theta} & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \theta} & \frac{\partial Q_{flow}}{\partial V} \\ \frac{\partial I_{mag}}{\partial \theta} & \frac{\partial I_{mag}}{\partial V} \\ \frac{\partial V_{mag}}{\partial \theta} & \frac{\partial V_{mag}}{\partial V} \\ 0 & \frac{\partial V_{mag}}{\partial V} \end{bmatrix} \quad (16)$$

- Elements corresponding to real power injection measurements:

$$\frac{\partial P_i}{\partial \theta_i} = \sum_{j=1}^N V_i V_j (-G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij}) - V_i^2 B_{ii} \quad (17)$$

$$\frac{\partial P_i}{\partial \theta_j} = V_i V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (18)$$

$$\frac{\partial P_i}{\partial V_i} = \sum_{j=1}^N V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) + V_i G_{ii} \quad (19)$$

$$\frac{\partial P_i}{\partial V_j} = V_i (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \quad (20)$$

- Elements corresponding to reactive power injection measurements:

$$\frac{\partial Q_i}{\partial \theta_i} = \sum_{j=1}^N V_i V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) - V_i^2 G_{ii} \quad (21)$$

$$\frac{\partial Q_i}{\partial \theta_j} = V_i V_j (-G_{ij} \cos \theta_{ij} - B_{ij} \sin \theta_{ij}) \quad (22)$$

$$\frac{\partial Q_i}{\partial V_i} = \sum_{j=1}^N V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) - V_i G_{ii} \quad (23)$$

$$\frac{\partial Q_i}{\partial V_j} = V_i (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (24)$$

- Elements corresponding to real power flow measurements:

$$\frac{\partial P_{ij}}{\partial \theta_{ij}} = V_i V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (25)$$

$$\frac{\partial P_{ij}}{\partial \theta_j} = -V_i V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (26)$$

$$\frac{\partial P_{ij}}{\partial V_i} = -V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) + 2V_i (g_{si} + g_{ij}) \quad (27)$$

$$\frac{\partial P_{ij}}{\partial V_j} = -V_i (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (28)$$

- Elements corresponding to reactive power flow measurements:

$$\frac{\partial Q_{ij}}{\partial \theta_{ij}} = -V_i V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (29)$$

$$\frac{\partial Q_{ij}}{\partial \theta_j} = V_i V_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}) \quad (30)$$

$$\frac{\partial Q_{ij}}{\partial V_i} = -V_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) - 2V_i (b_{ij} + b_{si}) \quad (31)$$

$$\frac{\partial Q_{ij}}{\partial V_j} = -V_i (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}) \quad (32)$$

- Elements corresponding to voltage magnitude measurements:

$$\frac{\partial V_i}{\partial V_i} = 1, \frac{\partial V_i}{\partial V_j} = 0, \frac{\partial V_i}{\partial \theta_i} = 0, \frac{\partial V_i}{\partial \theta_j} = 0 \quad (33)$$

- Elements corresponding to current magnitude measurements (ignoring the shunt admittance of the branch):

$$\frac{\partial I_{ij}}{\partial \theta_i} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} V_i V_j \sin \theta_{ij} \quad (34)$$

$$\frac{\partial I_{ij}}{\partial \theta_j} = -\frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} V_i V_j \sin \theta_{ij} \quad (35)$$

$$\frac{\partial I_{ij}}{\partial V_i} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} (V_i - V_j \cos \theta_{ij}) \quad (36)$$

$$\frac{\partial I_{ij}}{\partial V_j} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} (V_j - V_i \cos \theta_{ij}) \quad (37)$$

## III. NETWORK OBSERVABILITY ANALYSIS

### A. Introduction

Power system state estimator uses the set of available measurements in order to estimate the system state. Given a set of measurements and their locations, the network observability analysis will determine if a unique estimate can be found for the system state. This analysis may be carried out off-line during the initial phase of a state estimator installation, in order to check the adequacy of the existing measurement configuration. If the system is found not to be observable, then additional meters may have to be placed at particular locations. Observability analysis is also done on-

line, prior to running the state estimator. Network observability analysis identifies all the observable islands, unobservable branches, and locations for placement of meters to make the system observable. Temporary unobservability may occur due to telecommunication failure [6].

*B. Network Observability Using Numerical Method Based On the Nodal Variable Formulation*

Numerical observability analysis can also be carried out by using the nodal variables. Nodal variable vector is denoted by  $x$  and represents the vector of magnitude and phase angle of all bus voltages in the system [6]. Consider the linearized measurement model, where measurement errors are ignored due to their irrelevance in the observability analysis.

Using the weak coupling between  $P - V$  and  $Q - \delta$ , the linearized model can be decoupled as:

$$\Delta z_A = H_{AA} \Delta \theta \quad \text{and} \quad \Delta z_R = H_{RR} \Delta V \tag{38}$$

$\Delta z_A, \Delta z_R$  is the real and reactive power measurement mismatch vectors respectively.

$H_{AA} = \frac{\partial h_A}{\partial \theta}$  is the decoupled Jacobian for the real power measurements.

$H_{RR} = \frac{\partial h_R}{\partial V}$  is the decoupled Jacobian for the reactive power measurements.

$$\Delta \theta = \theta - \theta_0 \quad \text{and} \quad \Delta V = V - V_0$$

Assuming that the  $P, Q$  measurements come in pairs,  $P - \theta$  and  $Q - V$  observability can be separately tested. Note that, unlike  $\theta$ , the voltage solution requires a measured reference bus. Hence, following the  $P - \theta$  analysis, it should be further checked to ensure that at least one voltage measurement exists per observable island. A network is said to be  $P - \theta$  observable if the rank of matrix  $H_{P\theta}$  is equal to  $N-1$ . Also, network is said to be  $Q - V$  observable if the rank of the  $H_{QV}$  matrix is equal to  $N$ .

It should be noted that the system observability is independent of the branch parameters as well as the operating state of the system. So, all system branches can be assumed to have an impedance of  $j1.0$  p.u. and all bus voltages can be set equal to  $1.0$  p.u. for the purpose of observability analysis. Then, the d.c. power flows along these system branches can be written as:

$$P_b = A\theta \tag{39}$$

where  $P_b$  is the vector of branch flows,  $A$  is the branch-bus incidence matrix,  $\theta$  is the vector of bus voltage phase angles.

If the estimated state  $\hat{\theta}$  is zero, then all branch flows will be zero as given by Equation (39). Using the DC measurement model:  $z_A = H_{AA}\theta$ , the WLS estimate for  $\theta$  will be given by:  $\hat{\theta} = G_{AA}^{-1}t_A$

A null estimate for  $\hat{\theta}$  will be obtained for an observable system when all system measurements  $\Delta z_A$ , i.e. flows and injection measurements, are all zero. If there exists an estimate  $\hat{\theta}$  which satisfies the measurement equation  $Z_A=0$ , yet yields a non-zero branch flow  $P_b \neq 0$ ,  $\hat{\theta}$  will be called an unobservable

state. Furthermore, those branches carrying nonzero flows, will be referred to as unobservable branches.

*C. Identification of Observable Islands*

This procedure needs to be carried out recursively, each time eliminating the irrelevant injections until all observable islands are identified. Irrelevant injections are those that are incident to unobservable branches. The algorithm is given below.

1. All irrelevant branches are eliminated. These are branches that have no incident measurements.
2. Form the decoupled linearized gain matrix for the  $P - \delta$  estimation problem:  $G_{AA} = H_{AA}^T H_{AA}$ .
3.  $G_{AA}$  is factorized by modifying the zero pivots and the right side vector as described above. The state variables are determined and then the branch flows are found.
4. All unobservable branches and all injections that are incident to these unobservable branches are identified and eliminated.
5. If no more unobservable branches are found in step 4, then the observable islands are determined separated by the unobservable branches and stopped. Else considering the modified network, go to step 2.

*D. Measurement Placement to restore Observability*

Once the observable islands are identified, measurements can be added to merge these islands so that eventually a single observable island can be formed. The candidate measurements that can merge islands are:

- the line flows along branches that connect observable islands, and
- the injection at the boundary buses of observable islands.

The following measurement placement algorithm can then be implemented based on the above defined matrices:

1. Form the gain matrix (from  $G = H^T H$ ) and compute the factor matrix  $L$ . Modify the matrix  $L$  with diagonal elements as 1.
2. Since  $G = L D L^T$  diagonal matrix  $D = L^{-1} G (L^{-1})^T$ . Check whether  $D$  has only one zero pivot. If yes, stop, the system is observable.
3. Else assemble  $W$  matrix from the rows of  $L^{-1}$  matrix corresponding to zero pivot positions.
4. Form the candidate measurement list. The list will contain flow and injection measurements incident to branches connecting observable islands, which have already been identified by the observability analysis.
5. Build the measurement Jacobian  $H_c$  for the candidate measurements.
6. Compute  $B = H_c W^T$ . Obtain the reduced row echelon form of  $B$  as  $E$ . The linearly independent rows of  $E$  will correspond to all the measurements required to be placed.

**IV. ARTIFICIAL NEURAL NETWORK**

**A. ANN Configuration**

The ANN's elemental unit is the mathematical neuron, which receives a fraction  $w$  of the input data  $x$ , simplified by a bias  $b$ , and processes the information according to its activation function  $F$  [1].

$$y = F(n) = F(\sum_j w_j \cdot x_j + b) \tag{39}$$

Neural networks are known to be universal approximators for any non-linear function and they are generally used for mapping error tolerant problems that have much data trained in noise. A neural network consists of many layers namely: an input layer, a number of hidden layers and an output layer. The input layer and the hidden layer are connected by synaptic links called weights and likewise the hidden layer and output layer also have connection weights. When more than one hidden layer exists, weights exist between such layers. Neural networks use some sort of "learning" rule by which the connections weights are determined in order to minimize the error between the neural network output and desired output.

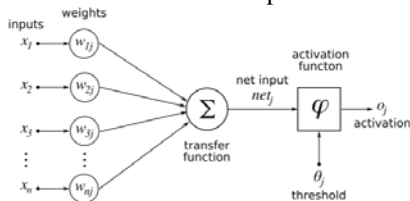


Figure 3: Basic Layout of ANN

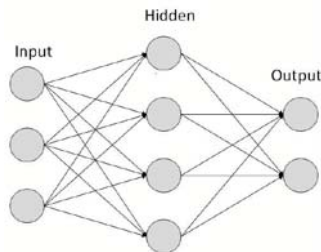


Figure 4: Multi Layered ANN

**B. Polynomial Curve Fitting**

The problem is to fit a polynomial to a set of  $N$  data points by the technique of minimizing an error function [7]. Considering the  $M^{\text{th}}$  order polynomial given by

$$y(x) = w_0 + w_1x + \dots + w_mx^m = \sum_{j=0}^m w_j \cdot x^j \tag{40}$$

This can be regarded as a non-linear mapping which takes  $x$  as input and produces  $y$  as output. The precise form of the function  $y(x)$  is determined by the values of the parameters  $w_0, \dots, w_m$ , which are analogous to the weights in a neural network.

The desired outputs  $y$  for a specific input  $x^n$  is called target values  $t^n$ . In order to find the suitable values for the coefficients in the polynomial, it is convenient to consider the error between the desired output  $t^n$ , for a particular input  $x^n$  and

the corresponding value predicted by the polynomial function given by  $y(x^n; w)$ . Standard curve fitting procedures involve minimizing the square of this error, summed over all data points, or mean squared error, given by

$$E = \frac{1}{N} \sum_{n=1}^N \{(x^n; w) - t^n\}^2 \tag{41}$$

In order to assess the capability of the polynomial to generalize to new data, it is convenient to consider the root mean square (RMS) error given by

$$E^{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N \{(x^n; w) - t^n\}^2} \tag{42}$$

**C. Bayesian Learning of Network Weights**

The Bayesian approach considers a probability distribution function over weight space, representing the relative degrees of belief in different values for the weight vector. Once the data has been observed, it can be converted to a posterior distribution through the use of Bayes' theorem and can be used to evaluate the predictions of the trained network for new values of the input variables [7]. In the Bayesian framework, we consider a probability distribution over a weight values. In the absence of any data, this is described by a prior distribution which we shall denote by  $p(w)$ . Here  $w = (w_0, \dots, w_M)$  denotes vector of adaptive weight (and bias) parameters. Let the target data from the training set be denoted by  $D = (t^1, \dots, t^N)$ . Once we observe the data  $D$  we can write down an expression for the posterior probability distribution for the weights, which we denote by  $p(w/D)$ , using Bayes' theorem.

$$p(w/D) = \frac{p(w/D)p(w)}{p(D)} \tag{43}$$

The normal objective function is the Mean Squared Error (MSE) function.

**V. ANN BASED STATE ESTIMATOR AND OBSERVABILITY ANALYSIS**

The ANN is trained to perform a nonlinear mapping between input measurements and output state variables and other estimated measurements. Once the ANN is trained, state estimation is a direct procedure, without having the need of going through the iterative procedure of WLS method that consumes a lot a computational memory. Also,  $Y_{BUS}$  matrix is also not required to be stored. Thus ANN saves time, computational burden reduced [1].

In case of observability analysis, ANN is trained with various patterns of input measurements [3]. In this case the inputs are the locations of measurements available in the system. If an input is available at a particular location, injections or flows, 1 is taken as an input to the ANN, and if it's not available, then input to the ANN is 0. The output of the ANN is the locations of the measurement placement to restore observability of the system. Flow measurements are the candidate measurements here. So the outputs of the ANN are the branches where measurement placement is to be done. If













