Economic Operation and Control of Power System

Dr. Gururaj Mirle Vishwanath

Department of Electrical Engineering

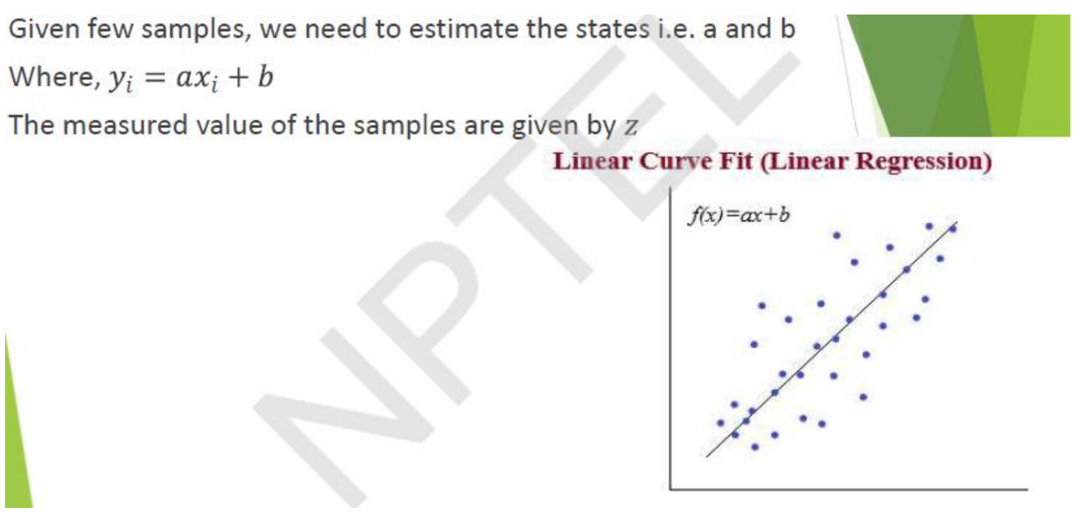
IIT Kanpur

Week - 09

Lecture – 45

Hello and good morning everyone, welcome you all for the NPTEL online course on Economic Operation and Control of Power Systems. So, we will discuss State Estimation. So, let us say there are few samples have been given to you, we need to estimate the states that is A and B. So, states are those variable by knowing them you can know the unknown things. So, we can express output variable as a function of state variable and the input variable, y is equal to Ax plus B. So, the measured value of the samples are given by z, this is the actual measured value.

Now, you are trying to estimate the states of that measured value actually. So, this can be expressed using a curve fit formula:



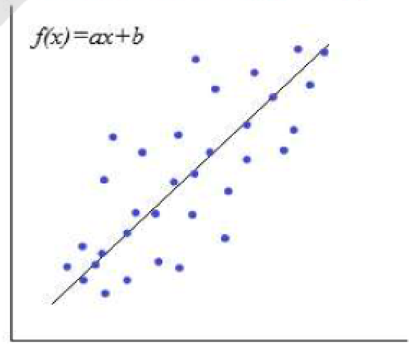
f of x is equal to Ax plus B that is what is output that you need to find it out. So, this is depending upon the state variables A and B and the input variable x. So, the objective is to obtain that specific curve fitting function which will be able to estimate the states of the given measurement.

So, the error obviously there will be error know the error is the difference between the measured value and the estimated value which is a function of state variables. So, E1 is equal to:

$$\varepsilon_1 = z_1 - (ax_1 + b)$$
$$\varepsilon_2 = z_2 - (ax_2 + b)$$

Minimize $\quad \sum_{i=1}^{m}(\varepsilon_i)^2 = \min\limits_{a,\, b} \sum_{i=1}^{m}\left(z_i - (ax_i + b)\right)^2$

Linear Curve Fit (Linear Regression)



In matrix form,

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} - \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ & \vdots \\ x_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{Or} \quad \varepsilon = z - hs$$

Where, z is measurement matrix and $h = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ & \vdots \\ x_m & 1 \end{bmatrix}$

Now, the objective is what to minimize the error because error could be negative and positive. So, henceforth in order to avoid that what we are doing is error square. So, even if there is a negative, so we will make it as a positive.

So, that means minimizing this actual measurement and the estimated measurement estimated value. So, in matrix form if you want to represent it error matrix is nothing, but the measured matrix minus this H matrix we call it as H matrix. So, function matrix basically x11, x21, xm up to 1 into this state variable matrix. These are the matrix this is the matrix that you need to find out A and B. So, or you can express epsilon or error matrix is equal to z minus H into S.

Minimize $\quad J = \min \sum_{i=1}^{m}(\varepsilon_i)^2 = \min\limits_{a,\, b} \sum_{i=1}^{m}\left(z_i - (ax_i + b)\right)^2 = \min \varepsilon^T \varepsilon$

Where, $\varepsilon = z - hs$

$\therefore \varepsilon^T \varepsilon = (z^T - s^T h^T)(z - hs)$

- $\min J = \min \varepsilon^T \varepsilon = \min(z^T z - s^T h^T z - z^T hs + s^T h^T hs)$

$$= \min(z^T z - 2z^T hs + s^T h^T hs)$$
$$\nabla_s J = -2h^T z + 2h^T hs$$

$min\, J$ is achieved when, $\nabla_s J = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \therefore s = (h^T h)^{-1} h^T z$

Example: Perform linear regression on (1,1.5), (2, 0.8), (3, 2.2)

- $z = \begin{bmatrix} 1.5 \\ 0.8 \\ 2.2 \end{bmatrix}$, $H = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$
- $h^T h = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$
- $S = (h^T h)^{-1} h^T z = \begin{bmatrix} 0.35 \\ 0.8 \end{bmatrix}$
- $y = 0.35 * x + 0.8$

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

Where,

where, J is the objective function to be minimized

- We now define a quantity called the measurement residual, r_i, for the i^{th} measurement, as shown below.

$$r_i = z_i - \mu_i = z_i - E(z_i) \qquad \forall\, i = 1, \dots, m$$

- Where, $E(z_i)$ is the expected value of the measurement z_i
- Notice that the expected value of z_i can be determined from the equations or functions of the measurements in terms of the system states.

Now, the objective is to minimize this error or minimize that this matrix z which is actual measurement and this estimated matrix Ax plus B. Now, what is this? This can be in mathematics form expressed as e t into e transpose matrix into e because the square matrix. So, in matrix form you can express this as e t into e. So, e t what is e we do not we know z minus H S. So, find out what is e t into e that means t minus this matrix whole transpose into the matrix.

So, if you have to minimize e t into e, so you will have to minimize this whole matrix. You just multiply it multiply one to one. So, you will get this term. Now, minimizing means the differential equation. You need to differentiate and that should be equal to 0.

So, gradient of this that should be equal to 0. Now you differentiate with respect to S. Now what you will get? This is a constant term z t into z because that is a measurement value which you are getting it is a constant term. So, this is anyway constant derivation will be 0. Now you have to differentiate these two terms with respect to S.

Now what you will get? Minus 2 into z t into H because dou S by dou S is 1 plus this is what S t into S this is nothing but S square basically. So, what you will get? 2 into S basically when you differentiate that is nothing but 2 into H t into H into S. Now this is equal to 0. If you make this as 0 what you will get? Minus 2 H t z plus 2 H t H S is equal to 0. You need to find out S basically.

S means 2 H t z multiplied by H t into H whole inverse. This is what we have done. Of course 2 2 gets cancelled. So, you will ultimately get H t into H whole inverse into H t z. This is what you will get.

That is a state matrix that you are going to find out. You know what is the measurement matrix, actual matrix with the because you are getting samples you know what is z. Having known this model function matrix which is having these variables x1, x2 and all these parameters then you have to do this mathematics only H t H inverse into H transpose and you will get the state variable matrix. Now perform an example. Perform linear regression on 1, 1.5, 2.8, 3, 2.2. Here 1.5, 0.8, 2.2 are the measured values. So that is why we have kept it in as a z matrix and remaining you have the H matrix.

This measurement is taken for a specific function and that value is considered as 1. So this measurement is taken at a specific condition where that value of that function matrix is 2, similarly 3. So the function matrix you get is H is equal to 1, 1, 2, 1, 3, 1. Now you do it H t into H and then H t H inverse H t into z then you will get the state matrix that is 0.35, 0.8. So basically for this sort of measurement the state variable is 0.358. By knowing 0.358 and having different value of 1 to H matrix you can find out the value of Y which should be ultimately closer to the z matrix, the measurement matrix basically. So let us find out maximum likelihood estimator.

So the objective is to minimize this. What is this? This is the measurement value. This is average. This is standard deviation. This need to be minimized.

We have discussed in greater detail about Gaussian matrix and how the data should be closely present with respect to the true value. So where J is the objective function to be minimized we now define a quantity called the measurement residual r i for the ith measurement as shown below. This residual value for this z measurement at the ith sample, at the ith measurement so that is nothing but the difference. So this is the actual measurement, this is the average. So average is nothing but expectation of z where e z i is the expected value of the measurement z i.

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

Where,

where, J is the objective function to be minimized

- We now define a quantity called the measurement residual, r_i, for the i^{th} measurement, as shown below.

$$r_i = z_i - \mu_i = z_i - E(z_i) \qquad \forall\, i = 1, \ldots, m$$

- Where, $E(z_i)$ is the expected value of the measurement z_i
- Notice that the expected value of z_i can be determined from the equations or functions of the measurements in terms of the system states.

Notice that the expected value of z i can be determined from the equations or functions of the measurements in terms of the system states. The expected value of z can be obtained from the system states. So basically this is what the measurement matrix and this is expectation matrix. So expected value should be close to the true value. That means you want residual item should be 0.

That will be the objective. The error, this is basically error. So for example, power flow measurement at a bus can be determined from the power flow equation given the values of the bus voltage, magnitude and angles. Their bus voltage and angle, they are the state variables. Correct? So, let us say you need to find out power flow, that p is nothing but we know this way, V i V j Y ij cosine of theta i minus theta j minus alpha ij. Right? Similarly, reactive power flow is V i V j Y ij sin of theta i minus theta j minus alpha ij.

Power flow is depending upon voltage, magnitudes and the phase angles, whether active or reactive power. So these are the state variables here. And it also depends upon transmission line parameters. Here that is been taken care by this Y matrix. Correct? Having known this voltage, magnitude, phase, angle and this transmission line parameters, you would be able to find out what is the line flow.

Right? Now, let us now define the weight assigned to the ith measurement as W i is equal to 1 by sigma i square. Right? So, above means that the measurement that have higher variances are assigned less weights. That means you are getting a measured value and that is too much deviated from the true value actually, has lot of variances. That is what is standard deviation identifies us. Now, if you have the data which is far away from the standard value of actual measurement, so that for that specific measurement we are giving very less weightage.

That means you do not want a specific measurement to have a lot of deviation or variance. So this weight matrix is very important. So the problem of determining the maximum likelihood estimates then becoming then becomes a following. So this is a minimize, this is the function that you need to minimize subject to Z i is equal to H i of x plus E i. What is H i? It is the measurement function for the ith measurement.

X is the vector, that means we are assuming that there is a measurement and it may have some error. This H of x is the true actual measurement. The objective is to find out that H of x now. The measurement has some error and that error we are assuming that there will be that error and there will be some true value also. So we are trying to find out the true value.

Now x is the vector of the system states because that depends upon the system states. Actual value depends upon the system states basically consisting of bus voltage magnitude and phase angles and E is the vector of measurement errors. Correct? Above formulation for the state estimation is the so-called weighted least square state estimation problem. Now the ith measurement Z i is modeled as given below expressing all the measurements in a compact form.

Now just express in the matrix form. This is the measurement matrix that is equal to H of x plus error, where H is the vector of measurement functions of the measurement vector Z. Following assumptions are taken regarding the statistical properties of the measurements, some of the assumptions. The first assumption is expectations of the measurement errors are zero. So ultimately as a true system state estimator, I do not want any error to be present in my estimation. So the measurement value should have zero error.

So that means expectation of error, do you want to have high expectation? No, you should have minimum or zero expectation. So expectation of the actual measurement with respect to average that should be zero for all the measurements and measurement errors are independent. That means you are getting measurement from different devices let us say and they are and they have each measurement device may have its own error and that error may not have dependency another error, which is coming from some other measurement device. So basically what we are trying to find out is, is there any interdependency in error that is being populated from different measurement devices? That we are going to find out. So what we are saying is, let us say E i is a measurement taken from some measurement device, E j is a measurement taken from other measurement device.

So their dependency is zero basically. That means:

- Assuming the measurement errors to be independent, the error-covariance matrix, **R**, mentioned in the last equation becomes a diagonal matrix, with the inverse of the measurement variances as the diagonal elements. The non-diagonal elements are zeros. The *i*th diagonal element, $\mathbf{R}(ii)$ of $\mathbf{R}$ is given by,

$$E(e_i e_i) = E\{(z_i - \mu_i)(z_i - \mu_i)\} = E\{(z_i - \mu_i)^2\} = \sigma_i^2$$

- The objective function can then be written as

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

Now this is a diagonal matrix. If there is any dependency in this matrix, the off diagonal elements will be non-zero. Now this is a dependency on the first error with respect to that measurement itself. This is a diagonal element. This is second measurement with that there, there will be dependency on that variable, that measurement with respect to that measurement itself. So this is second entry, that second measurement basically.

This diagonal entry surely it will be non-zero matrix, non-zero entry. But off diagonal it will be zero if there is no dependency on each measurement. Now this diagonal matrix and now the entire covariance matrix, let us term it as R basically. Now assuming the measurement error to be independent, the error covariance matrix R mentioned in the last equation becomes a diagonal matrix with inverse of the measurement variances as the diagonal elements.

The non-diagonal elements are zeros basically. The ith diagonal element RiI of Ri is given by, now this is expectation Ei into Ei because we are just looking at that specific measurement itself. That is E into Zi minus Mi into Zi minus Mi and that is what Zi minus Mi whole square, that is expectation of this is nothing but the standard deviation, standard deviation. Now the objective function can then be written as G of x is equal to:

$$\mathbf{g}(\mathbf{x}) = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = -\sum_{i=1}^{m} \left\{ \frac{z_i - h_i(\mathbf{x})}{\sigma_i^2} \right\} \frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}}$$
$$= -\mathbf{H}^T \mathbf{R}^{-1} [\mathbf{z} - \mathbf{h}(\mathbf{x})] = 0$$

Where, H is the measurement jacobian matrix, given by

$$\mathbf{H} = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

- Let us now find the Hessian matrix of the measurements, because it will be needed soon in our derivations.

$$\mathbf{G}(\mathbf{x}) = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x}^2}$$
$$= \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \left[\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}}\right] \left[\frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}}\right]^T - \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \{z_i - h_i(\mathbf{x})\} \frac{\partial^2 h_i(\mathbf{x})}{\partial \mathbf{x}^2}$$

- Neglecting the second order terms in the above equation, we can write

$$\mathbf{G}(\mathbf{x}) \approx \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$$

- Above is called the gain matrix for WLS state estimation.

- Now, expanding the elements in the vector g(x) in Taylor series around the state vector, x^k, at the k^{th} iteration, we get,

$$\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}^k) + \mathbf{G}(\mathbf{x}^k)\Delta \mathbf{x}^k + \ldots = 0$$

- Where, $\mathbf{x}^{k+1} - \mathbf{x}^k = -\left[\mathbf{G}(\mathbf{x}^k)\right]^{-1} \mathbf{g}(\mathbf{x}^k)$ is the increment in the state vector, given by,

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

- The iterative equation for WLS estimation, therefore, can be written as

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \left[\mathbf{G}(\mathbf{x}^k)\right]^{-1} \mathbf{g}(\mathbf{x}^k)$$
$$= \mathbf{x}^k + \left[\mathbf{G}(\mathbf{x}^k)\right]^{-1} \mathbf{H}^T \mathbf{R}^{-1} [\mathbf{z} - \mathbf{h}(\mathbf{x})]$$

- Inversion of the gain matrix in the above equation is a computationally demanding task. Efficient numerical methods, such as LU-factorization, Cholesky factorization, etc. are typically used to effectively invert the gain matrix, G.

- The estimation process starts with the following equation

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

Where, $\Delta\mathbf{x}^{k+1} = \mathbf{x}^{k+1} - \mathbf{x}^k$.

There are some set of procedure that you need to find out, follow, Steps for implementation:

- Step 1: Start iteration; set iteration index, k = 0.
- Step 2: Initialize the state vector, x. Typically, all the bus voltage magnitudes are set to ones, and bus angles are set to zeros.
- Step 3: Calculate the gain matrix, $G(x^k)$.
- Step 4: Calculate the right-hand side of the above equation

- Step 5: The gain matrix, $G(x^k)$, is decomposed, and is solved for $\Delta\mathbf{x}^{k+1}$.
- Step 6: If max $|\Delta\mathbf{x}^{k+1}| \leq \epsilon$, where ϵ is a predefined threshold, the iteration is taken to be converged.
- Step 7: If above is not satisfied, update, $\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}^{k+1}$, k = k + 1, and go to step 3.

What is R inverse? Basically this is a weighted matrix, weight matrix. I told you each measurement device will be given some weightage based on its previous track record.

Now how, let us say 1999 times out of 1000 it is giving a true value. Then you will be giving it a very high weight actually. The measurement from that device is considered to be a most accurate measurement. Let us say a specific measurement device which is giving true value 500 times out of 1000. It is 50 percentage weightage basically.

So that means you are giving less weightage for that specific measurement device. So it depends upon the previous history. So you have this entry already there because you have lot of data being collected for time, very long time. So you have this is a constant matrix. And you also know what is Hessian matrix that we have already defined.

So by knowing this you can find out the gain matrix. So once you have find out the gain matrix, you know what is H, this Hessian matrix. And this is the previous state values, voltage and angle in the previous iteration. And this is a weighted matrix that also you know. And Z minus, this is the Z is the actual measurement at this instant. And H of XK is that model, you know that we have model function that we have discussed, X1, X2 and all these values.

So you can now calculate the right hand side of the above equation. So the gain matrix G of XK is decomposed and it is solved for delta XK plus 1. Ultimately you need to find out delta XK plus 1. You know what is G of X, G of XK, you know what is the right hand side. Find out what is delta XK plus 1. Now if this maximum of this deviation, if it is less than the error, you have some threshold value already predefined.

The iteration takes to be converged. Now you assume that it is converged. Now if above is not satisfied, then update X of K plus 1 with XK plus delta XK plus 1 which you have got in the previous iteration. So again you follow the same procedure. So the gain matrix G of XK is rarely inverted directly to avoid numerical problems. Numerically efficient techniques such as Cholesky factorization or Gauss elimination are usually used to decompose the gain matrix.

So one example you can see here. This is just a DC power flow. This is a impedance and phase angles have been given here. You need to find out the phase angle basically. So you have got this power flow.

This is a measurement, measurement matrix Z. You have this line parameters. With this line parameters you are going to find out that H matrix basically. So you have to find out the state variables, theta 1, theta 2. So bus 3 is considered as a reference bus.

That means this angle is 0. You need to find out phase angles of bus 1 and bus 2 basically. So how do you do it? So Z1, Z2, this is the line flow measurement, actual measurement. And that is a dependence matrix. I am just expressing in the matrix form. That is nothing but 1 by X12.

That is Z12 is nothing but theta 1 minus theta 2 by X12 phase angle difference divide by reactance. Similarly you get Z13, you will get Z32. The objective is to find out theta 1 and theta 2, the states. How do you find out? Again follow the previous procedure. I gave the steps. So find out the gain matrix, hessian matrix and you have the measurement matrix and then by which you will find out the system states.

So then reasons for using state estimator. Why we are doing state estimation? State estimator can give the value of a line flow that is not measured. And state estimator can tell if one or more measurements are bad, that is detecting and identifying bad measurements. If you suddenly get a value which is true far from the variance, then what would be the actual value? So from previous track record can you able to find out the possible close value? So there state estimator will help.

So some drawback of SCADA based state estimator. The refresh rate of SCADA data is slow. Remote terminal units are scanned at every 2 to 5 second. The state estimator gives output every 1 to 5 minute. After running bad data detection, identification, identifying and processing, it takes a lot of time. So it is technically difficult to obtain the voltage phase angles in real time from widely dispersed location with respect to a common reference by using conventional measurements in the state estimator.

So basically the synchronization is not so easy. This has different time scale. So real time visualization of power system is therefore not possible by using this SCADA system. It is not actual real time. It takes a lot of time for the, you get the measurement, then you take a lot of time for state estimation, check whether the data is true or not, then use that measurement for load flow analysis.

This is a long procedure. So real time visualization of power system is very tedious thing. So phasor measurement technology elevates the above problem. Even you can see here PMUs can provide measurements at sub-second rate, typically up to 50 and 60 times per second for 50 and 60 hertz systems compared to 2 to 5 second scan rate of RTUs. So PMUs provide direct measurement of the phase angles which was technically infeasible so far with the existing SCADA based measurement. Earlier SCADA based measurements would not give you phase angle.

This technology offers wide area visibility. This in turn facilitates the capability for distributed sensing and coordinated control action. So controlling would be more faster. So PMUs are typically much more accurate than the conventional SCADA based measurements. So integration of PMUs into state estimator. So the design and deployment of the hybrid state estimator is still in the experimental stage and the following are some of the benefits that can be derived as a result of including PMU with the state estimator.

Improved accuracy of the state estimator because you are getting more samples. So more samples means more accuracy. Better observability, enhanced convergence characteristics, better handling of bad data. So PMU based state estimator will be better than SCADA based state estimator. Some source of error in state estimation. The quality of the measurements used in state estimator vary and the use of a state estimator gives engineers and power system operators a mean to cope with errors.

A partial list of error source are as follows. Modelling error, data error, transducer error and sampling error. These are the 4 typical errors that one can see. I will just give you a light about each of these. So modelling error. These have to do with the fact that the model of the transmission system may be incorrect.

You need model, the H of X matrix may be incorrect to begin with. Impedance data for transmission lines and transverse can be calculated wrong because there could be measure wrong in the calculation or may have drifted due to environmental factors. So transmission line may expand. There could be some practical scenarios where there could be different value such as ground moisture and all these things. Finally, modelling errors can be present because breaker and disconnect switch status indications can be wrong due to human error or due to telemetry error through the SCADA system.

There could be human error also which may be creating an issue. Telemetry error means the data which has been logged. So that is also there is some issue. So data error by this are meant the errors that come with incorrect data specifications that indicate which current transformer is connected to which transducer, which are the positive and the negative sides of the measurement devices, which breaker status contacts are voided to which terminal strip connections in the substation building etc.

So these are some of the cases by which you get data error. Sometimes you may correct a measuring device opposite. Instead of getting a positive measurement, you may get a negative measurement. Then transducer error. These are errors in the values transmitted from current transformers, potential transformers CT and PT and transducers that produce signals proportional to P and Q flows. These may be simple bias errors that give a constant offset to the measurement value or they may be random errors giving an error that is distributed around a mean value representing the true measurement value and there is sampling error. These are due to the fact that SCADA systems do not sample all measurements simultaneously nor all the measurements sampled at the same frequency.

Thus we are using measurements of time varying quantities that were taken at different times and can be in error simply due to dynamic nature, time changing values of the quantity being measured. Hence the SCADA based measurement is less accurate compared to PMU where PMU you get time sample data. So state estimator would be more accurate. So with this we will conclude state estimation and we will consider a new topic. Thank you very much. .