Power system state estimation based on fuzzy linear regression technique

Amr Abou-Ghazala [a] and Nabil H. Abbasy [b]
[a] *Electrical Eng. Dept., Alexandria University, Alexandria, Egypt*
[b] *Electrical Eng. Dept., College of Technological Studies, Shuwaikh, Kuwait*

A novel approach for power system state estimation is presented in this paper. The proposed approach is based on a fuzzy linear regression model, and uses on-line measurements, available from remote terminal units. The problem is formulated as a fuzzy linear programming optimization problem, where the objective is to minimize the spread of the estimated states, over the total number of available measurements. Effects of measurements accuracy, as well as the degree of fuzziness on the estimated states are discussed in the paper via simulated examples on a 6-bus system. Superior bad data detection and identification is obtained and verified via the simulation.

يقدم هذا البحث طريقة جديدة لتقدير الحال الإستاتيكي لمنظومة القوى. تعتمد الطريقة الجديدة على التمثيل الخطي للمفاهيم الغيمية حيث تستخدم القياسات اللحظية المتوفرة من وحدات القراءة البعيدة. تم تشكيل المسألة كمسألة برمجة خطية غيمية حيث تمثل دالة الغرض مجموع انتشار الحالات المحسوب عبر جميع القياسات المتاحة. يناقش البحث تأثير دقة القياسات ودرجة الغيوم على دقة حساب حالات المنظومة عبر مثال مكون من منظومة تحتوي على ٦ قضبان. أثبتت نتائج التمثيل فاعلية الطريقة المقترحة في اكتشاف وتحديد القياسات الرديئة في المنظومة، مع المحافظة على دقة وسرعة حالة المنظومة.

Keywords: Bad data, Linear regression, Power system state estimation, Fuzzy set theory

1. Introduction

The problem of monitoring the power flows and voltages in power systems is very important in maintaining system security. Measurement discrepancies cannot be avoided due to metering nature. Power System State Estimation (PSSE) performs the task of smoothing out small random errors in meter readings, detect and identify gross measurement errors (bad data) [1].

Power systems are large, complex, geographically widely distributed and influenced by unexpected events. These facts make it difficult to effectively deal with many power system problems through strict mathematical approaches. Therefore, intelligent techniques such as fuzzy algorithms have emerged in recent years in power systems as a complement to mathematical approaches and have proved to be effective when properly coupled [2, 3, 4].

This paper presents a novel technique, based on fuzzy set theory [5, 6], for PSSE with superior performance for bad data detection and identification. The proposed approach is based on solving the power system linearized model as fuzzy linear regression problem. Both system measurement variables and system state variables (bus voltage magnitudes and bus angles at all buses, except for the reference bus) are assumed to be a fuzzy numbers, having a certain middle ‘p’ and a spread value ‘c’. The network parameters are assumed to be known and a non-fuzzy ‘crisp’ values. The objective function is to minimize the spread of the estimated states over the data set available. The output is fuzzy system states and fuzzy measurement vector in which the degree of fuzziness is set by the user. Bad data detection is based on exceeding a preset threshold for the output membership value. Measurements with maximum spread and minimum membership value are prime suspects of bad data.

2. Mathematical formulation

2.1. Fuzzy linear regression: a background

The general fuzzy model has the form:

$$\underline{Y} = f(x, \underline{A}) = \sum_{i=1}^{n} A_i\, x_i \;, \qquad (1)$$

where $\underline{Y}$ is a fuzzy dependant variable, x is crisp independent variable, $\underline{A}$ is a set of Ai coefficients, A_i is the i[th] fuzzy coefficient (usually a fuzzy number). The fuzzy parameter A_i is a function of two parameters, p and c, known as the middle value and the spread, respectively. The spread denotes the fuzziness of the function. The membership function of the fuzzy coefficient A_i is assumed triangular with a middle value p and spread c on each side as shown in fig. 1. This can be expressed as [7]:

$$\mu_{(A_i)}(a_i) = \begin{cases} 1 - \dfrac{|p_i - a_i|}{c_i} & p_i - c_i \le a_i \le p_i + c_i \\ 0 & otherwise \end{cases} . \qquad (2)$$

The fuzzy parameters $A = (A_1, \ldots, A_n)$ can be denoted in the vector form of $\underline{A} = \{\underline{P}, \underline{C}\}$, where $\underline{P} = \{p_1, \ldots p_n\}$ and $\underline{C} = \{c_1, \ldots, c_n\}$are the middle and spread vectors. Therefore, the output $\underline{Y}$ can be written as:

$$\underline{Y} = (p_1, c_1)x_1 + (p_2, c_2)x_2 + \ldots + (p_n, c_n)x_n \;, \qquad (3)$$

and the membership function for the output fuzzy parameter, as shown in fig. 2, $\underline{Y}$ is given by [8]:

$$\mu_Y(\underline{Y}) = 1 - \frac{\left| y - \sum_{i=1}^{n} p_i x_i \right|}{\sum_{i=1}^{n} c_i x_i} . \qquad (4)$$

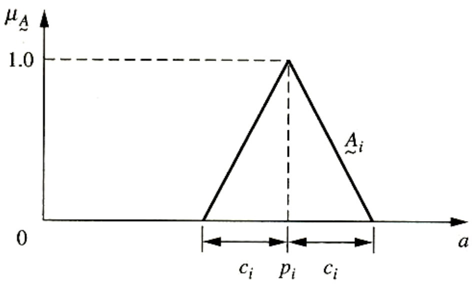

Fig. 1. Membership function for the fuzzy parameters A.

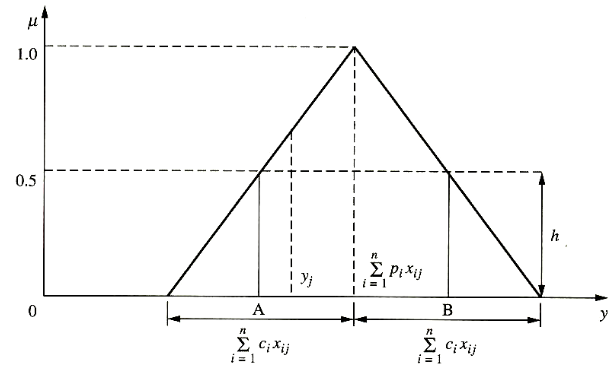

Fig. 2. Fuzzy output function.

In regression we seek to find the fuzzy coefficients that minimize the spread of fuzzy output for the whole data set. The objective function that has to be minimized:

$$O = min \sum_{j=1}^{m} \sum_{i=1}^{n} c_i x_{ij} \;, \qquad (5)$$

where m is the number of samples, x_{ij} is the independent variable x_i at data set j. This objective function is minimized subject to two constraints given by:

$$y_j \ge \sum_{i=1}^{n} p_i x_{ij} - (1-h) \sum_{i=1}^{n} c_i x_{ij} \quad , j = 1, \ldots, m \;, \qquad (6)$$

$$y_j \le \sum_{i=1}^{n} p_i x_{ij} + (1-h) \sum_{i=1}^{n} c_i x_{ij} \quad , j = 1, \ldots., m \;, \qquad (7)$$

where h, defined as fuzzification degree, is specified by the user. As h increases, the fuzziness of the output increases. This is due to the need of a wider spread to validate the input measured values in condition of satisfying higher h [5]. Since each data set produces two constraints, there is a total of 2 m constraints. The problem formulated in eqs. (5) and (6) and (7) is a standard optimization problem and can be solved using any optimization software available.

2.2. Problem formulation

The linearized power system measurement model about a given point x_o is given by:

$$\Delta Z_j = \sum_{i=1}^{n} h_{ij}\Delta x_i \qquad , j = 1,...,m , \tag{8}$$

where Zj is the system j^{th} measurement, m and n are the total number of system measurements and states, respectively, hij is the ij^{th} element of the system measurement jacobian matrix, x_i is the i^{th} system state.

Eq. (8) can be rewritten as:

$$\Delta \underline{Z} = \Delta x_1 \underline{h}_1 + \Delta x_2 \underline{h}_2 + ... + \Delta x_n \underline{h}_n , \tag{9}$$

where $\Delta \underline{Z}$ is the (mx1) measurement difference vector, $\underline{h}_i$ is the (mx1) i[th] column of the measurement jacobian matrix. Eq. (9) is a linear model in the state variable Δx_i, i=1, 2, .., n. If these state variables are fuzzy numbers, then eq. (9) is a general fuzzy linear regression model. The system state variables are assumed to have a middle value p_i, i=1,2,…, n. The spread value represents a measure to the variable fuzziness. In a fuzzy linear regression the objective is to minimize the spread of the state variables over the number of measurements available. This can be expressed as:

$$O = min \sum_{i=1}^{n} \sum_{j=1}^{m} c_i h_{ji} \quad , \tag{10}$$

subject to satisfying the inequality constraints on each measurement:

$$\Delta Z_j \geq \sum_{i=1}^{n} p_i h_{ji} - (1-h) \sum_{i=1}^{n} c_i h_{ji} \quad . \tag{11}$$

$$\Delta Z_j \leq \sum_{i=1}^{n} p_i h_{ji} + (1-h) \sum_{i=1}^{n} c_i h_{ji} \, . \tag{12}$$

where h is specified by the user indicating the degree of fuzziness. The optimization problem defined by eqs. (10,11), and (12) is a linear programming optimization problem and can be solved using any available software available. Having estimated the fuzzy parameters p's and c's, all system operating conditions (e.g. complex bus voltages, line flows, ..etc) will be readily available using fuzzy set rules. Bad data detection is easily determined through exceeding a preset threshold for measurements membership value. Identification is done by discarding the measurement that has a membership value of exactly h and maximum spread, and rerunning the optimization program.

3. Implementation and results

The proposed algorithm is tested using a 6-bus system with 11 states and 62 measurements [9]. System layout and data are shown in fig. 3 and table 1. A MATLAB code was used to apply the algorithm. Convergence was achieved after the third iteration (p<1e-5). Output fuzzy states for different fuzzification factors are shown in table 2. Table 3 shows the input crisp measurements, traditional output using least square state estimation, and the output fuzzy measurements with its membership values. From these results, one can conclude that the degree of fuzziness, h, has no effect on the middle, p, of the estimated states, and directly proportionally, as expected, to the spread. It seems that high discrepancy in measurements affects the accuracy of the output fuzzy states when applying fuzzy linear regression technique. Yet, this algorithm can easily detect and identify bad data.

Bad data detection was based on setting a threshold (was set to .3 pu) for the spread of the measurements with barely h membership value. Exceeding the threshold value flags a bad data occurrence. Measurements with exactly h membership value and high spread are the primary suspects of bad data. Removing the highest spread measurement and rerunning the program reveals the removal of the bad data. Tables 4, 5 show the examples of P_{35} power flow reversal and 300% error increase in Q_{14}. In conclusion, it can be said, through extensive bad data runs, that the measurement with exactly h membership value and highest spread is most probably the bad data.

4. Conclusions

A new technique for PSSE based on fuzzy linear regression is presented in this paper.

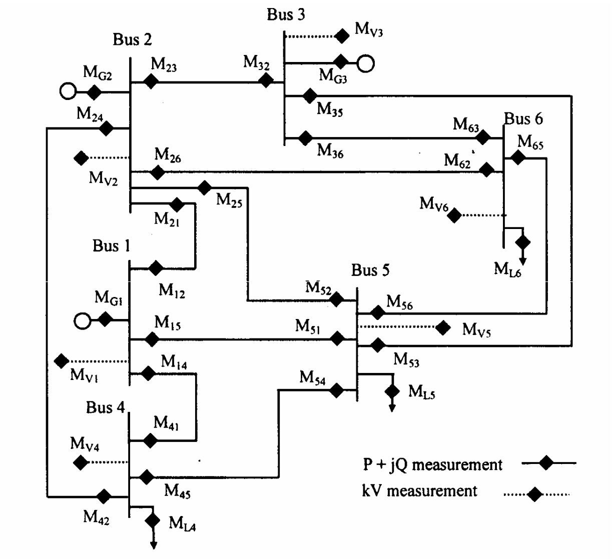

Fig. 3. Six-bus system layout with measurements.

Table 1
Bus system line data, PU (230 kV, 100 MW base)

From	To	*R*	*X*	Bcap (0.5 line cap.)
1	2	0.1	0.2	0.02
1	4	0.05	0.2	0.02
1	5	0.08	0.3	0.03
2	3	0.05	0.25	0.03
2	4	0.05	0.1	0.01
2	5	0.1	0.3	0.02
2	6	0.07	0.2	0.025
3	5	0.12	0.26	0.025
3	6	0.02	0.1	0.01
4	5	0.2	0.4	0.04
5	6	0.1	0.3	0.03

Table 2
Output states using least square state estimation and fuzzy state estimation with different *h*

	V_{se} kV	V_{fuzzy} *(p,c)* kV *h*=0.1	*h*=0.5	*h*=0.9
V_1	240.5	239.8, 1.7	239.8, 3.1	239.8, 7.7
V_2	239.8	239.9, 2.3	239.9, 4.1	239.9, 8.3
V_3	244.6	247.4, 3.6	247.4, 6.6	247.4, 19.6
V_4	226	225.1, 2.1	225.1, 3.9	225.1, 10
V_5	225.1	226.1, 4.5	226.1, 8.7	226.1, 15
V_6	229.9	230, 1	230, 1.9	230, 5
	$Theta_{se}$ (degree)	Theta $_{Fuzzy}$ (degree) *h*=0.1	*h*=0.5	*h*=0.9
$Theta_1$	0	0	0	0
$Theta_2$	-3.8283	-4.07, .13	-4.07, .45	-4.07, .63
$Theta_3$	-4.4656	-4.620, .3	-4.620, .7	-4.620, 1
$Theta_4$	-4.3428	-4.113, .25	-4.113, .5	-4.113, .85
$Theta_5$	-5.5103	-5.741, .2	-5.741, .7	-5.741, 1.1
$Theta_6$	-6.1560	-6.14, .1	-6.14, .4	-6.14, .8

The problem is formulated as a linear optimization problem, where the objective is to minimize the spread of measurements available. Fuzziness factor did not affect the middle value of the estimated states, but it was directly proportional to measurements spread, as expected in fuzzy linear regression techniques. The presented algorithm had a high performance in terms of bad data detection and identification. Bad data was flagged when the spread of any measurements with barely *h* membership value exceeds a threshold value. Identification, through numerous runs, leads always to the measurement with highest spread.

Table 3
Input measurements values and output fuzzy values

Measured quantity	Measured value	State estimation Value	Fuzzy *O/P* Value*(p,c)*, μ								
			h=0.1			*h*=0.5			*h*=0.9		
			p	*c*	μ	*p*	*c*	μ	*p*	*c*	μ
P_1	1.131	1.119	1.1185	0.0946	0.8675	1.1185	0.1703	0.9264	1.1185	0.8517	0.9853
P_2	0.484	0.475	0.4662	0.2741	0.8513	0.4662	0.4934	0.6396	0.4662	1.4672	0.9279
P_3	0.551	0.595	0.5734	0.3117	0.8474	0.5734	0.5611	0.9152	0.5734	2.8055	0.9830
P_4	-0.718	-0.702	-0.6774	0.1679	0.1625	-0.6774	0.3022	0.7347	-0.6774	1.111	0.9269
P_5	-0.72	-0.718	-0.6924	0.2386	0.8844	-0.6924	0.4294	0.9358	-0.6924	2.1472	0.9872
P_6	-0.723	-0.689	-0.6763	0.1965	0.4536	-0.6763	0.3537	0.6853	-0.6763	1.3687	0.9471
Q_1	0.202	0.187	0.1654	0.2885	0.6960	0.1654	0.5193	0.8755	0.1654	1.5966	0.9551
Q_2	0.719	0.703	0.7541	0.5636	0.9378	0.7541	1.0145	0.9654	0.7541	5.0724	0.9931
Q_3	0.906	0.874	1.0780	0.5371	0.6798	1.0780	0.9668	0.8221	1.0780	4.8342	0.9644
Q_4	-0.719	-0.702	-0.7438	0.2179	0.6073	-0.7438	0.4722	0.7818	-0.7438	1.8608	0.9564
Q_5	-0.677	-0.694	-0.6814	0.3810	0.6633	-0.6814	0.8458	0.8129	-0.6814	3.2288	0.9626
Q_6	-0.609	-0.658	-0.6816	0.2862	0.2942	-0.6816	0.4951	0.6079	-0.6816	2.4754	0.9216
P_{12}	0.315	0.304	0.3278	0.0371	0.6554	0.3278	0.0667	0.8086	0.3278	0.3335	0.9617
P_{14}	0.389	0.448	0.4176	0.0318	0.1000	0.4176	0.0572	0.5000	0.4176	0.2859	0.9000
P_{15}	0.357	0.368	0.3731	0.0258	0.3760	0.3731	0.0465	0.6533	0.3731	0.2323	0.9307
P_{23}	0.086	0.03	0.0307	0.0615	0.1000	0.0307	0.1107	0.5000	0.0307	0.5533	0.9000
P_{24}	0.328	0.324	0.2329	0.1056	0.1000	0.2329	0.1901	0.5000	0.2329	0.9506	0.9000
P_{25}	0.174	0.156	0.1352	0.0431	0.1000	0.1352	0.0776	0.5000	0.1352	0.3882	0.9000
P_{26}	0.223	0.259	0.2230	0.0268	0.9988	0.2230	0.0483	0.9993	0.2230	0.2415	0.9999
P_{35}	0.177	0.192	0.1747	0.1072	0.9787	0.1747	0.1930	0.9882	0.1747	0.9652	0.9976
P_{36}	0.433	0.433	0.3584	0.1419	0.4744	0.3584	0.2554	0.7080	0.3584	1.2772	0.9416
P_{45}	0.007	0.043	0.0459	0.0433	0.1000	0.0459	0.0779	0.5000	0.0459	0.3893	0.9000
P_{56}	-0.021	0.013	0.0117	0.0363	0.1000	0.0117	0.0653	0.5000	0.0117	0.3267	0.9000
P_{21}	-0.349	-0.294	-0.3156	0.0371	0. 1000	-0.3156	0.0667	0.5000	-0.3156	0.3335	0.9000
P_{41}	-0.401	-0.436	-0.4070	0.0289	0.7930	-0.4070	0.0520	0.8850	-0.4070	0.2601	0.9770
P_{51}	-0.366	-0.356	-0.3624	0.0230	0.8423	-0.3624	0.0414	0.9124	-0.3624	0.2071	0.9825
P_{32}	-0.021	-0.03	-0.0297	0.0626	0.8609	-0.0297	0.1126	0.9227	-0.0297	0.5631	0.9845
P_{42}	-0.298	-0.309	-0.2163	0.0957	0.1470	-0.2163	0.1723	0.5261	-0.2163	0.8617	0.9052
P_{52}	-0.117	-0.151	-0.1321	0.0391	0.6139	-0.1321	0.0705	0.7855	-0.1321	0.3523	0.9571
P_{62}	-0.196	-0.254	-0.2183	0.0248	0.1000	-0.2183	0.0446	0.5000	-0.2183	0.2231	0.9000
P_{53}	-0.251	-0.181	-0.1649	0.0956	0.1000	-0.1649	0.1722	0.5000	-0.1649	0.8608	0.9000
P_{63}	-0.468	-0.423	-0.4464	0.0851	0.1000	-0.4464	0.1432	0.5000	-0.4464	1.0162	0.9000
P_{54}	-0.021	-0.042	-0.0447	0.0445	0.4674	-0.0447	0.0801	0.7041	-0.0447	0.4003	0.9408
P_{65}	0.01	-0.012	-0.0116	0.0366	0.4088	-0.0116	0.0659	0.6716	-0.0116	0.3293	0.9343
Q_{12}	-0.132	-0.144	-0.1771	0.0733	0.3858	-0.1771	0.1320	0.6588	-0.1771	0.6601	0.9318
Q_{14}	0.212	0.212	0.2147	0.0863	0.9683	0.2147	0.1554	0.9824	0.2147	0.7768	0.9965
Q_{15}	0.094	0.118	0.0878	0.1089	0.6411	0.0878	0.2019	0.8006	0.0878	1.0597	0.9601
Q_{23}	-0.119	-0.126	-0.1745	0.1100	0.4958	-0.1745	0.1981	0.7199	-0.1745	0.9904	0.9440
Q_{24}	0.383	0.453	0.4432	0.1080	0.1000	0.4432	0.2204	0.550	0.4432	1.1022	0.91
Q_{25}	0.22	0.148	0.1009	0.1323	0.1000	0.1009	0.2381	0.570	0.1009	1.1906	0.92
Q_{26}	0.15	0.108	0.1266	0.0695	0.6629	0.1266	0.1252	0.8127	0.1266	0.6258	0.9625
Q_{35}	0.239	0.229	0.2239	0.1797	0.9160	0.2239	0.3235	0.9533	0.2239	1.6173	0.9907
Q_{36}	0.583	0.583	0.5421	0.1402	0.3374	0.5421	0.2323	0.6319	0.5421	1.5616	0.9264
Q_{45}	-0.174	-0.051	-0.1002	0.0820	0.1000	-0.1002	0.1476	0.5000	-0.1002	0.7380	0.9000
Q_{56}	-0.008	-0.101	-0.0467	0.1056	0.6338	-0.0467	0.1901	0.7965	-0.0467	0.9505	0.9593
Q_{21}	0.097	0.119	0.1578	0.0737	0.1746	0.1578	0.1327	0.5414	0.1578	0.6633	0.9083
Q_{41}	-0.143	-0.209	-0.2132	0.0780	0.1000	-0.2132	0.1403	0.5000	-0.2132	0.7016	0.9000
Q_{51}	-0.175	-0.136	-0.1698	0.0869	0.1000	-0.1698	0.1505	0.5000	-0.1698	0.8525	0.9000
Q_{32}	0.102	0.062	0.1120	0.1172	0.9149	0.1120	0.2110	0.9527	0.1120	1.0552	0.9905
Q_{42}	-0.443	-0.444	-0.4305	0.0779	0.4460	-0.4305	0.1842	0.6922	-0.4305	1.0211	0.9384
Q_{52}	-0.222	-0.174	-0.1333	0.1196	0.2584	-0.1333	0.2153	0.5880	-0.1333	1.0764	0.9176
Q_{62}	-0.223	-0.145	-0.1652	0.0642	0.1000	-0.1652	0.1156	0.5000	-0.1652	0.5778	0.9000
Q_{53}	-0.299	-0.258	-0.2564	0.1556	0.7259	-0.2564	0.2801	0.8477	-0.2564	1.4003	0.9695
Q_{63}	-0.511	-0.557	-0.5935	0.1139	0.1000	-0.5935	0.3850	0.60	-0.5935	1.1249	0.91
Q_{54}	-0.015	-0.025	0.0247	0.0832	0.5225	0.0247	0.1498	0.7347	0.0247	0.7492	0.9469
Q_{65}	0.029	0.044	0.0128	0.1081	0.6128	0.0128	0.1946	0.7849	0.0128	0.9728	0.9570
V_1	1.037	1.046	1.0418	0.0074	0.3566	1.0418	0.0133	0.6425	1.0418	0.0666	0.9285
V_2	1.034	1.043	1.0430	0.0100	0.1000	1.0430	0.0181	0.5000	1.0430	0.0903	0.9000
V_3	1.09	1.064	1.0757	0.0159	0.1000	1.0757	0.0287	0.5000	1.0757	0.1435	0.9000
V_4	0.981	0.983	0.9788	0.0094	0.7599	0.9788	0.0169	0.8666	0.9788	0.0843	0.9733
V_5	0.979	0.980	0.9952	0.0316	0.4892	0.9952	0.0569	0.7162	0.9952	0.2845	0.9432
V_6	0.995	1.00	0.9991	0.0046	0.1000	0.9991	0.0082	0.5000	0.9991	0.0411	0.9000

Table 4
Measurements of highest spread when P_{35} is reversed

Measurements with $\mu = 0.5$	Original measured value (MW)	Output fuzzy value(middle, spread) (MW)
P_{35}	0.177	0.045, 0.46
P_{25}	0.174	0.02, 0.3
P_{56}	-0.021	0.13, 0.29
Q_{25}	0.22	0.086, 0.27
P_{63}	-0.468	-0.34, 0.24

Table 5
Measurements of highest spread when Q_{14} has 300% increase

Measurements With $\mu = 0.5$	Original measured value (MW)	Output fuzzy value (middle, spread) (MW)
Q_{14}	0.636	0.38, 0.54
Q_{24}	0.383	0.64, 0.49
Q_{41}	-0.143	-0.35, 0.421

References

[1] J. Grainger, W. Stevenson Jr, Power System Analysis, McGraw-Hill, NY (1994).
[2] K. Momoh, Tomsovic, "Overview and Literature Survey of Fuzzy Set Theory in Power Systems", Power Systems, IEEE Transactions on , Vol. 10, pp. 1676 –1690 (1995).
[3] F. Shabani, N. Prasad and H. Smolleck, "State Estimation with Aid of Fuzzy Logic", Proceedings of the Fifth IEEE International Conference on Fuzzy Systems, pp. 947-953 (1996).
[4] R. Saric, Ciric, "Integrated Fuzzy State Estimation and Load Flow Analysis in Distribution Networks", Power Delivery, IEEE Transactions on, Vol. 18 (2), pp. 571 -578 (2003).
[5] J. Ross, Fuzzy Logic with Engineering Applications, McGraw-Hill, NY (1995).
[6] J. Jang, C. Sun and E. Mizutani, Neuro-Fuzzy and Soft Computing, Prentice Hall, NJ (1997).
[7] S. M. Labuda, Shahidehpour, "A Fuzzy Multi Objective Approach to Power System State Estimation", Expert Systems Applications to Power Systems, pp. 218-223 (1993).
[8] S. Soliman and N. H. Abbasy, "Power System Frequency Relaying Based on Fuzzy Linear Regression Technique", MEPCON Conf., Alexandria University, pp. 552-558 (1997).
[9] J. Wood, Power Generation Operation and Control, John Wiley & Sons (1984).

Received March 30, 2003
Accepted July 30, 2003