NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

UNIVERSITY OF ALBERTA

STATE ESTIMATION ALGORITHMS FOR FREQUENCY RELAYING IN POWER SYSTEMS

BY

© NAN LIU

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

FALL 1990

ISBN 0-315-65102-4

TO MY PARENTS

ABSTRACT

Frequency deviation from its rated value and its rate of change are indications of load imbalance in power systems. Frequency relays which detect frequency deviation and its rate of change and react accordingly are implanted in power systems to ensure the safe and efficient operation of power systems. This thesis describes a new algorithm for measuring the frequency, its deviation and the rate of change of frequency accurately even with the presence of bad data contamination, which is based on a recently developed least absolute value estimation technique. Compared with the least error square technique, which has been used widely, this new algorithm is shown to be a viable alternative. Moreover, under certain circumstances, the new technique proves to be superior to the existing least error squares technique.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Dr. G. S. Christensen and co-supervisor D. H. Kelly for their support, guidance and encouragement, without which this thesis would never has been possible. I also would like to thank Dr. Soliman for his valuable suggestions and frequent discussions and Mr. R.H. Rouhi for allowing me to use some of his programs to carry out part of the computation in this thesis.

I gratefully acknowledge the financial support that has been provided by the Department of Electrical Engineering, University of Alberta and the Natural Sciences and Engineering Research Council of Canada.

Last, but not the least, thanks are due to my family for their love, support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In the steady state operation of a power system, the total power generated is equal to the system load plus the losses, and the system frequency is a constant value, 60 Hz, which also is termed rated frequency. The huge rotating masses of turbine-generator rotors act as repositories of kinetic energy: when there is insufficient mechanical power input to the system, the rotors slow down, supplying energy to the system; conversely, when excess mechanical power is input, they speed up, absorbing energy. Any change in speed causes a proportional frequency change.

Under normal operating conditions, there is always a continuous frequency dither as the load may vary slightly from instant to instant, and the control of load sharing between machines requires some frequency variation. That is a change in load is always accompanied by a change in the frequency, which is normally corrected by a load-frequency controller. This dither is small, and to a great extent, is determined by the sensitivity of the governors on the prime movers and the load-frequency controllers [9].

Under abnormal conditions, such as a fault or loss of transmission capability, different parts of the system will experience a local generation-load imbalance. This leads to abnormal currents and disconnection of parts of the system. In extreme situations, transient instability may occur,

leading to the failures of the system (often referred as blackouts). There have been cases where transient instability, per se, has been avoided but failure has occured. The Northeast Power Pool (NEPP) blackout of 1965 is a prime example. In this failure, some areas were left with an excess of load over generation and the frequency dropped. As the frequency drops, the performance of motors,..., etc is degraded and, in this case, critical equipment, such as feed-water pumps for boilers was shut down. This in turn caused the shutdown of more generators, causing more severe generation-load imbalance. This eventually led to a nearly complete shutdown of generation between Niagara Falls and New York City. The whole process of failure took about 15 minutes, but the restoration took several days. The cost, both monetary and social were extremely high. This incident led directly to the implementation of underfrequency relays, which selectively shed load when an unacceptable frequency drop is experienced, to prevent the risk of uncontrolled generation shutdown [9].

1.1 Frequency Relays in Power Systems.

Frequency relays are classified into two groups, underfrequency relays and overfrequency relays according to their functions.

Underfrequency relays, as mentioned earlier in this chapter, are used to detect the frequencies a certain amount

lower than the rated frequency and disconnect some loads automatically to maintain the generation-load balance. This is called load shedding, in which discrete amounts of the demanded power are dropped in a preplanned order upon a detection of a power deficiency. In recent years, digitized underfrequency relays, which sense not only frequency deviation, but also its rate of change have been designed and built [1,2,3,4,6,8]. The advantage of knowing the rate of change of frequency is that it is then possible to predict if a critical frequency drop will occur, before it occurs, providing better selectivity.

Overfrequency relays, on the other hand, detect the frequencies above the rated value at the generator terminals, and decrease the generation or shut down some generating units, if necessary, to keep the nominal frequency to protect the generators from overspeeding during start-ups or faults which often cause loss of load.

1.2 Estimation of Frequency Deviation and Its Rate of Change

As explained in the previous section, frequency relays require accurate estimates of the frequency. Many techniques have been developed in the past twenty years to measure the voltage phasor parameters, including frequency, amplitude and phase angle. Some of these are based on the Discrete Fourier Transform (DFT), which tracks the local system frequency and its rate of change [3,4]. It is assumed that the input signal is a pure sine wave of a fundamental

frequency and must be band-limited to avoid errors due to aliasing. This technique may be affected by harmonics and the presence of bad data and white noise.

Another algorithm, which measures the frequency and the amplitude simultaneously from the sampled values of the system voltage at the relay location, is based on the Least Error Squares (LS) curve-fitting technique [1]. The LS technique finds the mean value of a set of measurements. It is known that the mean value is the best estimate when the set of measurements has a Gaussian error distribution, in other words, the noise present is pure white noise. However, for error distributions other than the Gaussian distribution, the LS technique will not produce the best estimate. This is especially true when the set of measurements is contaminated with bad data [9].

Two techniques have been developed for the optimal tracking of power system voltage amplitude and frequency deviation [6]. The first one uses a two-state linear Kalman Filter model (linear model) and the second one uses a three-state extended Kalman Filter model (non-linear model). Since the Kalman Filter was derived on the basis of LS approximations, eventually these two techniques face the same problem as the LS technique.

In this thesis, a newly developed technique for estimating the frequency deviation and its rate of change is presented. This algorithm is based upon the Least Absolute Value Algorithm (LAV), which minimizes the sum of the

absolute values of the errors. Two models of the voltage phasors are used, namely the Constant Frequency Model (CFM) and the Variable Frequency Model (VFM).

1.3 Outline of the Thesis.

In chapter II, the parameter estimation problem is introduced. Different algorithms are then discussed and compared. The advantages of the new LAV technique are shown through a few examples. Chapter III discusses the two mathematical models of the voltage phasors. Then, in chapter VI, the new LAV and other algorithms are tested in an off-line mode. Different factors which may affect the algorithms are explored. Advantages and disadvantages of different algorithms are compared. The LS and the LAV techniques use a set of measurements, which may be considered as a static approach. In chapter V, the use of the Kalman Filter, a dynamic approach, is presented as an alternative to the static ones. And finally, the conclusions are drawn in the last chapter.

CHAPTER II

PARAMETER ESTIMATION ALGORITHMS

Parameter estimation problem involves the estimation of *n* unknown system parameters from *m* measurements and a knowledge of the system structure. Equation (2.1) describes the relationship between the measurements and the system parameters.

$$\mathbf{z} = H\,\theta + \mathbf{v} \tag{2.1}$$

where $\mathbf{z}$ is the *m* by *1* (*m*>*n*) vector of measurements, θ is the *n* by *1* vector of parameters to be estimated, H is the *m* by *n* matrix which describes the mathematical relationship between the measurement vector $\mathbf{z}$ and the parameter vector θ in the absence of measurement error. Each element of $\mathbf{v}$ represents the error in a measurement.

The parameter estimation problem is to estimate the elements of the parameter vector θ, with the H matrix and the measurement vector $\mathbf{z}$ given, and the elements of $\mathbf{v}$ unknown. If *m* = *n*, an estimate of θ can be obtained by using $\theta = [H]^{-1}\mathbf{z}$. In this case, the parameter estimate fits exactly the measurement vector, i.e.., $\mathbf{z} - H\theta = \mathbf{v} = 0$, and is of poor quality, since this estimation process assumes that the error vector $\mathbf{v}$ contains only 0's. Therefore the estimate does not account for, or filter out measurement errors.

In most cases the number of measurements exceeds the number of system parameters, and we can obtain a good quality estimate according to the standard used.

Define an m by 1 vector $\mathbf{r}$ of residuals as

$$\mathbf{r} = \mathbf{z} - H\theta \tag{2.2}$$

Three common estimates of the parameter vector θ are l_1, l_2 and l_∞ estimates, which have cost functions to be minimized as follows [23]:

$$\text{For } l_1 \text{ estimation: } J_1(\theta) = \sum_{i=1}^{m} | z_i - H_i\theta | = \sum_{i=1}^{m} |r_i| \tag{2.3}$$

$$\text{For } l_2 \text{ estimation: } J_2(\theta) = \sum_{i=1}^{m} (z_i - H_i\theta)^2 = \sum_{i=1}^{m} (r_i)^2 \tag{2.4}$$

$$\text{For } l_\infty \text{ estimation: } J_\infty(\theta) = \lim_{k\to\infty} \left(\sum_{i=1}^{m} |z_i - H_i\theta|^k \right)^{1/k}$$

$$= \max|r_i| \tag{2.5}$$

where r_i, z_i, H_i are the rows of $\mathbf{r}$, $\mathbf{z}$, and H, corresponding to the ith measurement of m measurements.

The l_1 estimate minimizes the sum of the absolute values of the residuals. The l_2 estimate minimizes the sum of the squares of the residuals. The l_∞ estimate minimizes the largest absolute value of the residuals. The l_1 is called the least absolute value (LAV) estimate, the l_2 is called the least error squares (LS) estimate and l_∞ is called Chebyshev estimate.

For data $\mathbf{z}$ whose errors are normally distributed, l_2 is a good choice. If the data $\mathbf{z}$ are plagued with outliers, l_1 is a good choice because in l_1, outliers tend to be completely ignored. Finally, if the errors in $\mathbf{z}$ are negligible (e.g., if the residuals are expected to be much larger than

the errors in $\mathbf{z}$), then l_∞ will produce the smallest maximum deviation $|r_i|$ [23].

2.1 Least Error Squares Estimation (LS).

The first application of LS estimation took place in 1795 according to J.M. Mendel [10]. LS estimation has been applied to many estimation problems since then, and continues to be widely used.

It can be shown [21] that the LS estimation is given by

$$\theta^* = [H^T H]^{-1} H^T \mathbf{z} \tag{2.5}$$

In more general cases, weights are assigned to each measurement so that the measurements assigned with larger weights influence the least squares estimate more than those assigned with smaller weights. The estimates are then called the weighted least squares estimates. The cost function for weighted least squares estimation is

$$J_2(\theta) = \sum_{i=1}^{m} w_i (z_i - H_i \theta)^2 \tag{2.6}$$

where w_i is the weight assigned to the ith measurement.

Equation(2.6) can be rewritten as

$$J_2(\theta) = (\mathbf{z} - H\theta)^T W (\mathbf{z} - H\theta) \tag{2.7}$$

where W is a diagonal m by m matrix that contains weights, or $W = \text{diag.}(w_i,\ i=1,\ldots m)$. It can be shown [10] that the weighted least squares estimation is given by

$$\theta^* = [\ H^T W H\ ]^{-1} H^T W z \qquad (2.8)$$

The LS estimate possesses a number of interesting properties. First, the L_2 estimate gives the best estimation (maximum likelihood) when the measurement errors follow a Gaussian or normal distribution and the weighting matrix is equal to the inverse of the covariance matrix. It has been also stated [20] that, in cases where the measurement error distribution is not Gaussian, but the number of measurements greatly exceeds the number of unknown parameters, the method of least squares yields very good estimates.

Another valuable feature of the LS estimation is the ease with which the estimates can be calculated; equation (2.5) and (2.8) can be easily implemented on a computer. In contrast, no algorithm has yet been devised that makes the calculation of the LAV estimates as easy as that of the LS estimates.

There are many estimation problems for which the error distribution is not Gaussian and the number of measurements does not greatly exceed the number of unknown parameters. In these cases, least squares estimates are adversely affected by the presence of bad data. This problem has been recognized and addressed by several authors [11] who have proposed different ways of refining the least squares method so that l_2 estimates are less affected by bad data.

2.2 Least Absolute Value Estimation (LAV).

In contrast to the LS estimation technique, the LAV estimation technique has not been widely used for parameter estimation due to the following reasons [11].

(1) Computational difficulties in producing the numerical values of the LAV estimates in regression (lack of a closed form formula equivalent to that of the LS estimates).

(2) Lack of asymptotic theory for the LAV estimation in the regression model, and, in general, nonexistence of accompanying statistical inference procedures.

(3) Insufficient evidence to show the superiority of the small sample properties of the LAV as compared to the LS estimation when sampling from long tailed distributions.

2.2.1 Linear Programming Technique [12].

The most common approach used to calculate the LAV estimates is to first formulate the estimation problem as a linear programming problem and then to solve for the l_1 estimate using: the simplex method, the revised simplex method or a related technique.

A linear programming problem contains a cost function, which is to be maximized or minimized, and a set of constraints. The simplex method minimizes or maximizes the cost function within the bounds imposed by the constraints. The calculating procedure begins with an initial feasible solution which does not violate any of the constraints. Then new parameter values are selected which decrease the cost function. This is repeated until no new parameter values can be

found which decrease the cost function, then the solution of the linear programming problem is reached.

Following is a description of formulating LAV estimation as a linear programming problem [12].

Given an m by 1 measurement vector $\mathbf{z}$, and an n by 1 parameter vector θ, the LAV estimation problem is to estimate θ such that the cost function given by (2.9) is minimized

$$J_1(\theta) = \sum_{i=1}^{m} \left| z_i - \sum_{j=1}^{n} H_{ij}\theta_j \right| \tag{2.9}$$

where: z_i is the ith element of $\mathbf{z}$.

θ_j is the jth element of θ.

H_{ij} is the element in the ith row and the jth column of the m by n matrix H, which defines the relationship between $\mathbf{z}$ and θ.

Let
$$r_i = z_i - \sum_{j=1}^{n} H_{ij}\theta_j \quad i=1,2,\ldots,m \tag{2.10}$$

The linear programming formulation of the LAV estimation is as follows [12].

Minimize
$$\sum_{i=1}^{m} r_i \tag{2.11}$$

Subject to
$$r_i + \sum_{j=1}^{n} H_{ij}\theta_j > z_i \quad i = 1, 2, \ldots, m \tag{2.12}$$

$$r_i - \sum_{j=1}^{n} H_{ij}\theta_j > -z_i \quad i = 1, 2, \ldots, m \tag{2.13}$$

All the r_i's in the formulation will be non-negative because the minimum value that any r_i can reach, without vi-

olating a constraint, is the larger of expressions (2.14) and (2.15).

$$r_i > z_i - \sum_{j=1}^{n} H_{ij}\theta_j \tag{2.14}$$

$$r_i > \sum_{j=1}^{n} H_{ij}\theta_j - z_i \tag{2.15}$$

If one of the two expressions is negative, the other is positive, and r_i must be positive in order to satisfy both constraint expressions. If one of the two expressions equals zero, the other expression equals zero and therefore r_i is zero. Consequently r_i, when subject to the constraint expressions, represents the absolute value of the ith residual. Minimizing the sum of r_i's in the linear programming problem produces the LAV estimate.

Despite its popularity linear programming based LAV estimation has the following drawbacks [21].

1) It requires excessive memory storage; a typical formulation requires the manipulation of a matrix of size 2(mxn)xm [13], where m is the number of measurements and n is the number of unknown parameters.

2) It is an iterative technique and thus may use a considerable amount of CPU time and may be computational inefficient [14].

3) It may fail to identify the measurements with gross errors, leverage points [5].

4) The solution obtained may not be unique [15].

In the next section a new technique for LAV estimation is presented.

2.2.2 New LAV Technique.

A great deal of research in LAV estimation has involved attempts to prove the superiority of the l_1 estimate over the l_2 estimate and the development of more efficient linear programming based estimation algorithms. Very little research effort has been expended in finding and developing a relationship between the two types of estimates.

In 1973 Schlossmacher [13] presented an iterative technique which uses successive weighted least squares estimates to find an LAV estimate. His algorithm has the following steps [21]:

1) Obtain a weighted least squares estimate with all the weighting factors equal to one, i.e. { $w_i = 1$, $i = 1,2,\ldots,m$; m = the number of measurements }.

2) Use the generated weighted least squares solution to calculate the residuals r_i, $i = 1,2,\ldots,m$.

3) Set $w_i = 1/|r_i|$, $i = 1,2,\ldots,m$. If $r_i=0$, set $w_i = 0$.

4) Repeat step 2) and 3) until the changes in the r_i's between successive iterations approaches zero.

Although Schlossmacher's technique gives approximate estimates, it is an iterative technique and has been criticized as being computational inefficient [16].

In 1976-77, Sposito, Hand and McCormick suggested that the l_2 estimate may be used as starting point for linear programing based l_1 estimators [19]. Their research indi-

cates that starting a linear programming based l_1 estimator at the l_2 estimates saves many iterations. They found that in general the total computing CPU time (time to calculate the l_2 estimates to get the starting point + time to calculate l_1 estimates using linear programming formulation) of their technique is less than the time required to calculate an l_1 estimate from a flat start. The main drawback of their technique is that it still requires a linear programming algorithm to calculate the l_1 estimate.

In 1987, Christensen and Soliman [5] developed a new l_1 estimation technique. Their technique, which does not use linear programming, manipulates a simple relationship between the l_1 and l_2 estimates. It is also non-iterative.

In the rest of this section the new LAV technique developed by Christensen and Soliman is given in detail.

The new estimation technique utilizes the interpolation property that is stated in theorem 2.1.

Theorem 2.1 [18]

> If the column rank of the mxn matrix H is k ($k \leq n$), then the l_1 estimate interpolates at least k of the m measurements.

Since LAV estimates interpolate measurements, the estimation problem reduces to selecting the k points that the estimates should interpolate. The new technique assumes that H has full rank and therefore $k = n$. Given the measurement

equation (2.16), the first step of the new LAV algorithm is to calculate the least squares estimate of θ which is given by θ^* in equation (2.17).

$$\mathbf{z} = H\theta + \mathbf{v} \tag{2.16}$$

$$\theta^* = [\, H^T H \,]^{-1} H^T \mathbf{z} \tag{2.17}$$

The residuals of the l_2 estimate are calculated using equation (2.18)

$$r_i = z_i - H_i \theta^* \tag{2.18}$$

where r_i is the ith residual, z_i is the ith measurement, and H_i is the ith row of H.

Outlier measurements, assuming there are *p* of them and m-p > n, whose residuals are greater than the standard deviation of the LS solution residuals σ, are then rejected. This is based on the facts that the parameter σ is a measure of the amount of variability inherent in the regression model and a large value of σ will lead to observed (x_i, y_i)'s which are quite spread out about the true regression line, whereas when σ is small the observed points will tend to fall very close to the true line [23]. The standard deviation σ is given by

$$\sigma^2 = \frac{1}{m-n+1} \sum_{i=1}^{m} (r_i - \bar{r})^2 \quad .$$

The introduction of rejecting outliers according to the absolute values of the residuals is a improvement over the old version of this technique as described in the reference [21]. Some of the difficulties encountered in [21] now are solved with ease. The readers may refer to example 2.2 for details.

Recalculate the LS solution from the measurements of the remaining measurements and a new LS estimate $\overset{*}{\theta}_{new}$ is obtained.Use this new LS estimate to calculate the corresponding new residual vector $\mathbf{r}_{new}$.

The residuals are then ranked by their absolute values and then stored in the (m-p)x1 vector $\mathbf{r}$ with the smallest residual as r_1 and the largest one as r_{m-p}.

$$\mathbf{r}_{new} = \begin{bmatrix} r_1 \\ \cdot \\ \cdot \\ \cdot \\ r_{m-p} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{r}} \\ --- \\ \overset{*}{\mathbf{r}} \end{bmatrix} \begin{matrix} -- \\ nx1 \\ -- \\ (m-p-n)x1 \\ -- \end{matrix} \quad \begin{matrix} -- \\ \\ (m-p)x1 \\ \\ -- \end{matrix} \tag{2.19}$$

The rows of the $\mathbf{z}$ vector and the H matrix are also rearranged so that all the z's and all the rows of H correspond to the ranking of the absolute values of the residuals (2.20) and (2.21)

$$\mathbf{z}_{new} = \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_{m-p} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{z}} \\ --- \\ \overset{*}{\mathbf{z}} \end{bmatrix} \begin{matrix} --- \\ nx1 \\ --- \\ (m-p-n)x1 \\ --- \end{matrix} \tag{2.20}$$

$$H_{new} = \begin{bmatrix} H_1 \\ . \\ . \\ . \\ H_{m-p} \end{bmatrix} = \begin{bmatrix} \hat{H} \\ --- \\ * \\ H \end{bmatrix} \begin{matrix} --- \\ nxn \\ --- \\ (m-p-n)xn \\ --- \end{matrix} \qquad (2.21)$$

The sub-vectors $\hat{r}$, $\hat{z}$ and the sub-matrix $\hat{H}$ correspond to the n smallest residuals.

The final step of the algorithm is to calculate the LAV estimate from equation (2.22) and (2.23)

$$\hat{z} = \hat{H}\hat{\theta} \qquad (2.22)$$

$$\hat{\theta} = [\ \hat{H}\ ]^{-1}\hat{z} \qquad (2.23)$$

Once the final estimate is calculated, the LAV measurement residuals can be calculated using equation (2.24).

$$\hat{r} = z - H\hat{\theta} \qquad (2.24)$$

From equation (2.22) to (2.23), it can be easily seen that the first n residuals of equation (2.24) will be equal to zero. The n-m remaining residuals will either be zero or non-zero. The five steps in the new LAV algorithm can be summarized as follows.

1) Calculate the least squares estimate of θ using

$$\overset{*}{\theta} = [\ H^T H\ ]^{-1} H^T z$$

2) Calculate the residuals of the l_2 estimate using

$$\overset{*}{r}_i = z_i - H_i\overset{*}{\theta}, \qquad i=1,2,\ldots,m$$

3) Reject the outliers whose residuals are greater than the standard deviation of residual vector σ.Recalculate the LS solution and the associated residuals.

4) Select the n measurements that correspond to the residuals with the smallest absolute values and form $\hat{\mathbf{z}}$ and $\hat{H}$.

5) Solve for the LAV solution $\hat{\theta}$ using equation (2.23).

A useful property of the l_1 estimate is given by theorem 2.2.

Theorem 2.2 [19]

> For a least absolute value estimate of n parameters, if N_1 is the number of positive residuals and N_2 is the number of negative residuals, then an optimal least absolute value estimate $\hat{\theta}$ obeys the following equation, provided that no set of n+1 points lies on an optimal hyperplane in n dimensions

$$| N_1 - N_2 | < n \qquad (2.25)$$

Theorem 2.2 gives necessary, but not sufficient conditions for optimality. Thus l_1 estimates produced by the new LAV technique, or by any other technique, can be checked by using theorem 2.2.

The following example illustrates the application of the new technique.

Example 2.1. Fit the data points { (1,1), (4,5), 2,0), (2,2), (3,1.5), (3,2.5), (4,2), (5,3) } with a straight line in the form of $y = a_1 x + a_2$.

In this example, we have

$$z = \begin{bmatrix} 1.0 \\ 5.0 \\ 0.0 \\ 2.0 \\ 1.5 \\ 2.5 \\ 2.0 \\ 3.0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 1 \\ 4 & 1 \\ 2 & 1 \\ 2 & 1 \\ 3 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}, \quad \theta = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Step 1. Calculate the LS solution as

$$\hat{\theta}^* = [\, H^T H \,]^{-1} z$$

$$= \begin{bmatrix} 0.75 \\ -0.215 \end{bmatrix}$$

Step 2. Calculate the LS residuals as

$$\hat{r}^* = \begin{bmatrix} 0.375 \\ 2.125 \\ -1.375 \\ 0.625 \\ -0.625 \\ 0.375 \\ -0.875 \\ -0.625 \end{bmatrix}, \quad \text{with SD, } \sigma = 1.11$$

Step 3. Measurement #2 and #3 have residuals whose absolute values are greater than the standard deviation $\sigma = 1.1$. Reject these two measurements. The new LS solution is obtained

$$\hat{\theta}^*_{new} = \begin{bmatrix} 0.4 \\ 0.8 \end{bmatrix}$$

and the new LS residuals are

$$\overset{*}{\mathbf{r}}_{new} = \begin{bmatrix} -0.2 \\ 0.4 \\ -0.5 \\ 0.5 \\ -0.4 \\ 0.2 \end{bmatrix}.$$

The rank of the matrix H is two, hence, the estimator fits at least two measurements, and these two measurements have the smallest residuals. They are #1 and #6 of the new measurements set.

Step 4. Solve for the LAV solution.

$$\hat{\mathbf{H}} = \begin{bmatrix} 1 & 1 \\ 5 & 1 \end{bmatrix}, \quad \hat{\mathbf{z}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

which gives

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}.$$

and the straight line equation is

$$y = 0.5x + 0.5$$

Step 5. The LAV residuals are:

$$\mathbf{r} = \begin{bmatrix} 0 \\ 2.5 \\ -1.5 \\ 0.5 \\ -0.5 \\ 0.5 \\ -0.5 \\ 0 \end{bmatrix}.$$

with cost function $J = \sum_{i=1}^{8} |r_i| = 6$, $N_1 = 3$, and $N_2 = 3$. $| N_1 - N_2 | = 0 < n = 2$, so theorem 2.2 is satisfied. We can get the same solution by using linear programming (IMSL subroutine).

For cases where the absolute values of two residuals are equal, but there is only one place left in the to-be-interpolated measurement set, a tie-breaking procedure is implemented. Two LAV estimates, which correspond to the two different interpolated sets of measurements are calculated. Each interpolated set contains nearly the same elements, with the only difference being that that each set contains a different one of the two measurements that was involved in the tie. The estimate with the smaller cost function is the unique LAV estimate.

The next example illustrates the tie-breaking procedure.

Example 2.2. Fit the data points { (1,4.5), (2,7.0), (3,5.0), (4,6.0) } with a straight line y = ax + b

In this example, we have

$$\mathbf{z} = \begin{bmatrix} 4.5 \\ 7.0 \\ 5.0 \\ 6.0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \quad \theta = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Step 1. Calculate the LS solution.

$$\hat{\theta}^{*} = [\, H^T H \,]^{-1} H^T \mathbf{z}$$

$$= \begin{bmatrix} 0.25 \\ 5.00 \end{bmatrix}$$

Step 2. Calculate the LS residuals.

$$\overset{*}{\mathbf{r}} = \begin{bmatrix} -0.75 \\ 1.50 \\ -0.75 \\ 0.00 \end{bmatrix},$$

Step 3. Select the two measurements the correspond to the two smallest residuals and form $\hat{\mathbf{z}}$ and $\hat{\mathbf{H}}$.

The smallest residual is r_4. The absolute values of the 1st and the 3rd residuals are equal. Thus the tie-breaking procedure must be implemented and two estimates must be calculated.

The first estimate interpolates the 4th and the 1st measurements.

$$\hat{\mathbf{H}}_1 = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}, \quad \hat{\mathbf{z}}_1 = \begin{bmatrix} 6.0 \\ 4.5 \end{bmatrix}.$$

The other estimate interpolates the 4th and the 3rd data points.

$$\hat{\mathbf{H}}_2 = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}, \quad \hat{\mathbf{z}}_2 = \begin{bmatrix} 6.0 \\ 5.0 \end{bmatrix}.$$

Step 4. Solve for both LAV estimates.

$$\hat{\boldsymbol{\theta}}_1 = \begin{bmatrix} 0.5 \\ 4.0 \end{bmatrix}$$

$$J_1 = 2.5$$

and

$$\hat{\theta}_2 = \begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix}$$

$$J_2 = 4.5$$

The cost function corresponding to $\hat{\theta}_1$ is less than that corresponding to $\hat{\theta}_2$. Therefore $\hat{\theta}_1$ is the unique LAV estimate produced by the new technique.

It may has been noticed that the standard deviation of r^* is $\sigma = 1.061$. Measurement #2 can be rejected as explained earlier. The same optimal solution can be obtained without going through the tie-breaking procedure. The only purpose of this example is to show how to implement the tie-breaking equation in case it happens. On the contrary, if measurement #2 is not rejected according to the standard deviation criteria, as done in [21], the tie-breaking procedure is inevitable.

The following is a summary of the features the LAV possesses that make it an attractive alternative to linear programming approach.

1) It is non-iterative in nature.

2) It does not require an initial feasible solution (for linear cases), whereas the linear programming approach must first calculate an initial guess and then solve for the l_1 solution.

3) In most cases it produces a unique solution.

4) The new estimates are easily calculated and can often be carried out by hand for simple cases; whereas the

linear programming estimation almost always requires a computer.

CHAPTER III

MATHEMATICAL MODELS FOR FREQUENCY ESTIMATION

In the previous chapter, three parameter estimation techniques have been discussed, the LS approach, the linear programming based LAV approach, and the new LAV approach. This chapter presents the two models to be used to test the three estimation algorithms. They are the Constant Frequency Model (CFM) and the Variable frequency model (VFM).

3.1 Constant frequency model [1,2].

This model assumes that the voltage magnitude and frequency do not change during the data window used for measurement and it is sampled at the relay location. The system signal voltage can be written as:

$$v(t)=V\sin(2\pi ft+\phi) \tag{3.1}$$

where V is the peak value of the voltage, $\omega=2\pi f$, is the radian frequency, t is the time in seconds and ϕ is an arbitrary phase angle. Equation (2.1) can be rewritten as:

$$v(t)=(V\cos\phi)\sin 2\pi ft+(V\sin\phi)\cos 2\pi ft \tag{3.2}$$

Using the first four terms of the Taylor series $\sin 2\pi ft$ and $\cos 2\pi ft$ can be expanded in the neighborhood of f_0 as

$$\sin 2\pi ft=\sin 2\pi f_0 t+(2\pi t)\Delta f\cos 2\pi f_0 t-2\pi^2 t^2(\Delta f)^2\sin 2\pi f_0 t$$

$$-\frac{4}{3}\pi^3 t^3 (\Delta f)^3 \cos 2\pi f_0 t + R_{4\sin} \qquad (3.3)$$

$$\cos 2\pi ft = \cos 2\pi f_0 t + (2\pi t)\Delta f \sin 2\pi f_0 t - 2\pi^2 t^2 (\Delta f)^2 \cos 2\pi f_0 t$$
$$+\frac{4}{3}\pi^3 t^3 (\Delta f)^3 \sin 2\pi f_0 t + R_{4\cos} \qquad (3.4)$$

where Δf is the frequency deviation = $f-f_0$.

$R_{4\sin} = \frac{2}{3}\pi^4 t^4 (\Delta f)^4 \sin 2\pi(f_0+\alpha\Delta f)t$, the remainder term of the expansion of the sine function. $0<\alpha<1$. (3.3a)

$R_{4\cos} = \frac{2}{3}\pi^4 t^4 (\Delta f)^4 \cos 2\pi(f_0+\alpha\Delta f)t$, the remainder term of the expansion of the cosine function. $0<\alpha<1$. (3.4a)

Substituting the first four terms of equations (3.3) and (3.4) in equation (3.2), we obtain

$$v(t) = a_{11}(t)x_1+a_{12}(t)x_2+a_{13}(t)x_3+a_{14}(t)x_4+a_{15}(t)x_5 \qquad (3.5)$$
$$+a_{16}(t)x_6+a_{17}(t)x_7+a_{18}(t)x_8$$

where the following parameters are defined at time t as

$$a_{11}(t)=\sin 2\pi f_0 t, \qquad a_{12}(t)=2\pi t\cos 2\pi f_0 t$$
$$a_{13}(t)=\cos 2\pi f_0 t, \qquad a_{14}(t)=2\pi t\sin 2\pi f_0 t \qquad (3.6)$$
$$a_{15}(t)=-2\pi^2 t^2 \sin 2\pi f_0 t, \qquad a_{16}(t)=-2\pi^2 t^2 \cos 2\pi f_0 t$$
$$a_{17}(t)=-\frac{4}{3}\pi^3 t^3 \cos 2\pi f_0 t, \qquad a_{18}(t)=\frac{4}{3}\pi^3 t^3 \sin 2\pi f_0 t$$

and the state variables x_1 to x_8 are defined as

$$x_1 = V\cos\theta, \qquad x_2 = \Delta f V\cos\theta$$
$$x_3 = V\sin\theta, \qquad x_4 = -\Delta f V\sin\theta \qquad (3.7)$$

$$x_5 = (\Delta f)^2 V\cos\theta, \qquad x_6 = (\Delta f)^2 V\sin\theta$$
$$x_7 = (\Delta f)^3 V\cos\theta, \qquad x_8 = (\Delta f)^3 V\sin\theta$$

If the voltage is sampled at a pre-selected rate, we would obtain its samples at equal time interval, say Δt seconds. A set of m samples designated as $v(t_1)$, $v(t_2)$, ..., $v(t_m)$ is obtained. These are the digitized samples of the voltage, where $t_2 = t_1 + \Delta t$, $t_3 = t_1 + 2\Delta t$, ..., $t_m = t_1 + (m-1)\Delta t$ and t_1 is an arbitrary time reference. Equation (3.5) can then be written in matrix form as follows

$$\mathbf{z}(t) = A(t)\mathbf{x} + \mathbf{v}(t) \qquad (3.8)$$

where $\mathbf{z}(t)$ is an mx1 voltage samples vector, A(t) is an mx8 measurements matrix (in chapter two we use H), $\mathbf{x}$ is an 8x1 parameters vector to be estimated and $\mathbf{v}(t)$ is an mx1 error vector to be minimized.

The elements of the matrix A(t) depend on the time reference t_1 and the sampling rate Δt and can be pre-determined in an off-line mode. Also, all the x's of equation (3.8) are unknowns, so at least eight equations must be established. In other words, at least eight samples of the voltage would be required.

As a general case, let us assume that m samples are available, m>8, then equation (3.8) describes an over-determined system of equations. As described in the previous chapter, two techniques are used to solve this system. One

is the least error squares (LS) technique. The second technique, the new LAV technique, which is based on the least absolute value (LAV) approximations, uses some information from the LS solution. This second technique was developed in [5].

Having obtained the elements of the vector $\mathbf{x}$, the amplitude and frequency of the voltage signal can be derived from the equations established. For example, the amplitude of the sampled voltage can be estimated as

$$V = (x_1^2+x_3^2)^{\frac{1}{2}} \tag{3.9}$$

The frequency deviations Δf can be calculated as

$$\Delta f = \frac{x_2}{x_1} \tag{3.10}$$

Another possible approach is to estimate frequency deviation by using variables x_3 and x_4 as follows

$$\Delta f = -\frac{x_4}{x_3} \tag{3.11}$$

When the value $V\cos\theta$ is small, the frequency deviation estimated by equation (3.10) is not sufficiently accurate. Similarly, when the value of $V\sin\theta$ is small, the frequency deviation estimated by equation (3.11) is not sufficiently accurate. However, $\sin\theta$ is close to 1 p.u. when $\cos\theta$ is

small, and $\cos\theta$ is close to 1 p.u. when $V\sin\theta$ is small .So we can always use variables x_1, x_2, x_3 and x_4 to estimate frequency deviations as follows

$$(\Delta f)^2 = \frac{x_2^2+x_4^2}{x_1^2+x_3^2} \tag{3.12}$$

or use variables x_5, x_6, x_7 and x_8 as

$$(\Delta f)^2 = \frac{x_7^2+x_8^2}{x_5^2+x_6^2} \tag{3.13}$$

Note that the computation requirement for implementing equation (3.12) or (3.13) is higher than that for implementing equation (3.10) or (3.11).

Furthermore, the phase angle θ can be obtained from one of the following equations

$$\tan\theta = \frac{x_3}{x_1} = -\frac{x_4}{x_2} = \frac{x_6}{x_5} = \frac{x_8}{x_7} \tag{3.14}$$

3.2. The Variable Frequency Model [8].

When developing the CFM in the previous section, the frequency was assumed to be constant during a data sampling window. In practice, however, it may change at a rate depending on the situation the system may be in. If we use a very small window size, say 2 cycles or 0.0167 seconds, then the frequency change can be linearized on the starting point of the measurement sampling period. Also we assume that the amplitude of the voltage remains constant during each data

sampling window. Using these two assumptions, a voltage at a power system bus can be written as

$$v(t) = V\sin(\phi(t) + \theta) \tag{3.15}$$

Also

$$\omega(t) = 2\pi f(t) = \frac{d\phi(t)}{dt} \tag{3.16}$$

which can be written as

$$\phi(t) = \int_0^t 2\pi f(t)\,dt \tag{3.17}$$

We assume that the frequency changes linearly with time, and this can be expressed as

$$f(t) = a + bt \tag{3.18}$$

where a is the frequency in Hz at t = 0, b is the rate of change of frequency in Hz/sec. Substituting equation (3.18) into equation (3.17) for f(t), we obtain

$$\phi(t) = 2\pi (at + \tfrac{1}{2}bt^2) \tag{3.19}$$

Substituting for $\phi(t)$ from equation(3.19) into equation (3.15), we obtain

$$v(t) = V\sin (2\pi at + \pi bt^2 + \theta) \tag{3.20}$$

Note that if the frequency is constant during a data sampling window, we have a = f and b = 0. In this case

equation(3.20) becomes equal to equation(3.1). Equation (3.20) can be written as

$$v(t) = (V\cos\theta)\sin(2\pi at + b\pi t^2) + (V\sin\theta)\cos(2\pi at + b\pi t^2) \quad (3.21)$$

Using a Taylor series expansion, the above equation can be expanded in the neighborhood of $a = a_0$ and $b = b_0$ as

$$\begin{aligned} v(t) &= \sin(2\pi a_0 t)(V\cos\theta + 2\pi t\cos(2\pi a_0 t)((a - a_0)V\cos\theta) \\ &+ \cos(2\pi a_0 t)(V\sin\theta) - 2\pi t\sin(2\pi a_0 t)((a - a_0)V\sin\theta) \\ &+ \pi t^2\sin(2\pi a_0 t)(-bV\sin\theta - 2\pi(a - a_0)^2 V\cos\theta) \\ &+ \pi t^2\cos(2\pi a_0 t)(bV\cos\theta - 2\pi(a - a_0)^2 V\sin\theta) \\ &- 2\pi^2 t^3\sin(2\pi a_0 t)(b(a - a_0)V\cos\theta) \\ &- 2\pi^2 t^3\cos(2a_{\pi 0} t)(b(a - a_0)V\sin\theta \\ &- \frac{1}{2}(\pi t^2)^2\sin(2\pi a_0 t)(b^2 V\cos\theta) \\ &- \frac{1}{2}(\pi t^2)^2\cos(2\pi a_0 t)(b^2 V\sin\theta) \end{aligned} \quad (3.22)$$

Define the following state variables

$$\begin{aligned} y_1 &= V\cos\theta, & y_2 &= (a-a_0)V\cos\theta \\ y_3 &= V\sin\theta, & y_4 &= (a-a_0)V\sin\theta \\ y_5 &= -bV\sin\theta - 2\pi(a - a_0)V\cos\theta \\ y_6 &= bV\cos\theta - 2\pi(a - a_0)V\sin\theta \\ y_7 &= b(a - a_0)V\sin\theta \\ y_8 &= b(a - a_0)V\sin\theta \\ y_9 &= bV\cos\theta \\ y_{10} &= bV\sin\theta \end{aligned} \quad (3.23)$$

Also define the following parameters:

$$b_{11}(t) = \sin(2\pi a_0 t) \qquad b_{12}(t) = 2\pi t\cos(2\pi a_0 t)$$
$$b_{13}(t) = \cos(2\pi a_0 t) \qquad b_{14}(t) = -2\pi t\sin(2\pi a_0 t)$$
$$b_{15}(t) = \pi t\sin(2\pi a_0 t) \qquad b_{16}(t) = \pi t\cos(2\pi a_0 t)$$
$$b_{17}(t) = -2\pi t\sin(2\pi a_0 t) \qquad b_{18}(t) = -2\pi t\cos(2\pi a_0 t)$$
$$b_{19}(t) = -1/2(\pi t)\sin(2\pi a_0 t) \qquad b_{110}(t) = -(\pi t)\cos(2\pi a_0 t) \quad (3.24)$$

Then equation (3.22) can be written as

$$v(t) = b_{11}(t)y_1 + b_{12}(t)y_2 + \ldots\ldots + b_{110}(t)y_{10} \quad (3.25)$$

In this equation, y's are the unknowns to be estimated and they are functions of V, a,b, and θ and b's are the coefficients that can be evaluated by arbitrarily selecting the time reference.

If the voltage is sampled at a pre-selected rate, we would obtain its samples at equal time intervala, say Δt seconds. A set of m samples may be designated as $v(t_1)$, $v(t_2),\ldots,\ldots,v(t_m)$ is obtained. These are the digitized samples of the voltage. While $t_2 = t_1 + \Delta t$, $t_3 = t_1 + 2\Delta t$, $\ldots,\ldots$, $t_m = t_1 + (m - 1)\Delta t$ and t_1 is an arbitrary time reference. Equation (3.25) can be written in a way similar to equation (3.8) as

$$\mathbf{z}(t) = B(t)\ \mathbf{y} + \mathbf{w}(t) \quad (3.26)$$

The elements of the matrix B(t) depend on the time reference and the sampling interval Δt. These elements can be calculated in an off-line mode after the sampling rate and time reference have been selected.

Having obtained the elements of the vector y, the amplitude, frequency and rate of change of frequency of the voltage can be calculated as follows

$$V = (y_1^2 + y_3^2)^{1/2} \tag{3.27}$$

The frequency deviations can be calculated by using the following two alternative equations

$$(a - a_0) = \frac{y_2}{y_1} \quad \text{and} \tag{3.28}$$

$$(a - a_0) = \frac{y_4}{y_3} \tag{3.29}$$

We can also use all variables y_1, y_2, y_3 and y_4 to estimate the frequency deviation as

$$(a - a_0) = \sqrt{\frac{y_2^2+y_4^2}{y_1^2+y_3^2}} \tag{3.30}$$

If the value of $V\cos\theta$ is very small ($y_1 << 0$), the frequency deviations estimated from equation(3.28) will not be accurate. Also , if the value of $V\sin\theta$ is very small ($y_3 << 0$), the the frequency deviations estimated by using equation (3.29) will not be accurate. However, when $V\cos\theta$ is very small, $\sin\theta$ is close to 1 p.u.; when $V\sin\theta$ is very

small, $\cos\theta$ is close to 1 p.u.. So equation (3.30) can always be used with satisfactory accuracy.

The rate of change of frequency can be estimated as

$$b = \frac{y_7}{y_2} \qquad \text{and} \qquad (3.31)$$

$$b = \frac{y_8}{y_4} \qquad (3.32)$$

In the above two equations, y_2 and y_4 are very small when frequency has its nominal value or is in the neighborhood of its nominal value. Therefore, these two equations are not suitable for estimating the rate of change of frequency.

One can also use the following equations to estimate b

$$b = \left(\frac{y_9}{y_1}\right)^{\frac{1}{2}} \qquad \text{and} \qquad (3.33)$$

$$b = \left(\frac{y_{10}}{y_3}\right)^{\frac{1}{2}} \qquad (3.34)$$

which will run into similar difficulties when either of $V\sin\theta$ or $V\cos\theta$ is very small.

All these problems can be solved by using variables y_1, y_3, y_5 and y_6 by using the following equation

$$b = \frac{y_6 y_1 - y_5 y_3}{y_1^2 + y_3^2} \qquad (3.35)$$

3.3. The Rate of Change of Frequency.

As explained earlier, the power deficiency in the power systems is characterized by a drop in the frequency. Therefore, it has been a common practice for many years to shed the load with the assistance of frequency relays until a balance is restored and frequency back to its nominal value. In such cases, the load is disconnected in steps [17]. However, in cases of severe disturbances where frequency drops rapidly, the use of frequency level alone to decide when and how much of the load to be disconnected may be inadequate in that, due to system time constants, the action may be too late to be beneficial. Also, the use of the rate of change of frequency allows prediction of the frequency at a reasonable time in the future and allows anticipatory action. In addition, it is possible to recognize those situations where no shedding is required.

The VFM can be used to measure the rate of change of frequency as explained earlier in the previous section. In this section, we explain two alternative approaches using the CFM algorithm to estimate the rate of change of frequency. The first one is the direct approach, while the second one is the linear regression approach.

3.3.1 The Direct Approach.

This approach uses the frequency estimated by the CFM algorithm to determine the rate of change of frequency. If the frequency of a voltage changes from f_1 to f_2 in ΔT, the average value can be calculated as

$$f' = \frac{f_1 - f_2}{\Delta T} \text{ Hz/sec} \qquad (3.36)$$

Using the CFM may produce an inaccurate estimation for f' for the reasons explained earlier when f_1 and f_2 are estimated using different equations. Only equation (3.12) or (3.13) are accurate for estimation of f_1 and f_2.

3.3.2 Linear Regression Approach.

If the frequency of a signal is measured m times at regular intervals ΔT second, and the frequency can be expressed as a function of the time interval ΔT as follows

$$f = a + bk\Delta T(k=1,2,.....,m) \qquad (3.37)$$

where a is the estimate of the frequency at t = 0, and b is the estimate of the rate of change of frequency. These coefficients are to be estimated using the two proposed techniques. The above equation can be written as

$$\mathbf{z} = H\theta + \mathbf{v} \qquad (3.38)$$

where $\mathbf{z}$ is an mx1 measurement vector and is given by

$$\mathbf{z} = \text{col}(f_1, f_2, ..., ..., f_m), \quad (m>n) \qquad (3.39)$$

θ is an nx1 (n = 2) parameter vector to be estimated.

H is an mxn (m > n) matrix.

$\mathbf{v}$ is an mx1 error vector to be minimized.

CHAPTER IV

TESTING OF LS AND LAV ALGORITHMS FOR FREQUENCY ESTIMATION IN AN OFF-LINE MODE.

In this chapter, the LS, the LP based LAV and the new LAV are tested in the off-line mode for both CFM and VFM models to study the effects of data window size, sampling rate, time reference location and number of terms truncated from the Taylor series expansion. For this purpose, the voltage and frequency measurements were simulated in a software program. For CFM, the program generates a voltage having a constant frequency which is sampled at the pre-selected rates; for VFM, the program generates a voltage having a variable frequency, and we assume that this variation is proportional to the time as given by equation (3.18) and this voltage is sampled at a pre-selected rate. These digitized voltage samples are then used to calculate the steady state voltage phasor magnitude, the frequency deviation from the nominal frequency and the rate of change of frequency. The three algorithms, LS, LP based LAV and new LAV, are tested both for voltage samples containing no bad data, and for voltage samples containing bad data (gross error and/or gaussian noise).

4.1 Effect of Data Window Size.

Table 4.1 gives the estimated voltage and frequency deviation for different window sizes for the CFM model. It can be noticed from this table that, if the voltage samples

do not contain bad data the three algorithms give the same estimate for the phasor magnitude (= 1.4142) and the frequency deviation ($\Delta f = 0$), also it can be noticed that the sum of the absolute values of the residuals for the LAV technique is equal to that of the LP technique, which is supposed to generate the optimal estimate based on least absolute value deviation. The optimal estimate using the proposed technique is produced with no iterations, while it takes 8 to 16 iterations using the LP technique. On the other hand, when the voltage samples contain bad data (the location of bad data points are given in the table, bad data points are obtained just by reversing the sign of sample measurements marked in the table (gross error)), the new LAV technique as well as the LP technique give better estimates than the LS technique. The estimate obtained by the LS technique is improved as the number of samples increases (data window size increases). Also the new proposed technique gives the optimal estimate with no iterations, but the LP gives the optimal estimate in the range of 8 to 17 iterations, which is time consuming.

Table 4.2 gives the estimated voltage, the steady state frequency deviation and the rate of change of frequency for a VFM model for different window sizes when the voltage samples do not contain bad data and contain bad data. For no bad data contamination, the three algorithms give almost the same results. However, when the voltage samples are contaminated with the bad data, the new proposed algorithm as well

as the LP give the same optimal estimate, while the LS algorithm gives very poor estimates. Also, the new LAV algorithm gives the optimal estimate with no iterations, while the LP gives the optimal estimate in the range of 15 to 32 iterations.

Studying Table 4.1 and 4.2 reveals that the new proposed technique is superior to either the LS algorithm which is adversely affected by the presence of bad data or the LP algorithm which requires a large number of iterations.

Table 4.1, The estimated phasor voltage, and the frequency deviation, $\Delta f = f - f_0$, $f_0 = 60$ Hz, for sampling frequency = 720 Hz, with the time reference at the middle of the data window and the number of parameters to be estimated; x = 8. Constant frequency model CFM.

Case number			1	2	3	4	5	6
No. of samples			12	24	36	48	60	72
Window size (msec.)			16.67	33.33	50.00	66.67	83.33	100
	LS	V	1.414	1.414	1.414	1.414	1.414	1.414
No		Δf	0.000	0.000	0.000	0.000	0.000	0.000
	LAV	V	1.414	1.414	1.414	1.414	1.414	1.414
Bad		Δf	0.000	0.000	0.000	0.000	0.000	0.000
	LP	V	1.414	1.414	1.414	1.414	1.414	1.414
Data		Δf	0.000	0.000	0.000	0.000	0.000	0.000
	LS	V	1.360	0.914	1.464	1.351	1.375	1.349
With		Δf	137.0	14.18	1.590	2.110	1.150	1.103
	LAV	V	1.416	1.414	1.414	1.414	1.414	1.414
		Δf	0.062	0.002	0.002	0.000	0.000	0.000
Bad		$\Sigma\lvert r_i \rvert$	2.730	2.730	2.730	2.730	3.470	5.840
	LP	V	1.414	1.414	1.414	1.414	1.414	1.414
		Δf	0.000	0.000	0.000	0.000	0.000	0.000
Data		$\Sigma\lvert r_i \rvert$	2.730	2.730	2.730	2.730	3.470	5.460
# of iter.			10	8	17	11	12	13
Location of Bad Data(#)			10	10	1, 2	1,11	1,10	1,10,20

Table 4.2a. The estimated phasor voltage, the frequency deviation and the rate of change of frequency, with the time reference shifted to the left of the middle of the data window by $3\Delta t$, x = 10, sampling frequency = 720 Hz. Variable frequency model (VFM). No bad data.

Case #		1	2	3	4	5
No. of sample		24	36	48	60	72
Window size (msec.)		33.33	50.00	66.67	83.33	100.00
LS	V^+	1.414	1.414	1.414	1.414	1.414
Algorithm	Δf^x	0.000	0.000	0.000	0.000	0.000
	b^*	0.194	0.201	0.200	0.200	0.200
LAV	V^+	1.414	1.414	1.414	1.414	1.414
Algorithm	Δf^x	0.000	0.000	0.000	0.000	0.000
	b^*	0.235	0.198	0.201	0.200	0.201

* The actual value for b = 0.2

x The actual value for $\Delta f = 0.0$

\+ The actual value for V = 1.414

Table 4.2b. Continued from Table 2a, the measurements set is contaminated with gross error. (Bad data)

Case #		1	2	3	4	5
No. of Samples		24	36	48	60	72
Window Size (msec.)		33.33	50.00	66.67	83.33	100.00
LS	V	1.148	1.107	1.112	1.545	1.127
	Δf	6.601	2.380	3.960	0.726	1.317
Algorithm	b	fail	fail	fail	fail	fail
LAV	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	b	0.222	0.199	0.199	0.200	0.199
	$\Sigma\|r_i\|$	2.732	7.465	6.927	8.926	12.931
LP	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	b	0.200	0.201	0.200	0.200	0.200
	$\Sigma\|r_i\|$	2.732	7.465	6.927	8.926	12.931
# of iter.		14	18	28	22	27

4.2 Effect of Sampling Rate.

The algorithm is examined at different sampling frequencies or sampling rates. Table 4.3 gives the results obtained for CFM when the sampling frequency is 180, 360, 540, 720 and 900 Hz when the set of measurements contains no bad data and when it contains bad data. Studying Table 4.3 reveals that with no bad data the three algorithms give the same results, and when the measurements are contaminated with bad data the LAV algorithm and the LP algorithm give good estimates for the variables, they are almost equal to the true values, in contrast to the LS, which is adversely affected by the presence of bad data (it gives a bad estimate for the variables). We can also notice from Table 4.3 that the linear programming based LAV algorithm produces the optimal estimate after a large number of iterations, varying between 11 to 16,

Also it can be noticed that as the sampling frequency increases, it does not affect the optimal estimate for LP algorithm and the new LAV algorithm, but it does affect the LS algorithm.

Table 4.4 reports the same results obtained for a VFM. The above discussions hold true for this table. In conclusion, a suitable combination of data window size and sampling rate must be selected for the LS algorithm, but it is not necessary for the LP or LAV algorithms.

Table 4.3a. The estimated phasor voltage, the frequency deviation for a number of samples = 60, number of unknowns = 8 and the time reference is at the middle of the window. CFM. No bad data.

Case #		1	2	3	4	5
Sampling frequency (Hz)		180	360	540	720	900
Data window size (msec.)		333.3	166.6	111.1	83.33	66.60
LS	V	1.414	1.414	1.414	1.414	1.414
Algorithm	Δf	0.000	0.000	0.000	0.000	0.000
LAV	V	1.428	1.414	1.415	1.414	1.414
Algorithm	Δf	0.000	0.000	0.000	0.000	0.000
LP	V	1.414	1.414	1.414	1.414	1.414
Algorithm	Δf	0.000	0.000	0.000	0.000	0.000

Table 4.3b. The estimated phasor voltage and the frequency deviation for number of samples = 60, number of unknowns = 8 and the time reference is at the middle of the window for CFM. Measurements are contaminated with gross error (bad data).

Case #		1	2	3	4	5
Sampling frequency		180	360	540	720	900
Data Window (msec.)		333.3	166.6	111.1	83.33	66.66
LS	V	1.467	1.425	1.476	1.375	1.380
Algorithm	Δf	0.069	0.080	0.281	1.545	1.028
LAV	V	1.414	1.414	1.414	1.414	1.415
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	$\Sigma\|r_i\|$	4.901	2.829	5.571	3.500	4.929
LP	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	$\Sigma\|r_i\|$	4.900	2.829	5.571	3.464	2.900
	# of iter.	11	13	11	15	16

Table 4.4a. The estimated phasor voltage, the frequency deviation and the rate of change of frequency. Number of unknowns = 10, Number of samples = 60, time reference at he middle of the window, VFM, and no bad data.

Case #		1	2	3	4	5
Sampling frequency (Hz)		180	360	540	720	900
Data Window Size (msec.)		333.3	166.7	111.1	83.33	66.66
LS	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	b	0.200	0.200	0.200	0.200	0.200
LAV, LP	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	b	0.200	0.200	0.200	0.200	0.200

Table 4.4b. Continued from Table 4a. Measurements contaminated with gross error (bad data).

Sampling Frequency (Hz)		180	360	540	720	900
Data Window Size (msec.)		333.3	166.7	111.1	83.33	66.66
LS	V	1.389	1.413	1.400	1.514	1.487
Algorithm	Δf	0.069	0.078	0.316	1.267	0.848
LAV	V	1.383	1.433	1.414	1.414	1.414
	Δf	0.076	0.000	0.000	0.000	0.000
Algorithm	b	0.205	0.201	0.199	0.202	0.204
	$\Sigma\|r_i\|$	6.948	2.856	5.573	3.469	2.693
LP	V	1.414	1.414	1.414	1.414	1.414
	Δf	0.000	0.000	0.000	0.000	0.000
Algorithm	b	0.200	0.200	0.200	0.200	0.200
	$\Sigma\|r_i\|$	4.934	2.844	5.572	3.467	2.693
# of iter.		26	19	26	17	23

4.3. Effect of Taylor Series Expansion Truncation Approximation.

It can be noticed from equations 3.3a and 3.4a that if the data widow size is small ($t=m*\Delta t<<1$) a few terms of Taylor series expansion of sine and cosine functions would be adequate for reasonable accuracy. But, if the measurement window size is large, more terms are needed. Table 4.5 reports the results when the number of terms kept from the Taylor series is 3, 4 and 5 of sine and cosine functions. With no bad data, the three algorithms produce the same optimal estimate for CFM, but when the measurements are contaminated with gross error the LS algorithm produces bad estimates, while the LAV and LP algorithms produce good estimates (optimal estimates) and they are the same. However, the LP based LAV algorithm produces the optimal estimate after a great number of iterations. It can be noticed, from Table 4.5, that the number of terms truncated from the Taylor series expression does not affect the optimal estimate for the new LAV and LP based LAV algorithms, but it does affect the LS algorithm.

Table 4.5. The estimated phasor voltage and frequency deviation for different truncated from Taylor series of sine and cosine functions. Number of samples = 60, sampling frequency = 720Hz, time reference at the middle of the window size, CFM.

Case #			1	2	3
Number of variables			6	8	10
No Bad Data	LS	V	1.414	1.414	1.414
		Δf	0.000	0.000	0.000
	LAV	V	1.414	1.414	1.414
		Δf	0.000	0.000	0.000
	LP	V	1.414	1.414	1.414
		Δf	0.000	0.000	0.000
With Bad Data	LS	V	1.362	1.351	1.511
		Δf	0.540	2.110	1.267
	LAV	V	1.414	1.414	1.414
		Δf	0.000	0.000	0.000
		$\Sigma\|r_i\|$	3.460	3.470	3.470
	LP	V	1.414	1.414	1.414
		Δf	0.000	0.000	0.000
		$\Sigma\|r_i\|$	3.460	3.460	3.460
		# of iter.	10	12	14

4.4. Effect of Time Reference.

To reduce the computing time, the time reference may be chosen at he middle of the data window size for the LS algorithm utilizing the fact that by doing so the elements of the rows of the matrix $[A^TA]^{-1}A^T$ will be symetrical. However, the position of the time reference has no effect on the results of the optimal estimate for the all three algorithms.

4.5. Rate of Change of Frequency Via Linear Regression.

The new LAV algorithm together with the LP based LAV and LS algorithms are used to estimate the rate of change of frequency. Table 4.6 gives the results obtained from a computer simulation when the number of measurements, used to estimate the coefficients a and b, are 3, 6, 9 and 12 measurements. With no bad data the three algorithms produce the same optimal estimate, but with bad data both LAV and LP algorithms produce the the optimal estimate, while the LS algorithm fails to estimate the coefficients a and b.

Table 4.6. The estimated nominal frequency and the rate of change of frequency using linear regression. Sampling frequency = 720 Hz.

Case Number			1	2	3	4
Number of samples used			3	6		12
No	LS	a	60.00	60.00	60.00	60.00
		b	0.200	0.200	0.200	0.200
Bad	LAV	a	60.00	60.00	60.00	60.00
		b	0.200	0.200	0.200	0.200
Data	LP	a	60.00	60.00	60.00	60.00
		b	0.200	0.200	0.200	0.200
With	LS	a	56.70	54.70	55.00	55.00
		b	0.200	411.6	240.2	126.1
	LAV	a	60.00	60.00	60.00	60.00
Bad		b	0.200	0.200	0.200	0.200
		$\Sigma\|r_i\|$	10.00	20.00	30.00	50.00
	LP	a	60.00	60.00	60.00	60.00
Data		b	0.200	0.200	0.200	0.200
		$\Sigma\|r_i\|$	10.00	20.00	30.00	50.00
# of iter.			2	4	2	3

CHAPTER V

KALMAN FILTER - A DYNAMIC APPROACH

In Chapter II, III, and VI, three curve fitting techniques, the least squares, the linear programming based LAV, and the new LAV, and their application to frequency estimations have been presented and explored. As the name suggested, these three algorithms use parameter estimation approach to estimate states in power systems, frequency and amplitude of voltage phasors. In order to carry out the estimation, a set of digitized measurements are required. In this frequency estimation case, a certain number of digitized measurements of the voltage phasors have to be taken and stored in the microprocessor before each estimation is carried out, which is often referred as the static approach. In other words, we used a static approach to estimate a dynamic system. Compared with dynamic approaches, where only one measurement is needed to update the estimation, the curve fitting techniques obviously take more time to carry out estimations. In this chapter, the application of Kalman filter for tracking the voltage frequency is presented.

5.1 State model of the signal [6].

The noise-free voltage signal can be expressed as

$$v(t) = A \cos(\omega_0 t + \phi) \qquad (5.1)$$

where A is the amplitude and ϕ is the phase angle. Expand (5.1) into (5.2)

$$v(t) = A\cos\phi\cos\omega_0 t - A\sin\phi\sin\omega_0 t \qquad (5.2)$$
$$= x_1\cos\omega_0 t - x_2\sin\omega_0 t$$

where $x_1 = A\cos\phi$, $x_2 = A\sin\phi$.

Therefore we have the state equation and the measurement equation of the voltage signal as follows

$$\mathbf{x}_{k+1} = \phi_k\mathbf{x}_k + \mathbf{w}_k \qquad (5.3)$$

$$z_k = H_k\mathbf{x}_k + v_k$$
$$= [\cos(\omega_0 k\Delta t) - \sin(\omega_0 k\Delta t)]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + v_k \qquad (5.4)$$

$$k = 1, 2, \ldots..$$

where $\mathbf{x}_k$ is the 2x1 process state vector at step k.

$\phi_k = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix}$, is the state transition matrix.

$\mathbf{w}_k$ is the 2x1 plant noise vector uncorrelated sequence with known covariance structure Q_k.

z_k is the 1x1 measurement at step k.

H_k is the 1x2 matrix giving the ideal relationship between the measurement and the state vector.

v_k is the 1x1 measurement error-assumed to be uncorrelated with known covariance structure R_k.

Having the two state variables estimated using the equations described in the next section, the amplitude and

the phase angle can be calculated simultaneously using (5.5) and (5.6)

$$A = \sqrt{x_1^2 + x_2^2} \tag{5.5}$$

$$\phi = \tan^{-1}\frac{x_2}{x_1} \tag{5.6}$$

And using the fact that the frequency deviation in Hz is linearly related to the average rate of change of the phase angle [6], we have the following equation

$$\frac{\partial \phi}{\partial t} \text{ (radians/sec)} = 2\pi\Delta f \tag{5.7}$$

Frequency deviation can be calculated from (5.7).

5.2 Kalman Filter Algorithms.

A complete derivations of Kalman filter equations are beyond the scope of this thesis. Only the filter recursive equations are presented with a brief description of their use.

Having an apriori estimate $x_{k/k-1}$ and its error covariance $P_{k/k-1}$ the general recursive Kalman filter equations are as follows [6]

(1) Compute Kalman filter gain K_k:

$$K_k = P_{k/k-1}\, H_k^T \left(H_k\, P_{k/k-1}\, H_k^T + R_k \right)^{-1} \tag{5.8}$$

(2) Update estimate with measurement z_k:

$$\mathbf{x}_{k/k} = \mathbf{x}_{k/k-1} + K_k(z_k - H_k\mathbf{x}_{k/k-1}) \qquad (5.9)$$

(3) Update error covariance:

$$P_{k/k} = (I - K_kH_k)P_{k/k-1} \qquad (5.10)$$

(4) Extrapolate the state vector and its covariance:

$$\mathbf{x}_{k+1/k} = \phi_k\mathbf{x}_k \qquad (5.11)$$

$$P_{k+1/k} = \phi_kP_k\phi_k^T + Q_k \qquad (5.12)$$

where

R_k is the covariance matrix of observation errors at instant k.

Q_k is the covariance matrix of plant noise at instant k

It can be noticed that equations (5.8), (5.10),and (5.12) are independent of measurement and the estimate. Therefore only equations (5.9) and (5.11) are to be computed on-line.

5.3 Computer Simulation and Results.

The Kalman filter was first tested with different frequency deviations to find out the range of accuracy. In the test, a zero mean Gaussian white noise, with a standard deviation of 0.01, was generated and imposed on the sinusoidal waveform to represent the noisy measurements. The sampling frequency was 360 Hz. The test results are shown in Figure 5.1 and Figure 5.2. We can conclude from these tests that the Kalman filter converges to the exact values quickly

(in about 20 sampling intervals, or 3.333 cycles) and accurately, and is hardly affected by the amplitude of frequency deviations.

Then the effect of the noise level was studied. The standard deviation of the noise was taken at three different values, 0.05, 0.1 and 0.2. The results are shown in Figure 5.3, from which we can see that the estimates are slightly affected by the noise level. It takes a little longer for the higher level noisy data to converge than the lower level noisy data (about 5 sampling intervals, or 0.833 cycle). However, the Kalman filter is the least affected by the noise level than other algorithms, provided that the noise is white [6].

The Kalman filter was also tested with continuously increasing frequency deviations. The result of this test is shown in Figures 5.4, 5.5, 5.6, and 5.7. The rate itself has little effect on the performance of the filter, and it can be computed using curve fitting techniques, or simply average the change rate as the filter progresses.

Finally, the Kalman filter was tested with measurements which were contaminated with bad data (gross error) to check its capability of rejecting the bad data. In this test, the signs of samples #21 and #41 were reversed. It can be seen that Kalman filter can not reject the bad data (Figure 5.8). Neither can it converge to the actual value fast enough so that measures can be applied to avoid the wrong estimate. For example, 3.3 cycles after #41 (bad data point) is

introduced, there is still an error of approximately 0.5 Hz present (Figure 5.8), which is not acceptable for frequency relays.

So, we can conclude that the Kalman filter gives the best estimate on measurements with white noise, but fails on measurements contaminated with gross errors.

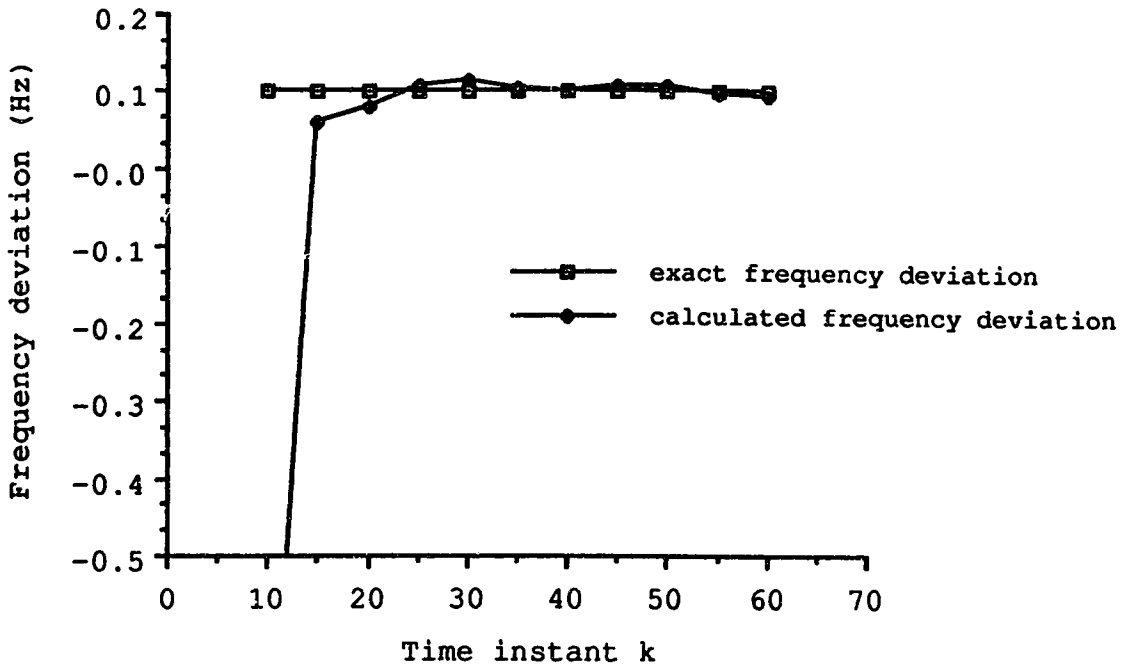

Figure 5.1 Computed frequency deviation for noisy data and a constant frequency deviation 0f 0.1 Hz.

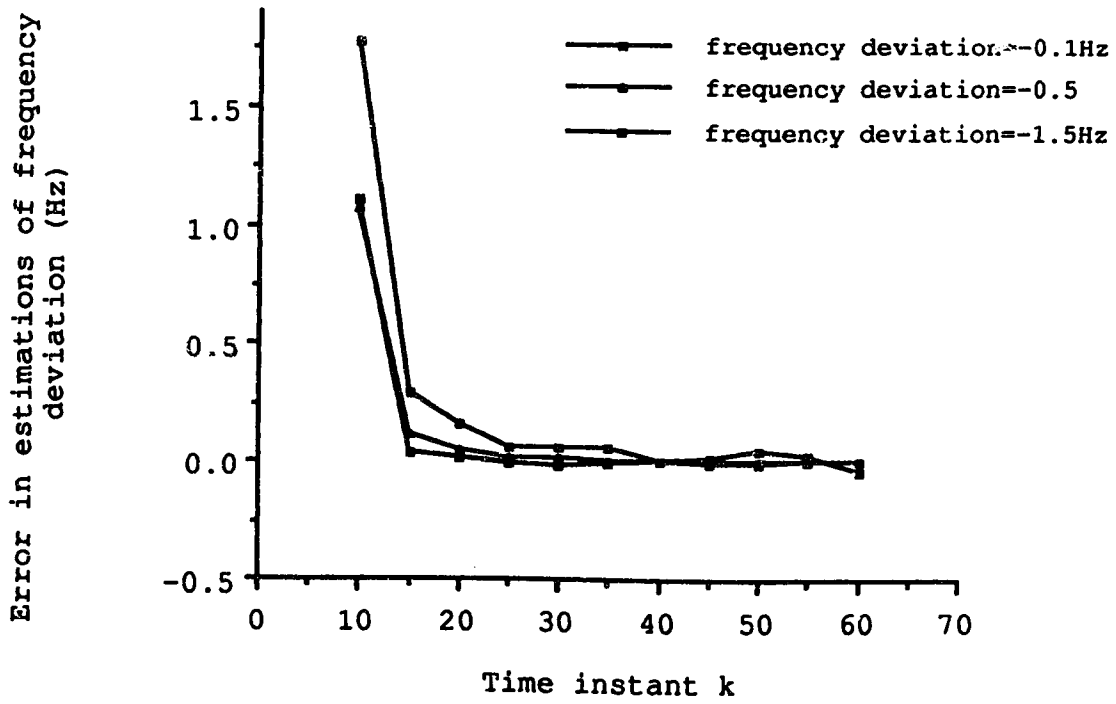

Figure 5.2 Comparison of errors in estimations of different frequency deviations, standard deviation of white noise=0.01

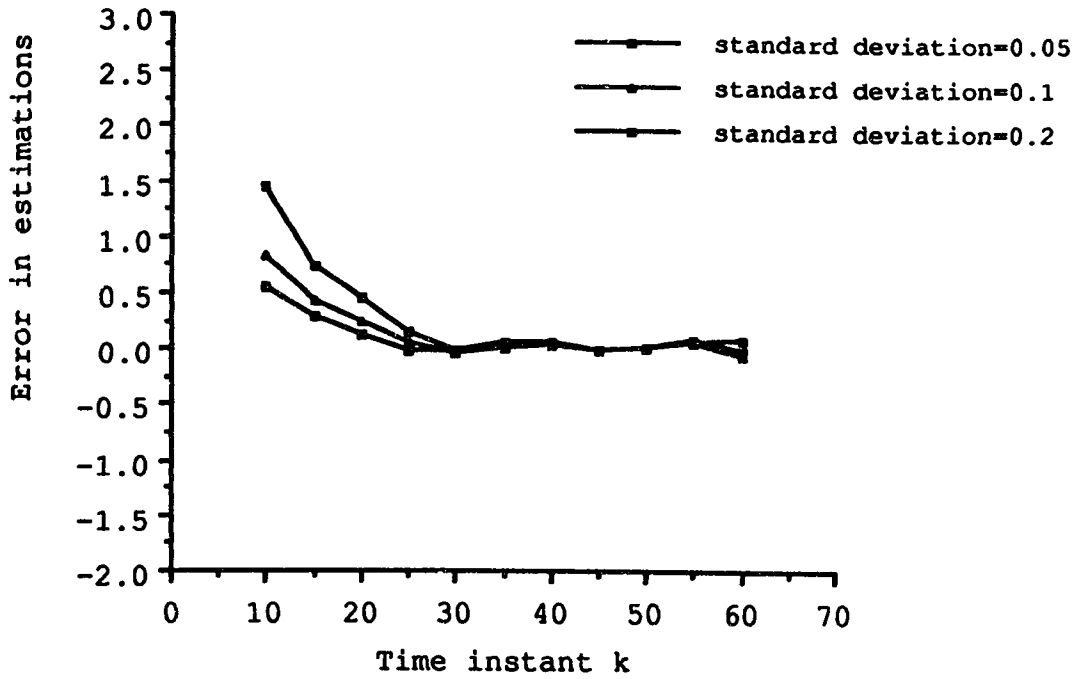

Figure 5.3 Comparison of errors in estimation for noisy data of different standard deviations, frequency deviation=-0.5Hz

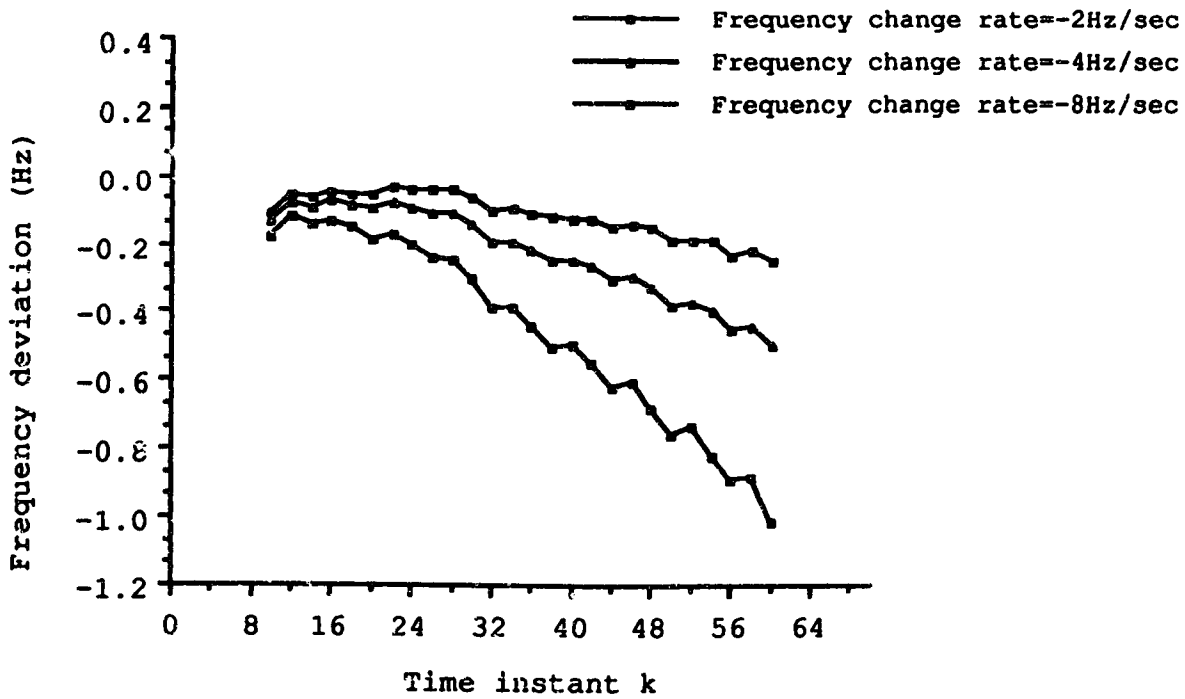

Figure 5.4 Test results on noisy data and continuously incresing frequency deviation.

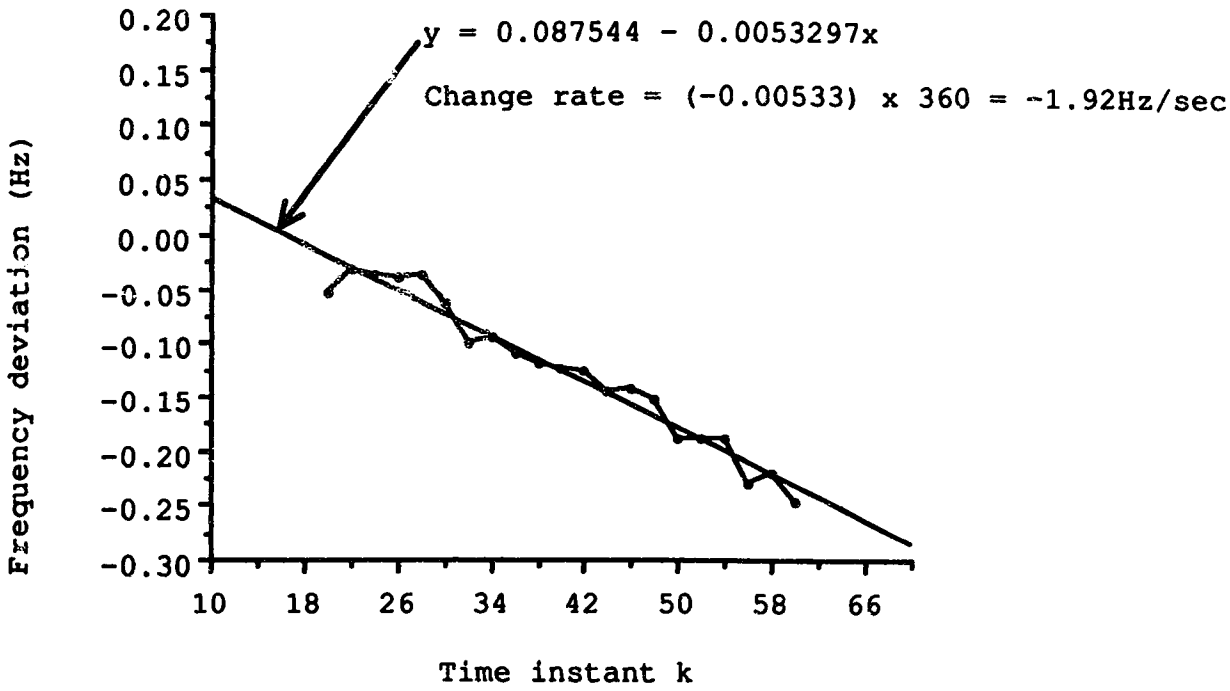

Figure 5.5 Computed frequency change rate for noisy data and continuously increasing frequency deviation, with actual change rate=-2Hz/sec, for k≥20.

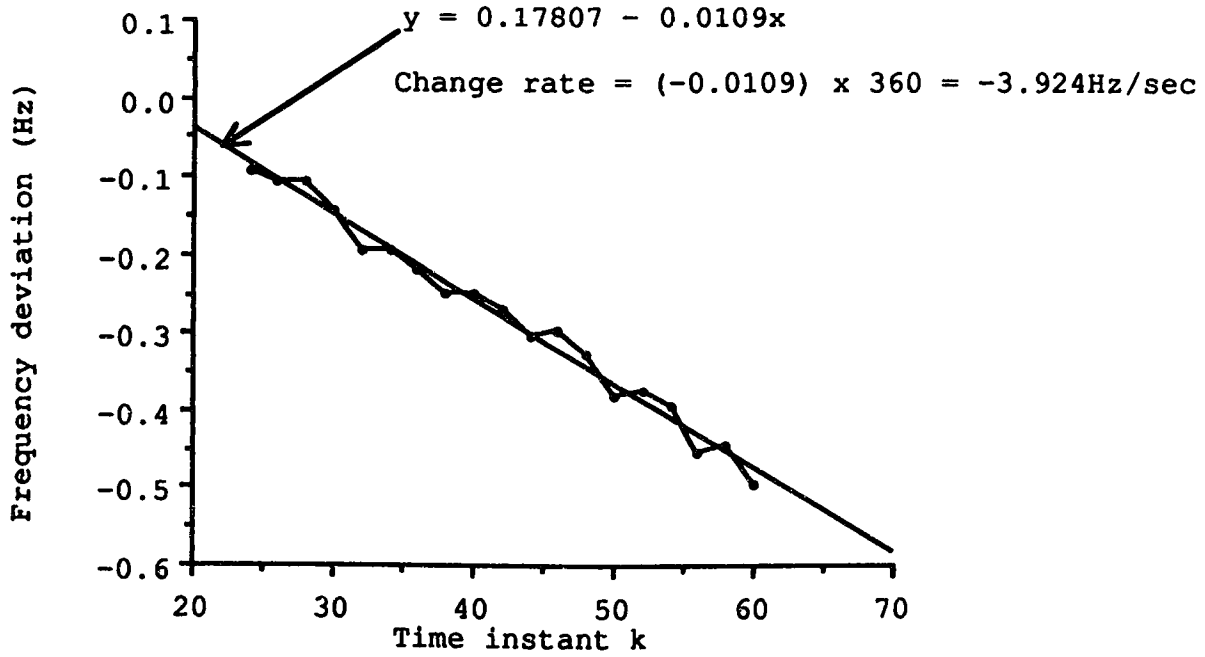

Figure 5.6 Computed frequency change rate for noisy data and continuously increasing frequency deviation, with actual change rate=-4Hz/sec, for k≥24.

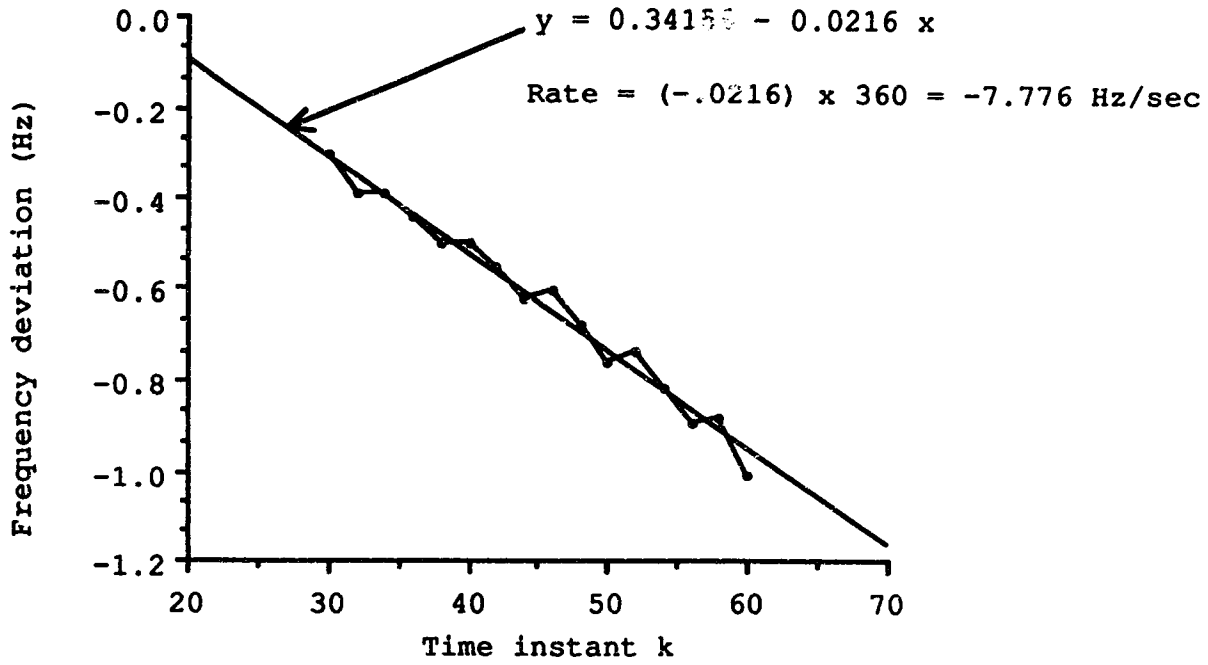

Figure 5.7 Computed frequency change rate for noisy data and continuously increasing frequency deviation, with actual change rate=-8Hz/sec, for k≥30.

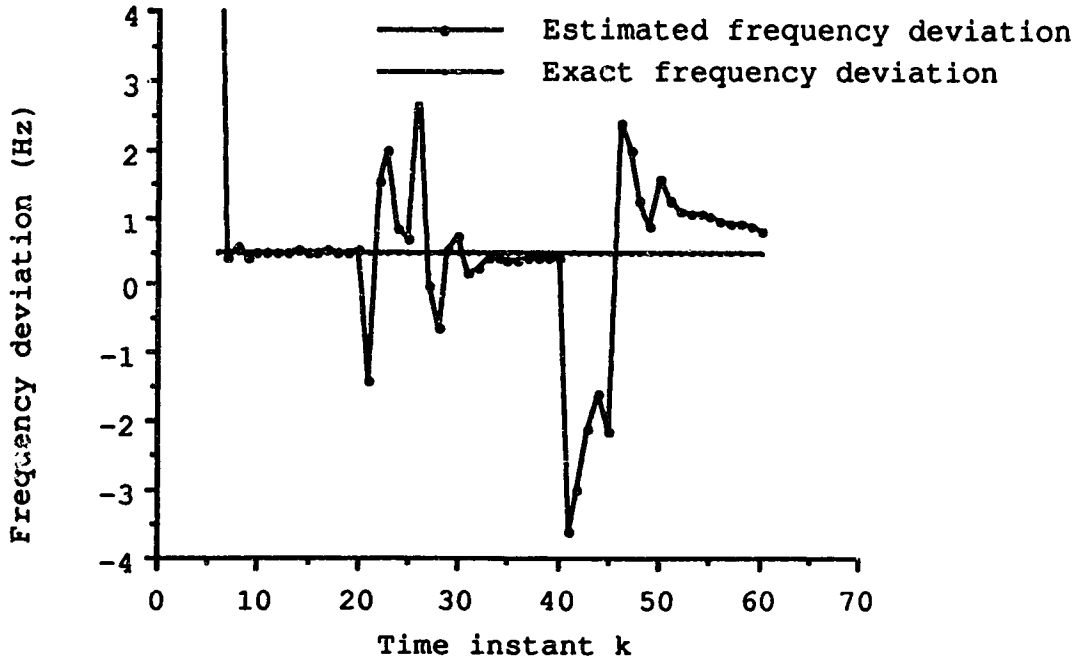

Figure 5.8 Frequency estimation of Kalman filter when the measurements are contaminated with bad data (gross error), the signs of samples #21 and #41 are reversed.

CHAPTER VI

CONCLUSIONS

This thesis deals with frequency measurements at a bus for relaying purposes. In the first part of the thesis (chapter II, III, and VI), three parameter estimation techniques, least squares (LS), linear programming based LAV, and the new LAV, have been applied to frequency estimations. Two models were considered in this work, namely, a constant frequency model (CFM) and a variable frequency model (VFM). The frequency measurements are used to estimate the true values of the frequency. These three different estimation techniques are used and in each case the results are compared for situations with no bad data present and situations with bad data present (gross error). It was invariably found that the three methods: LS, new LAV, and linear programming based LAV all give the same results with no bad data present, whereas when bad data is present the least squares estimation gives considerably poorer results than linear programming based LAV and the new LAV. In addition, the linear programming based LAV requires many iterations whereas the new LAV is not iterative in nature. Hence from the results presented in this thesis, the new LAV method is best suited for problems in which most measurements are accurate with only a few bad data.

In chapter V, the Kalman filter was used to estimate the frequency changes and the rate of frequency change. As a

dynamic filter, it is faster than the fixed window techniques. However, since the Kalman filter is based on the LS standard, it is affected by the presence of bad data.

Therefore, it is concluded that the new LAV estimator is a better choice for situations where most of the measurements are correct and only a few are bad data.

The new LAV is a static state estimator. It recalculate the states after every sampling window. We are expecting to develop a dynamic state estimator, like the Kalman filter, which will have the same property as the static one, that is it can reject the bad data as it progresses.

In this thesis, the models of the voltage phasors were assumed as a pure sinusoidal wave, with constant frequency or linearly changing frequency in the sampling data widow size. Obviously this is not exactly the case in the power system. Further research require more sophisticated models. For example harmonics may exist in the measurements; the amplitude of the voltage phasors may also change as a result of load change.

References.

1. M.S. Sachdev and M.M. Giray, " A Least Error Squares Technique for Determining Power System Frequenc , IEEE Trans. on Power System and Apparatus, Vol. 104, N 2, pp 437-444, 1985.

2. M.S. Sachdev and M.M. Giray, " Measurement of Local Frequency from Digitized Bus Voltage Samples", Trans Canadian Elect. Assoc. Engineering and Operation Division, Vol. 23, Part 3, 1984, Paper No. 84-sp-149.

3. M.S. Sachdev and M.M. Giray, " Amplitude and Frequency Measurements by Discrete Fourier Transform ", Trans. CEA, Power System Planning & Operation Section, Vol. 24, 1985, Paper No. 85.sp.150.

4. A.G. Phadek, J.S. Thorp and M.G. Adamiak, " A new Measurement Technique for Tracking Voltage Phasors, Local System Frequency, and Rate of Change of Frequency", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-102, No. 5, pp 1025-1038, 1983.

5. S.A. Soliman, G.S. Christensen and A. Rouhi, " A New Technique for Curve Fitting Based on Minimum Absolute Deviations" Comput. Statistics & Data Analysis, Vol. 6, No. 4, pp 341-351, 1988.

6. A. A. Girgis and T.L. Daniel Hwang, " Optimal Estimation of Voltage Phasors and Frequency Deviation Using Linear and Non-Linear Kalman Filtering : Theory and Limitations ". IEEE Trans. on Power Apparatus and Systems, Vol. PAS-103, No. 10, October 1984.

7. M.S. Sachdev, H.C. Wood and N.G. Johnson, " Kalman Filtering Applied to Power System Measurements for Relaying",

Trans. on Power Apparatus and Systems, Vol. PAS-104, No. 12, pp 3565-3575, 1985.

8. M.M. Giray, " A Least Error Squares Technique For Estimating The Magnitude And Frequency of A Voltage Signal", Ph.D. Thesis, University of Saskatchewan, Saskatoon, Canada, October 1986.

9. S.A. Soliman, G.S. Christensen, D.H. Kelly and N. Liu, " An Algorithm For Frequency Relaying Based on Least Absolute Value Approximations", Electrical Power Systems Research, in Press April 1990.

10.J.M. Mendel, " Lessons in Digital Estimation Theory", Prentice Hall, Englecliffs, New Jersey, 1987.

11.Y. Dodge, " An Introduction to Statistical Data Analysis L_1-Norm Based", Statistical Data Analysis Based on the L_1-Norm and Related Methods. Y. Dodge Editor, Elsevier Science Publishing Company, New York, 1987, pp. 1-21.

12.V. Chvatal, " Linear Programming", W.H. Freeman and Company, New York, 1983.

13.E.J. Schlossmacher, " An Iterative Technique for Absolute Deviations Curve Fitting", Journal of the American Statistical Association, Vol. 68, No. 344, December 1973, pp. 857-859.

14.F.R. Glahe and J.G.Hunt, " The small sample properties of Simultaneous Equation Least Absolute Estimators Via-A-Vis Least Squares Estimators", Econometrica, Vol. 38, No. 5, September 1970, pp. 742-753.

15.J.E. Gentle, " Least Absolute Value Estimation: An Introduction", Communications in Statistics - Simulation and Computation, B6(4), 1977, pp. 313-328.

16.P. Bloomfield and W. Steiger, " Least Absolute Deviations Curve Fitting", SIAM Journal Scientific and Statistical Computing, Vol. 1, No. 2, June 1980, pp. 290-301.

17.E.A. Udren, "Load Shedding and Frequency Relaying", Westinghouse Electric Corporation, Relay-Instrument Division , " Applied Protective Relaying ", 1979, pp. 21.1-21.16.

18.I. Barrodale and F.D.K. Roberts, " An Improved Algorithm for Discrete L_1 Linear Approximation ", SIAM Journal of Numerical Analysis, Vol. 10, No. 5, 1973, pp. 839-848.

19.G.F. McCormick and V.A. Sposito, " Using the L_2-Estimator in L_1-Estimation ", SIAM Journal of Numerical Analysis, Vol. 13, No. 3, June 1976, pp. 337-343.

20.C. Eisenhart, " The Meaning of 'Least' in Least Squares", Journal of the Washington Academy of Sciences, Vol. 54, 1964, pp. 24-33.

21.A.H. Rouhi, " A New Technique For Absolute Value Power System State Estimation ", M.Sc. Thesis, University of Alberta, Edmonton, Alberta, 1988.

22.G.S. Christensen, S.A. Soliman and A. Rouhi, " Discussion of 'An example showing that a new technique for LAV estimation breaks down in certain cases'", Computational Statistics & Data Analysis, 9, 1990, pp. 203-213.

23.R.I. Shrager and E.Hill, Jr., " Non-linear Curve-Fitting in the L_1 and L_∞ Norms", Mathematics of Computation, Vol. 34, No. 150, April 1980, pp. 529-541.