Non-Technical Loss Detection Using Smart Distribution Network Measurement Data

Yuan-Liang Lo, Shih-Che Huang, and Chan-Nan Lu

***Abstract*--Nowadays distribution system operators are equipped with many real time operation data from SCADA, intelligent electronics devices (IED), automatic meter reading (AMR) and advanced metering infrastructure (AMI) systems. Customer and transformer loading data obtained from AMI system at different time intervals give an opportunity to develop a distribution state estimation (DSE) algorithm. DSE is enabler of many smart distribution applications. In this paper, a DSE based quasi real time search for potential irregularity of electricity usage is presented. The proposed approach can be used to reduce non-technical loss in the distribution network.**

***Index Terms*--Advanced metering infrastructure, distribution state estimation, meter tampering, non-technical loss detection.**

I. INTRODUCTION

Electric power loss is the difference between the total energy generated/bought and what is billed. Losses can be divided into two types: 1) technical and 2) non-technical. Technical loss is due to physical characteristics of power equipment, i.e., energy lost during the transport, transformation, and measurement. Non-technical losses (NTL) generally involve energy which is delivered but not billed. They are strongly linked to illegal use of electricity.

Utility companies around the world have such losses. Besides improving the system operation efficiency, fraud detection is important in order to avoid non-technical energy losses. Intelligent and statistical methods were proposed to detect non-technical loss [1-3]. Authors of [3] presented a novel graph-based approach for automatic recognition of non-technical losses. The approach determines whether a user is becoming an illegal consumer by means of an optimum-path forest computation based classifier.

While other methods may provide more accurate estimates for non-validated data, the statistical approaches often use historical data to provide estimates that are representative of historical consumption. Common methods for estimation are [4]:

(1) Historic Estimation

The section of data needing estimation can be performed by averaging intervals from like day types to create a daily profile for the period to be estimated. There are two steps in the historic estimation procedure using baselines: Step 1: develop an average daily profile for each period to be estimated. Step 2: Use the average daily profile to estimate the usage data. The estimated value for each interval is simply the average interval value from the calculated daily profile. The technique generates estimates that are typical of recent behavior as opposed to trying to match historical usage to the profile of the meter read block being estimated.

(2) Class Load Profile

The most common intended use case of class load profile estimation is for intervals that cannot be historically estimated because of no “like days”. If the class profile is available, the class profile interval values are scaled using the average daily usage of the interval channel. It sums up the interval data values in the class profile and divides by the number of days in the class profile to obtain the average profile daily usage in kWh per day. The average daily usage is then obtained for the interval channel.

Note that the traditional statistic estimation methods assume that the historical days are a good match for the profile of the meter read block being estimated and implicitly assumes that no large changes in consumption behaviors have occurred. However, during meter data validation procedure, there is always some risk that the estimated value will differ from actual consumption. Every effort must be made to ensure each estimate reflects accrual consumption to the extent possible.

In the smart grid environment, the AMI functions require the system to measure and monitor system data, to manage the data for use by other systems such as the billing system, and to present the load profile data to customers for energy management. Meter reads of customer’s kW and kWh are transmitted to meter data management (MDM) database at certain interval. Non-valid data could be due to meter defect, communication problem or illegal electricity usage. AMI system shall be able to detect data abnormality. In order to perform efficient non-technical loss detection and avoid extensive check of millions of meter data, a DSE technique which provides a guided search of potential irregularity of electricity is proposed in this paper.

II. GUIDED SEARCH FOR NON-TECHNICAL LOSS

DSE is a key enabler for many smart distribution grid applications including outage management, loss reduction, demand response, adaptive over current protection and distributed generation dispatch [5-10]. DSE provides quasi real time network model to support system operations. It is also capable of providing potential locations and customers with abnormality of electricity usage.

The authors are with the Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. The first author is also with Taiwan Power Company. (e-mail: u360192@taipower.com.tw)

The DSE procedure used to estimate meter defect and/or plausible fraud location is shown in Fig. 1. In this study, it is assumed that distribution network configuration is available from distribution SCADA and outage management system (OMS). Feeder terminal units (FTU) provide operation data from automated feeder switching points. Feeder nodes (distribution transformers) demands are derivable from AMI/AMR database that provides customer demand data in every 15 minutes.

With sufficient measurements, weighted least square (WLS) technique is commonly used as a basis for state estimation in power systems. The WLS solution finds the minimal difference between measurements and estimated values, and takes account of variance for each measurement. The optimization problem is as follow:

$$\min_x \mathrm{JX} = \sum_{i=1}^{m} \frac{[z_i - h_i(x)]^2}{\sigma_{ii}^2} \tag{1}$$

where z_i is measurement, $h_i(x)$ is the calculated value of function corresponding to each measurement. σ_{ii}^2 is the variance for the i[th] measurement. Define R_z as an m×m diagonal matrix with each entry equals to $1/\sigma_{ii}^2$, and use the gradient method, the solution of (1) can be obtained by iteratively solving the following equation.

$$x_{k+1} = x_k + [H_k^T R_z^{-1} H_k]^{-1} H_k^T R_z^{-1}[z - h(x_k)] \tag{2}$$

where $H_k = \dfrac{\partial h(x)}{\partial x}$ is the Jacobian matrix of the measurement set. $H_k^T R^{-1} H_k = G\_matirx$ is the gain matrix of the normal equation. $z - h(x) = \Delta Z$ is the measurement mismatch vector. The final solution is obtained if the change in state between two iterations is less than a pre-defined tolerance.

When there exist measurement errors and/or errors in the network model, the estimation result would not satisfy the precision we desire. In this study, it is assumed that the network configuration used for DSE is correct. In the estimation process, the removal and downplaying the effect of bad data would result in a better DSE solution. The first step of bad data detection is to calculate the normalized residual (NR) by using the covariance matrix of measurement estimation. The measurement covariance matrix is defined as $R_{\hat{z}} = HG^{-1}H^T$, $G^{-1} = (H^T R_z^{-1} H)^{-1}$. Using the measurement covariance matrix and estimation residual $R_{\hat{r}} = R_z - R_{\hat{z}}$, the normalized residual ($\hat{r}^n$) can be calculated by the following:

$$\hat{r}^n = \Delta Z / \sqrt{diag(R_{\hat{r}})} \tag{3}$$

As shown in Fig. 2, if the objective function JX is greater than a statistical threshold determined based on the problem dimension and Chi-square distribution, there exists a bad data. Excluding the possible topology error, under normal condition, the measurement with the largest NR is identified as bad data. It is removed from the measurement set and the estimation process is repeated. After its removal and filling in reasonable estimates, the effect of bad data could be downplayed and result in better DSE solutions.

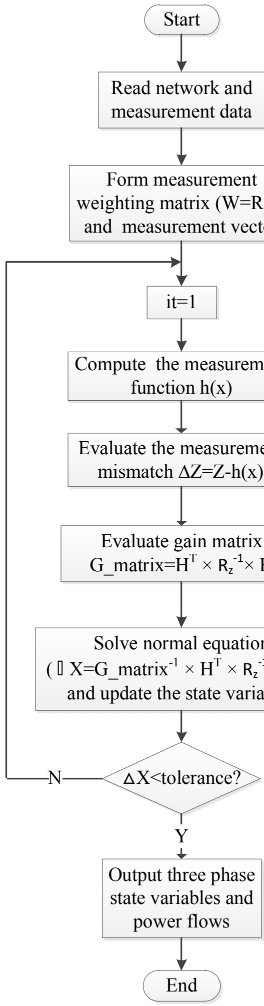

Fig. 1: Flow diagram of a three phase DSE

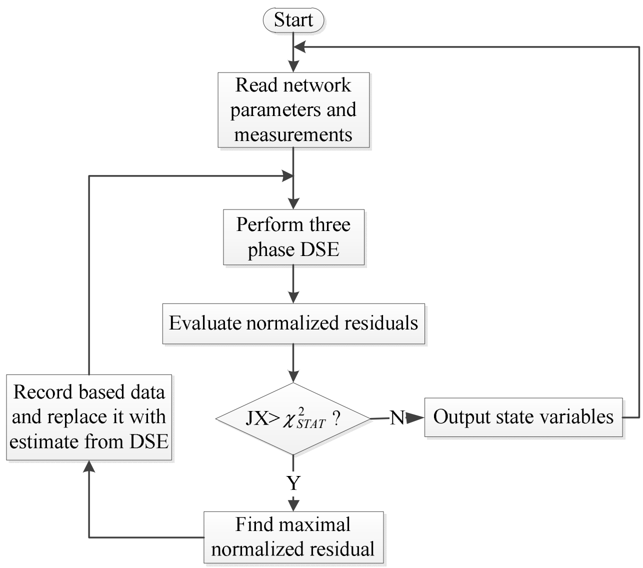

Fig. 2: Bad data detection

The DSE can be used to provide daily load profile at each feeder bus (distribution transformer). Based on significant behavior that emerges due to irregularities in consumption, feeder buses with bad (inconsistent) data would suggest locations of meter data abnormality or illegal connection. This results in a shortlist of problem locations for onsite inspection. Statistical methods can then be used to compare the demand differences in recorded customer load profile and baseline data of each customer. Potential bad data and fraud suspects could be identified.

III. NUMERICAL RESULTS

Fig. 3 shows a 11 bus test system modified from IEEE 13 bus test system [10]. Appendix A shows a sample load profile of the test system. The measurements used in the study are voltage and branch flow measurements at the main feeder branching point and voltage and consumption measurements at each load points (feeder bus). As shown in Fig. 3, there are three different types of customers, residential, commercial and industrial customers who are served from one, two and three phase distribution transformers.

Experience indicates that radial and weakly meshed distribution networks would have short branch length and lower differences in feeder bus voltages. A proper design of the measurement weightings is required when the distribution network under study has low branch loading. The weightings should tuned to parallel with the magnitude of ΔZ to ensure that a bad data can be detected while avoid identifying metering accuracy errors (e.g. 0.2% in the 0.2 class meter) as bad data. Table I shows the measurement weightings used in this study. Appendix A shows the load profile used in the tests and the results are shown in the followings.

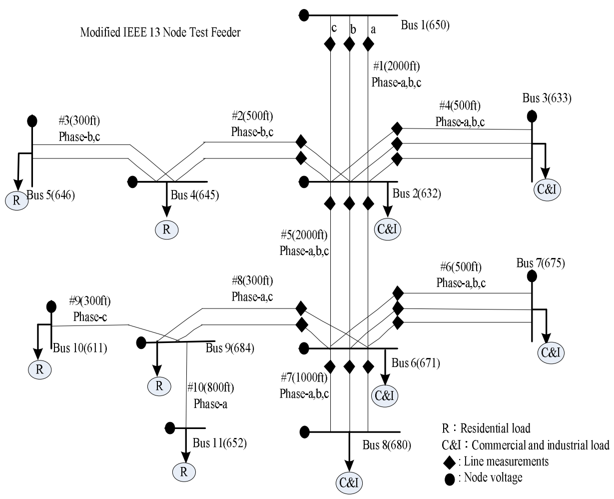

Fig. 3: 11 bus test system

TABLE I
STANDARD DEVIATION OF MEASUREMENTS

Data type	P_{inj} and Q_{inj}	P_{branch} and Q_{branch}	V_M
Standard deviation	0.001	0.001	10^{-5}

Case 1 Detection of a bad data in feeder bus demand

Table II shows the actual demands, measurements and DSE solution after the removal of bad data occurred at bus 10 of the test system at 12:00. Due to a defective meter reading at phase c of bus 10, a sum of 6.5 kW is obtained for the AMI system which is 10 kW lower than the actual usage. From the DSE results shown in Table II, it can be seen that the estimated demand at phase c is smoothed out by neighboring measurements, and the estimated load is close to the actual value. JX value in this case is 139.9757 before bad data removal which is higher than the threshold value 50.9985. JX reduces to 0.0245 after deleting the bad data and the demand estimated is 16.5 kW at the bus.

TABLE II
DSE SOLUTION AT 12:00 WITH BAD DATA AT BUS 10
(Unit: kW, kVAr, kV)

Meas. Type	Actual Demand Data			Meter Measurements			SE result after remaining bad data		
phase	a	b	c	a	b	c	a	b	c
δ_1	0	4.2	2.1	0	4.2	2.1	0	4.2	2.1
$P_{inj,9}$	-24.3	N/A	-29.4	-24.3	N/A	-29.4	-24.3	N/A	-29.4
$P_{inj,10}$	N/A	N/A	-16.5	N/A	N/A	-6.5	N/A	N/A	-16.5
P_{12}	539.6	815.6	770.5	540.6	816.4	759.8	540.5	816.5	770.0
$Q_{inj,9}$	-4.3	N/A	-5.3	-4.3	N/A	-5.3	-4.3	N/A	-5.3
$Q_{inj,10}$	N/A	N/A	-1.7	N/A	N/A	-0.7	N/A	N/A	-1.7
Q_{12}	174.9	262.7	244.0	175.2	263.2	242.3	175.2	263.1	244.5
V_{M1}	4.160	4.160	4.160	4.160	4.160	4.160	4.160	4.160	4.160
V_{M9}	4.128	N/A	4.119	4.127	N/A	4.120	4.128	N/A	4.119
V_{M10}	N/A	N/A	4.118	N/A	N/A	4.120	N/A	N/A	4.118

Case 2 Detection of two bad data in feeder bus demand

Table III shows the result of a case with two bad data at phase b of bus 4 and phase a of bus 11 at 12:00. The demands are 10 kW and 5 kW lower than actual usages. Bad data at bus 4 is detected first and removed. After fill in an estimate for bus 4 by DSE, the bad data at bus 11 is then detected and removed. DSE provide more accurate model that can be used for distribution operation applications.

TABLE III
DSE SOLUTION AT 12:00 WITH TWO BAD DATA AT BUS 4 AND BUS 11 (Unit: kW, kVAr, kV)

Meas. Type	Actual Demand Data			Meter Measurements			SE result after remaining bad data		
phase	a	b	c	a	b	c	a	b	c
δ_1	0	4.2	2.1	0	4.2	2.1	0	4.2	2.1
$P_{inj,2}$	-84.0	-216.7	-177.5	-83.9	-216.6	-177.6	-83.9	-216.6	-177.5
$P_{inj,4}$	N/A	-38.2	-29.6	N/A	-28.2	-29.6	N/A	-38	-30
$P_{inj,5}$	N/A	-24.1	-30.8	N/A	-24.0	-30.8	N/A	-24.1	-30.8
$P_{inj,9}$	-24.3	N/A	-29.4	-24.3	N/A	-29.4	-24.2	N/A	-29.4
$P_{inj,11}$	-18.1	N/A	N/A	-13.1	N/A	N/A	-18.1	N/A	N/A
P_{12}	539.6	815.6	770.5	535.4	816.4	770	540.4	816.4	770
V_{M1}	4.160	4.160	4.160	4.160	4.160	4.160	4.160	4.160	4.160
$Q_{inj,2}$	-23.9	-61.6	-50.5	-23.9	-61.6	-50.5	-23.9	-61.6	-50.6
$Q_{inj,4}$	N/A	-5.5	-4.3	N/A	-4.1	-4.3	N/A	-5.7	-3.9
$Q_{inj,5}$	N/A	-2.4	-3.1	N/A	-2.4	-3.1	N/A	-2.4	-3.2
$Q_{inj,9}$	-4.3	N/A	-5.3	-4.3	N/A	-5.3	-4.3	N/A	-5.3
$Q_{inj,11}$	-2.6	N/A	N/A	-1.9	N/A	N/A	-2.6	N/A	N/A
Q_{12}	174.9	262.7	244.0	174.4	261.3	244.6	175.2	263.1	244.5

Case 3 Single meter defect with higher than usual usage

There exists a possibility that the usage of identified customer with bad data might have large difference (30%) from the normal usage. Fig. 4 shows that the proposed method would have much better estimates than those using historical estimation. The baseline_3W curve shown in Fig. 4 is the customer usage baseline obtained from customer's past three week usage data. After checking with neighboring measurements and the Kirchhoff's laws during the studied period, it can be seen from Fig. 4 and Table IV that the DSE guided search method provides an estimate that is close to the actual usage of the customer if there is only one defective meter under the identified transformer. If the customer's usage baselines are adopted to calculate usage estimates as in the traditional meter data validation, estimation and editing (VEE) process, Table V shows the differences (i.e., the NTL) between actual usage and meter recorded data, values estimated by baselines and the proposed method. It can be seen that the power company will have revenue loss of 25.4 kWh if the meter defect is remained undetected. On the other hand, the NTL would reduce to 4.1 kWh if the proposed method is adopted. We conclude that if the customer has a larger variation in the energy consumption, the traditional VEE method would be less effective, and the proposed method can trace better the actual usage of the customer.

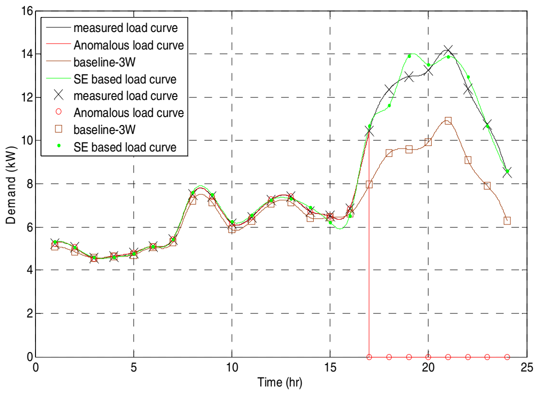

Fig. 4: Irregularly high usage with meter defect in IEEE test system (Case 3)

TABLE IV
RESULTS SUMMARY OF CASE 3

time	17hr	18hr	19hr	20hr	21hr	22hr	23hr	24hr
Actual data	10.4568	12.3455	12.9491	13.2301	14.1719	12.3947	10.7294	8.5221
Measured data	0	0	0	0	0	0	0	0
Baseline_3W	7.9684	9.4026	9.5911	9.9143	10.8989	9.0997	7.9066	6.2815
DSE solution	10.6598	11.6196	13.9074	13.4966	13.8860	12.9388	10.6565	8.5914

TABLE V
NONTECHNICAL LOSS IN CASE 3

Bus 5	Meter data	Baseline_3W	DSE
(Cus.3)	86.0616	25.3773	4.0959

IV. CONCLUSION

AMI meter data in conjunction with SCADA data are used for feeder bus load estimation and anomalous meter data detection. Test results indicate that DSE is a useful tool in providing daily load profile at each feeder node for distribution operations. When the estimation result does not satisfy the precision that we desire, there exist bad data in the SE model. If the network topology or parameters are validated, anomalous meter data would suggest locations of meter defects. This results in a shortlist of problem locations for data verification or onsite inspection. Based on the significant behavior that emerges due to irregularities in the recorded demand, NTL is recoverable by using the modern smart grid technology.

V. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the supports from Taiwan Power Company and National Science Council of Taiwan under contracts TPC-546-4840-0001, NSC 100–3113-P-110-004 and NSC 100-2221–E–110-001-MY3.

VI. REFERENCES

[1] A. Nizar, Z. Dong, and Y.Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 946–955, Aug. 2008.
[2] I. Davidson, "Evaluation and effective management of nontechnical losses in electrical power networks," in *Proc. 6th IEEE Africon Conf.* 2002, vol. 1, pp. 473–477.
[3] Caio César Oba Ramos, Andrá Nunes de Sousa, João Paulo Papa, and Alexandre Xavier Falcão, "A new approach for nontechnical losses detection based on optimum-path forest," *IEEE Trans. Power Syst.*, vol. 26, no. 1, Feb. 2011, pp., 181–189.
[4] VEE Standard for the Ontario Smart Metering System, Issue 3.0, March 31, 2010.
[5] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power Syst.*, vol. 9, Aug. 1994, pp., 1601–1609.
[6] J. Wan and K. N. Miu, "Weighted least squares methods for load estimation in distribution networks," *IEEE Trans. on Power Systems*, Vol. 18, No. 4, November 2003, pp. 1338-1345.
[7] C. N. Lu, J. H. Teng, and W. E.-H. Liu, "Distribution system state estimation," *IEEE Trans. on Power Systems*, Vol. 10, No. 1, Feb. 1995, pp. 229-240.
[8] Ke Li, "State estimation for power distribution system and measurement impacts," *IEEE Transactions on Power Systems*, Vol. 11, No. 2, pp. 911–916, May, 1996.
[9] A. Monticelli, State Estimation in Electric Power Systems A Generalized Approach,. Kluwer Academic Publishers, 1999.
[10] Haibin Wang, Distribution System State Estimation, Michigan Technological University, 2002.

VI. APPENDIX

TABLE A
AMI METER DATA USED IN THIS STUDY (UNIT: P.U. FOR BASE VALUE WITH 100KVA)

	Bus 2			Bus 3			Bus 4		Bus 5		Bus 6		
hour	phase a	phase b	phase c	phase a	phase b	phase c	phase b	phase c	phase b	phase c	phase a	phase b	phase c
1	0.6881	1.0555	0.5142	0.5941	0.6449	0.2689	0.3234	0.2828	0.3014	0.2759	0.6892	1.1892	0.4613
2	0.5866	0.9888	0.4922	0.5972	0.6836	0.3395	0.2920	0.2497	0.2536	0.2476	0.6034	1.14316	0.4475
3	0.6269	0.9688	0.4986	0.5958	0.6949	0.3312	0.2667	0.2312	0.2342	0.2342	0.6905	0.8862	0.5364
4	0.6377	0.8328	0.4435	0.7159	0.5636	0.2721	0.2645	0.2258	0.2196	0.2181	0.6287	0.9314	0.4937
5	0.5593	0.9289	0.4275	0.4872	0.5849	0.3339	0.2494	0.2112	0.1892	0.2136	0.5952	0.8799	0.4883
6	0.5761	1.0275	0.4363	0.5524	0.7136	0.3028	0.2385	0.2035	0.1995	0.2007	0.7249	0.9008	0.4431
7	0.5581	1.1416	0.4854	0.5767	0.7081	0.3708	0.2438	0.2215	0.1998	0.2112	0.7741	1.0354	0.4564
8	0.6545	1.5088	0.9506	0.5452	0.8380	0.5269	0.2648	0.2405	0.2122	0.2330	1.0013	1.3601	0.8166
9	0.7040	2.0967	1.6578	0.7396	1.1392	1.1129	0.3330	0.2786	0.2384	0.2764	1.3778	1.9586	1.0124
10	0.8546	1.7868	1.9062	0.9125	1.2228	1.1240	0.3553	0.2861	0.2468	0.2873	1.4424	1.9755	1.2969
11	0.7933	2.0786	2.2079	0.8979	1.3036	1.3567	0.3753	0.3072	0.2567	0.3061	1.4894	2.1191	1.6094
12	0.8388	2.1656	1.7755	0.6958	1.3443	1.0707	0.3823	0.2960	0.2402	0.3083	1.6266	2.1933	1.6406
13	0.6631	2.5113	1.8584	0.7279	1.3617	1.1587	0.3903	0.3055	0.2578	0.3147	1.6978	1.7653	1.3965
14	0.757	2.2033	1.8893	0.7547	1.6091	1.4336	0.3910	0.3079	0.2617	0.3117	1.5803	1.9570	1.5459
15	0.8563	2.6647	1.8784	0.9752	1.2941	1.5269	0.3668	0.3038	0.2353	0.3045	1.6451	2.2672	1.6486
16	0.7166	2.5166	2.0018	0.7220	1.4144	1.5143	0.3737	0.2959	0.2306	0.3011	1.578	1.9486	1.3245
17	0.7667	2.1774	1.7386	0.6982	1.2978	1.2061	0.3884	0.3293	0.2602	0.3196	1.3904	1.8851	1.2879
18	0.5760	2.0450	1.3686	0.6394	1.2776	0.8583	0.4185	0.3702	0.3070	0.3554	1.3376	1.5467	1.3710
19	0.5397	1.7308	1.2463	0.7052	1.2527	0.6568	0.4754	0.4203	0.3556	0.3991	1.1994	1.4544	1.1659
20	0.5545	1.8894	1.4441	0.5741	1.1728	0.7209	0.5047	0.4361	0.3601	0.4321	1.1336	1.2982	1.2710
21	0.5992	1.6191	1.1774	0.7357	0.9945	0.6585	0.5755	0.5063	0.4232	0.4857	0.9712	1.2288	1.1834
22	0.5916	1.3279	1.1170	0.6785	1.0672	0.4794	0.5461	0.4936	0.4472	0.4743	0.8483	1.0563	1.3326
23	0.5840	1.0579	1.0033	0.5239	1.0710	0.3143	0.4612	0.4497	0.4494	0.4111	0.9031	1.0948	0.9058
24	0.5542	0.9550	0.6878	0.5013	0.8381	0.2374	0.4212	0.3841	0.4020	0.3617	0.7778	0.8816	0.7688

	Bus 7			Bus 8			Bus 9		Bus 10	Bus 11
hour	phase a	phase b	phase c	phase a	phase b	phase c	phase a	phase c	phase c	phase a
1	0.7156	0.7158	0.9544	0.1815	0.1815	0.1556	0.2219	0.2755	0.1672	0.1568
2	0.6480	0.6480	0.8640	0.2708	0.2708	0.2321	0.1971	0.2384	0.1482	0.1394
3	0.6756	0.6756	0.9008	0.2243	0.2243	0.1922	0.1857	0.2250	0.1333	0.1276
4	0.6597	0.6597	0.8796	0.1700	0.1700	0.1457	0.1798	0.2132	0.1303	0.1253
5	0.5824	0.5824	0.7766	0.2409	0.2409	0.2065	0.1656	0.1990	0.1259	0.1221
6	0.6791	0.6791	0.9055	0.2675	0.2675	0.2293	0.1601	0.1924	0.1163	0.1172
7	0.7624	0.7624	1.0166	0.2311	0.2311	0.1981	0.1659	0.2022	0.1214	0.1252
8	0.8311	0.8311	1.1082	0.3143	0.3143	0.2694	0.1817	0.2171	0.1345	0.1466
9	1.0633	1.0633	1.4177	0.3400	0.3401	0.2915	0.2192	0.2608	0.1605	0.1699
10	1.1492	1.1492	1.5323	0.4437	0.4437	0.3803	0.2327	0.2762	0.1570	0.1711
11	1.2277	1.2277	1.6370	0.4586	0.4586	0.3931	0.2431	0.2944	0.1665	0.1849
12	1.2308	1.2308	1.6411	0.5624	0.5624	0.4820	0.2425	0.2939	0.1650	0.1811
13	1.2437	1.2437	1.6582	0.6297	0.6297	0.5397	0.2495	0.2921	0.1690	0.1839
14	1.1789	1.1789	1.5719	0.7371	0.7371	0.6318	0.2506	0.3045	0.1734	0.1839
15	1.2235	1.2235	1.6314	0.9157	0.9157	0.7849	0.2380	0.2877	0.1557	0.1863
16	1.1965	1.1965	1.5953	0.9404	0.9404	0.8060	0.2358	0.2810	0.1607	0.1757
17	1.2064	1.2064	1.6086	0.7633	0.7633	0.6542	0.2548	0.3103	0.1737	0.1974
18	1.1309	1.1309	1.5078	0.5755	0.5755	0.4933	0.2832	0.3390	0.1941	0.2196
19	0.9816	0.9816	1.3088	0.4685	0.4685	0.4016	0.3111	0.3869	0.2165	0.2466
20	1.1422	1.1422	1.5230	0.4914	0.4914	0.4212	0.3296	0.4049	0.2173	0.2496
21	1.1258	1.1258	1.5010	0.4088	0.4088	0.3504	0.3806	0.4622	0.2416	0.2935
22	0.8706	0.8706	1.1608	0.2994	0.2994	0.2567	0.3746	0.4521	0.2326	0.2665
23	0.9004	0.9004	1.2005	0.1715	0.1715	0.1470	0.3293	0.4010	0.2145	0.2308
24	0.7256	0.7256	0.9674	0.1077	0.1077	0.0923	0.2918	0.3547	0.2039	0.2059