

A SEPARATION TECHNIQUE FOR LOAD AND PHOTOVOLTAIC GENERATION OUTPUT IN DISTRIBUTION SYSTEM

Akihito Yasunaga, Ryoichi Hara and Hiroyuki Kita
 Hokkaido University
 Sapporo, Japan
 yasunaga@ee4-si.eng.hokudai.ac.jp

Abstract – Recently, a number of distribution generators (DG) such as photovoltaic generation have been installed in distribution systems for solving the global warming and the depletion of energy resources in the world. It will not only make the power flows in distribution systems more complex, but also cause a big problem in the restorative operation after load shedding due to the fault in grid. More specifically, since DGs are also disconnected from distribution system after the fault occurrence, the magnitude of load to be restored is different from that of net load which has been supplied from the distribution substation before the fault occurrence. Therefore, if system operators try to execute restorative operation based on the pre-fault net load, it would cause overloading and/or serious voltage drop. This paper proposes a method for separating the pre-fault net load measured into actual load and DG power output using Independent Component Analysis (ICA). Also, its validity and effectiveness of the proposed method are investigated through some computational simulations using MATLAB.

Keywords: *Independent Component Analysis, Distribution System, Photovoltaic Generation System*

1 INTRODUCTION

Generally, distribution systems are mainly constructed based on a radial configuration. Therefore, when a fault occurs at a point on the distribution feeder, the distribution systems downstream the fault point will experience supply interruption. System operators are required to detect the fault point quickly and restore the interrupted loads by operating sectionalizing switches appropriately. This problem is called the restorative control problems in power systems, and methods for determining the optimal configuration of the distribution systems so that voltage drops and/or distribution losses can be minimized have been presented [1] [2].

In the restorative operation, the magnitude of load to be restored has to be estimated in advance precisely to avoid overloading and/or voltage violation in distribution system. Conventionally, power which was flowing along the feeder before the fault occurrence has been considered as the magnitude of the current load to be restored. On the other hand, a large number of distribution generators (DGs) have been installed in distribution systems as a measure for solving the global warming and the depletion of energy resources. Especially, photovoltaic generation (PV) systems are expected to be

installed more and more in the near future. The electric energy produced by DGs is supplied partially to electric consumers in distribution system. Therefore, electric energy supplied from the distribution substation to consumers would be reduced by installing DGs. On the other hand, for the fault occurrence in the distribution system, most of DGs are disconnected from distribution system to avoid their islanding operation. More specifically, in the restorative operation after the supply interruption, more electric energy than the pre-fault state has to be sent from the distribution substation to consumers because DGs do not contribute to the restorative operation. Figure 1 shows the state transition in distribution systems caused by the fault occurrence. Figure 1(a) represents the pre-fault normal state. In this figure, it is assumed that the magnitude of load in every section is 200kW and every DG produces the electric power of 50kW. Then, the net load in every section is 150kW, which is supplied from distribution substation. When the fault occurs, the distribution system downstream the fault point experiences supply interruption

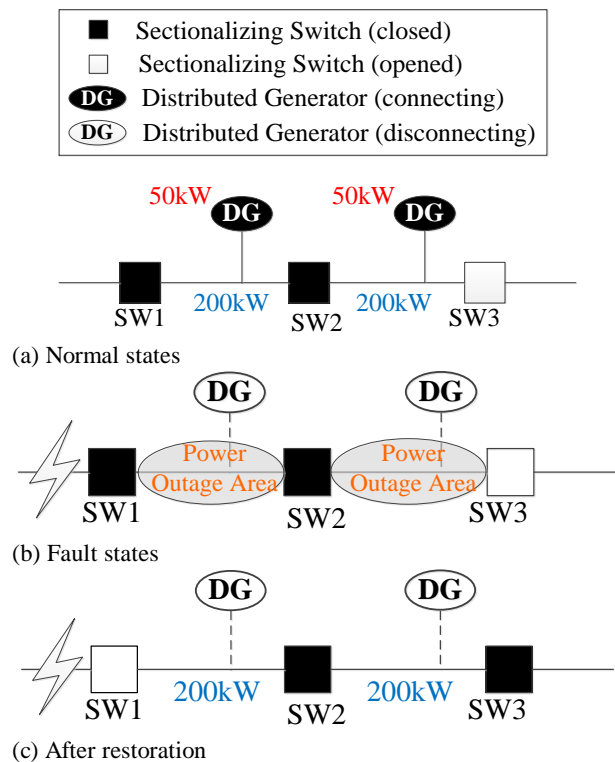


Figure1: Transition of system configuration.

and DGs are also disconnected from the distribution system as shown in Figure 1(b). Then, system operators have to restore the interrupted section by opening sectionalizing switch 1 (SW1) and closing SW3 as shown in Figure 1(c). However, since the DGs have already been disconnected, power of 200kW for the restorative operation has to be sent from the distribution substation to every section. In other words, if system operators try to send power of 150kW in the same manner as the pre-fault normal state for the restorative operation, it would be lack of 50kW. Therefore, it may cause voltage violation due to the temporary overloading. The more DGs are installed in distribution system in the future, the more overloading in the restorative operation would appear. Therefore, it may cause large voltage deviations and also may give serious effects on the system stability. More specifically, in the restorative operation, it is very important for system operators to estimate the actual load by separating PV power output from the net load.

In [3], load and DG output are estimated with Hybrid Particle Swarm Optimization. However, that is investigated under the conditions that average power output of DG, power factor and contracted kW power of every load at each section are known in advance.

This paper presents a method for separating the net load into the actual load and PV power output using Independent Component Analysis (ICA). Since the proposed method does not need the conditions mentioned before, any unknown time-series net load can be separated into actual load and PV power output.

In the following, this paper describes the outlines of the ICA in Section 2. In section 3, the details of the proposed method are explained. Then, in Section 4 the effectiveness of the proposed method is investigated through some computational simulations using MATLAB.

2 INDEPENDENT COMPONENT ANALYSIS^[4]

The ICA is a method to recognize unknown source signals from observed mixture signals. This method has been developed to resolve cocktail party problem [5] which is to separate the acoustic signals from the recorded signals, and it has been applied across the various disciplines, for example, the analysis of Electroencephalography [6] which expresses the recording of the brain's electrical activity. The ICA also has been applied to the estimation of load states in power systems [7].

The basic algorithm of the ICA is summarized as follows: Consider the relationship between the source signals $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$ and the observed signals $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ can be expressed by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t). \quad (1)$$

Where, \mathbf{A} is the $m \times n$ mixing matrix and has full rank. Equation (1) is called ICA model. The unknown source signals $\mathbf{s}(t)$ and mixing matrix \mathbf{A} can be estimated using the observed signals $\mathbf{x}(t)$ by ICA algorithm. In the following, for simplicity, the mixing matrix \mathbf{A} is assumed to be a non-singular matrix.

Generally, the following three conditions have to be satisfied to apply the ICA.

- (i) Each source signal is statistically independent.
- (ii) All the source signals are non-Gaussian distribution or only one of the source signals is Gaussian.
- (iii) The number of available observations is greater than or equal to the number of sources.

If the above conditions are satisfied, the distribution of observed signal $x_i(t)$ ($i \in m$) is closer to Gaussian than one of any source signals $s_j(t)$ ($j \in n$) by the Central Limit Theorem. Hence, if the vector which let observed signals $\mathbf{x}(t)$ be most non-Gaussian is found, one of the independent components is estimated. Therefore, the ICA is an optimization algorithm to estimate \mathbf{w} which is one of the rows of the inverse of \mathbf{A} by the maximization of *non-Gaussianity* of $\mathbf{w}^T \mathbf{x}(t)$. The flowchart of fastICA [8], which is one of the most general and popular algorithms, is shown in Figure 2. "Centering" means to transform the average of \mathbf{x} to 0 and "Whitening" means to transform the variance of \mathbf{x} to 1. These operations let the calculation of the ICA be simply and fast. The evaluation function employs the index to measure *non-Gaussianity*, for example, negentropy. Since the probability density function is needed to calculate negentropy, as mentioned in [4], the following approximation formula is employed.

$$J(\mathbf{w}^T \mathbf{x}) \propto [E\{G(\mathbf{w}^T \mathbf{x})\} - E\{G(\nu)\}]^2. \quad (2)$$

Where, ν is a Gaussian variable of zero mean and

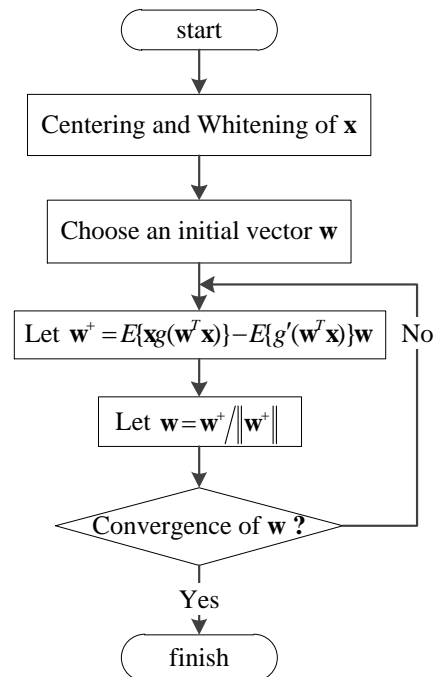


Figure2: Flowchart of fastICA.

unit variance. $E\{\cdot\}$ expresses an expectation value and G is non-quadric function, for example, $G(y) = \log \cosh(y)$. In addition, g and g' in Fig.2 are the derivative of G and g respectively. After iterative calculation, \mathbf{w} is calculated and one of the independent components is estimated. If all of \mathbf{w} , that is, the inverse matrix of \mathbf{A} is estimated, the source signals are achieved.

Since it is assumed that only the observed signals $\mathbf{x}(t)$ are known in the ICA, there are two problems that need to be solved. These problems are called permutation problem and scaling problem. In addition, if the third condition to apply the ICA mentioned before is not satisfied, some sort of devisal is needed. The methods to resolve these problems are mentioned in Section.3.

3 PROPOSED METHOD

3.1 ICA model

Figure 3 shows a section in the distribution feeder. In this figure, sectionalizing switches have sensors which can measure the active and reactive power and so on. The goal of this paper is to estimate the actual power of load, P_{load} and PV power output, P_{PV} from its net load, P in section between two switches in Figure 3 using information obtained from their sensors. To apply the ICA to this problem, the number of the observations has to be greater than two because the number of source signals to be estimated is two. Thus, this paper uses net active and reactive powers in the section between sectionalizing switches as the observed signals. The net active and reactive powers (P , Q) can be calculated using active and reactive power flows which are measured at two sectionalizing switches as follows:

$$\begin{cases} P = P_1 - P_2 \\ Q = Q_1 - Q_2 \end{cases} \quad (3)$$

Where, P_1 and Q_1 are the active and reactive power measured by SW1. Also, P_2 and Q_2 are the active and reactive power measured by SW2. Neglecting the power loss in the section, the observed P and Q are equal to the net load in the section and expressed by

$$\begin{cases} P = P_{load} - P_{PV} \\ Q = Q_{load} - Q_{PV} \end{cases} \quad (4)$$

Where, P_{load} and Q_{load} are the active and reactive powers consumed by load. Also, P_{PV} and Q_{PV} are the active and reactive power outputs produced by PV.

If the power-factor angles of load and PV are θ_{load} and θ_{PV} respectively, every reactive power can be expressed by

$$\begin{cases} Q_{load} = P_{load} \tan \theta_{load} \\ Q_{PV} = P_{PV} \tan \theta_{PV} \end{cases} \quad (5)$$

From eqs.(4) and (5), the relationship between the observed information $[P, Q]^T$ and the source signals $[P_{load}, P_{PV}]^T$ can be expressed by

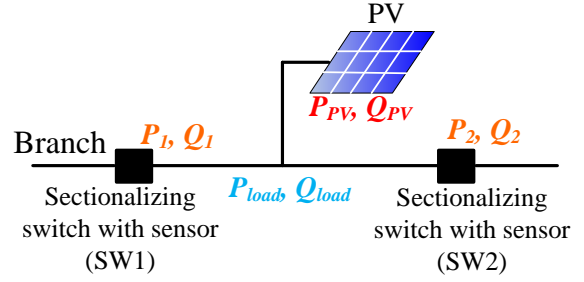


Figure3: Image of distribution system.

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ \tan \theta_{load} & -\tan \theta_{PV} \end{bmatrix} \begin{bmatrix} P_{load} \\ P_{PV} \end{bmatrix} \quad (6)$$

Since the power-factor angle of load is generally different from that of PV output, the coefficient matrix in the eq.(6) has full rank. Therefore, the method proposed in this paper has enough number of observations to estimate actual load and PV power output. Additionally, we can apply the ICA to eq.(6) because the relational equation between the source signals and the observed signals satisfies the ICA model (eq.(1)).

3.2 Statistical Independence

Generally, there is a possibility that actual load and PV output have some correlation. In this case, the statistical independence between source signals is not satisfied. Therefore, this paper applies the ICA to short cycle components (SCC) because it is expected that there is no correlation between actual load and PV power output for SCC.

The source signals can be expressed by the sum of SCC and long cycle component (LCC). Therefore, eq.(1) can be rewritten by

$$\mathbf{x}_{short}(t) + \mathbf{x}_{long}(t) = \mathbf{A}(\mathbf{s}_{short}(t) + \mathbf{s}_{long}(t)) \quad (7)$$

Where, $\mathbf{x}_{short}(t)$ and $\mathbf{x}_{long}(t)$ represent SCCs and LCCs of the observed signals. Also, $\mathbf{s}_{short}(t)$ and $\mathbf{s}_{long}(t)$ represent SCCs and LCCs of the source signals. Eq.(7) means even if LCC is filtered out, the mixing matrix \mathbf{A} is invariant. Therefore, the mixing matrix \mathbf{A} and its inverse matrix \mathbf{W} can be estimated by means of applying ICA for SCC of the observed signals. After that, multiplying the observed signals $\mathbf{x}(t)$ by \mathbf{W} , the source signals $\mathbf{s}(t)$ can be estimated.

3.3 Permutation Problem

Permutation problem means that the order of the independent components estimated by the ICA cannot be determined. More specifically, the estimated two components represent the actual load and PV power output. However, which is which cannot be discriminated. This paper solves the permutation problem by considering that PV power output in a section would have high correlation with PV power outputs in the neighboring section, although correlation between SCCs of the neighboring actual loads would be lower. Therefore, by investigating the correlation coefficient between components estimated by the ICA in the multiple sections,

the solutions which seem to express the PV power output can be found.

3.4 Scaling Problem

Scaling problem means the variances of the independent components which are estimated by the ICA cannot be determined. As shown in eq.(8), even if the estimated results are values multiplied the true source signals by a constant number, the observed signals are correct under the condition that the corresponding value of the estimated mixing matrix is a value divided the corresponding component of the true mixing matrix \mathbf{A} by the same number. Therefore, in the ICA, there is possibility that the estimated solutions are the values which are scalar multiplier in the true source signals \mathbf{s} .

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} 1/k_1 & -1/k_2 \\ \tan \theta_{load}/k_1 & -\tan \theta_{PV}/k_2 \end{bmatrix} \begin{bmatrix} k_1 P_{load} \\ k_2 P_{PV} \end{bmatrix}. \quad (8)$$

In other words, we have to determine these constant numbers k in eq.(8) in order to estimate actual load and PV power output accurately. The proposed method uses that first row's terms of the coefficient matrix in eq.(6) become a constant number respectively. More specifically, first row of the coefficient matrix in eq.(8) is normalized so that it becomes identical with the first row in eq.(6). Hence, since each source signal is also multiplied by a constant term as shown in eq.(8), actual load and PV power output can be estimated accurately.

The flowchart of the proposed method is summarized as shown in Figure 4. In this paper, actual load and PV power output are separated and estimated from the past time-series data of P and Q obtained in offline. In addition, the observed signals \mathbf{x} in Fig.4 mean $[\mathbf{P}, \mathbf{Q}]^T$ which is left-hand side in eq.(6). Various algorithms to solve the ICA have been developed [9] [10]. This paper uses the fastICA which has high reliability and is able to solve ICA fast (see Section.2). After applying the fastICA, the permutation and scaling problems for the solutions calculated by the ICA are solved by means of applying the above methods.

4 NUMERICAL ANALYSIS

4.1 Simulation Conditions

As mentioned previously, this paper separates into PV power output and actual load from daily net load data (P and Q) obtained in offline. Here, power factor of actual load was set to be 90% and power factor of PV output was set to be 100% considering that PV is connected through inverter. Of course, these parameters are given to obtain the observation data, and they are unknown values when we estimate actual load and PV power output using the ICA.

The time-series data of actual load was obtained by assuming Japanese residential customers in every season [11]. The daily load patterns in every season are shown in Figure 5. Since the load data in this figure

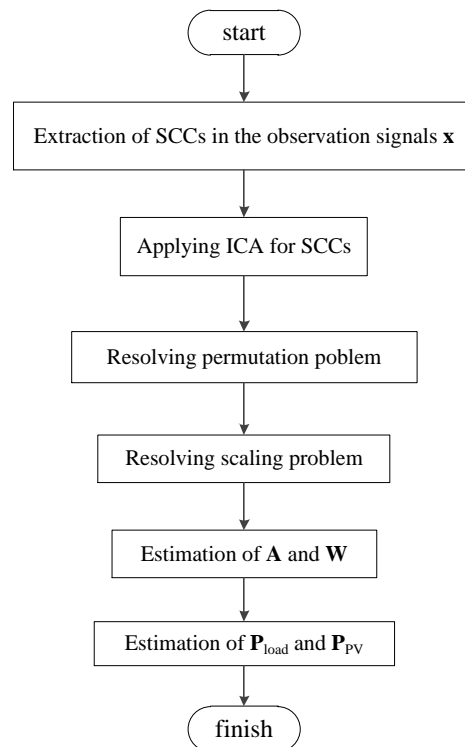


Figure4: Flowchart of the proposed method.

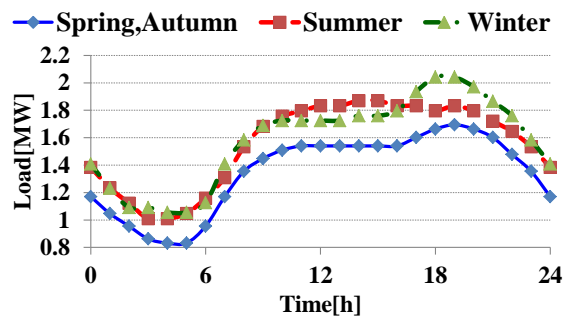
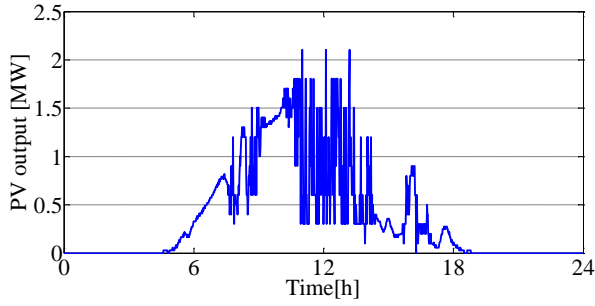


Figure5: Daily load pattern in each season.

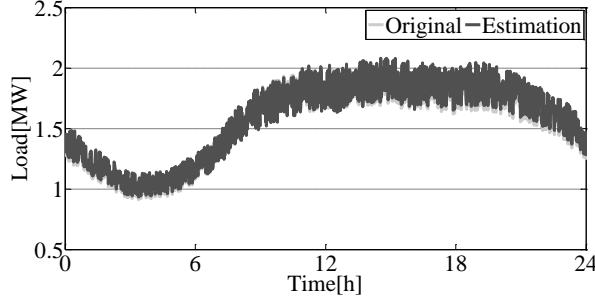
was given with hourly resolution, to derive one minute resolved load data random noise in the range of $\pm 10\%$ were added to the linear interpolated hourly mean load data. In addition, data of PV power output was made up using the dates in 2007 offered by Japan Meteorological Business Support Center [12]. The data also was one minute resolved data. The capacity of load and PV generation installed in a section were given based on data referred to *Ota, Gunma, Japan* [13]. The periodic components shorter than 20 minutes were extracted from original data to ensure the independence of source signals.

4.2 Simulation Results

Simulation results for two representative days are shown in Figs.6 and 7 respectively. Figures 6(a) and 7(a) show the actual PV power outputs for each day, and Fig. 6(b) and 7(b) show actual loads and estimated loads. Here, for simplicity, the permutation problem was solved by investigating the correlation between



(a) Actual PV output profile.



(b) Original and estimation of load profile.

Figure6: Example of a day estimated accurate.

actual PV power output and the results estimated by the ICA. As seen from comparison between Fig.6 and Fig.7, the accuracy of load estimation in Fig.6 is higher than that in Fig.7. On the other hand, the fluctuation in PV power output of Fig. 6(a) is larger than that of Fig. 7(a). In other words, there is a possibility that the accuracy of estimation depends on the magnitude of the fluctuations. Thus, the authors investigated the accuracy of loads estimated for every load patterns in 2007 except December. The obtained results are shown in Figure 8. The horizontal axis in this figure shows the negentropy defined in Section 2, and it was calculated from SCC of the actual PV power output assumed in this simulation. This index means how close the statistic characteristic of the stochastic variable is to Gaussian distribution. More specifically, the value of negentropy is closer to 0, the statistic distribution is closer to Gaussian distribution. Here, another approximation formula as the negentropy was used. As mentioned in [4], eq.(9) is essentially no different from eq.(2) though a way for approximation is different.

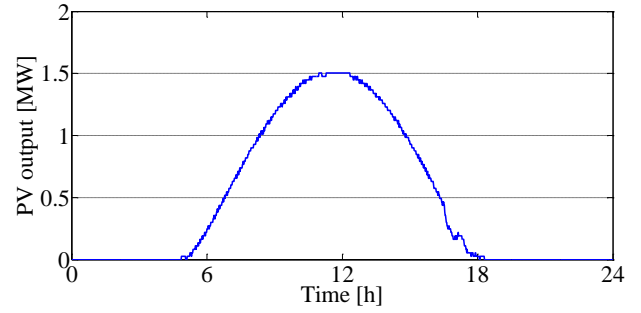
$$J(\mathbf{y}) = k_1 (E\{y_i \exp(-\frac{y_i^2}{2})\})^2 + k_2 (E\{\exp(-\frac{y_i^2}{2})\} - \frac{1}{\sqrt{2}})^2. \quad (9)$$

Where, $\mathbf{y}=[y_1, y_2, \dots, y_T]$ is the normalized SCC of actual PV power output which means its average value is 0 and variance value is 1. Also, $E\{\bullet\}$ expresses expectation value and

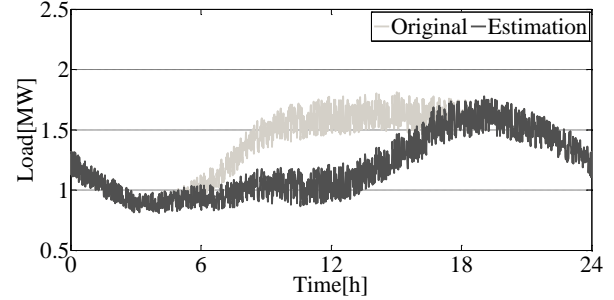
$$k_1 = \frac{36}{8\sqrt{3}-9}, k_2 = \frac{24}{16\sqrt{3}-27}.$$

In addition, the vertical axis of Fig.8 expresses the mean absolute error rate defined by

$$MAER = \frac{1}{N} \sum_{j=1}^N \left(\frac{1}{T} \sum_{i=1}^T \left| \frac{s_{true}(i,j) - s_{estimate}(i,j)}{s_{true}(i,j)} \right| \times 100 \right). \quad (10)$$



(a) Actual PV output profile.



(b) Original and estimation of load profile.

Figure7: Example of a day estimated inaccurate.

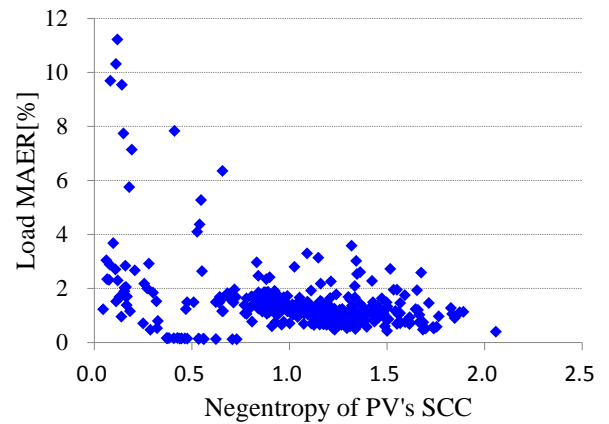


Figure8: Estimation results of load.

Where, N is the number of calculation and T is the evaluation time. In this paper, N and T are set to be 10000 and 1440(24hours \times 60minutes) respectively. The variable, $s(i, j)$ means a value of source signal in $\#i$ th minute of $\#j$ th calculation. In addition, s_{true} and $s_{estimate}$ express true value and estimated value respectively. As shown in Figure 8, when the negentropy for SCC of PV power output is larger, namely it is closer to non-Gaussian distribution, the actual load can be estimated accurately. This is because the condition for applying the ICA is that every source signal has to be non-Gaussian distribution as described in Section.2. Therefore, when the value of negentropy is smaller, it is considered that the accuracy in the estimated load becomes worse because the above condition is not satisfied.

However, as shown in Figure 8, there are some days with high accuracy of estimation even though the value of negentropy is smaller. Therefore, the authors investigated the difference between days with high accuracy

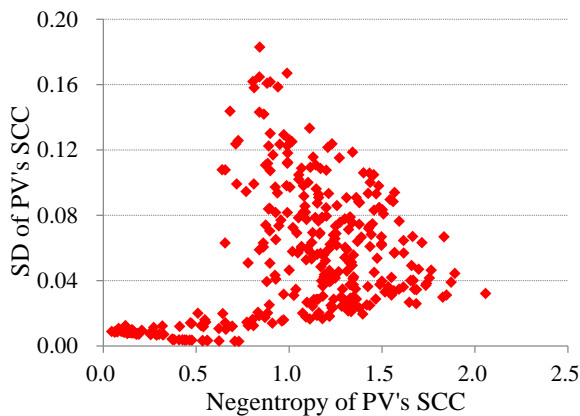


Figure9: Relationship between negentropy and SD for SCC of PV output.

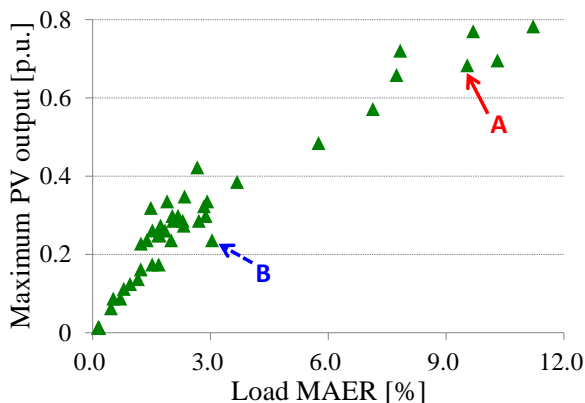


Figure10: Relationship between load MAER and maximum PV output.

and days with low accuracy when the value of negentropy is small. First, the authors calculated the standard deviation (SD) of SCC in PV power outputs for every day. The relation between the obtained SD and negentropy are shown in Figure 9. As seen from this figure, if PV power output does not have large fluctuation, the negentropy is also small; therefore, it is expected that the accuracy in the estimated load becomes worse. Secondly, the authors investigated the relation between the accuracy in the estimated load and the maximum value over daily PV power outputs. The obtained relation is shown in Figure 10. Here, only days with negentropy smaller than 0.5 are plotted in this figure. This figure shows that the smaller maximum PV power output is, the smaller MAER for the estimated load is. More specifically, even if the fluctuation in the PV power output is small, MAER becomes small when the maximum PV power output is small. Furthermore, the actual PV power outputs for two days (A and B) plotted on Figure 10 are shown in Figure 11. One day (A) is a clear day; therefore, the maximum PV power output is larger. In this case, the accuracy in the estimated load is lower as expected before. The other is a cloudy day; therefore, the maximum PV power output is smaller. In this case, the accuracy in the estimated load is higher. On a day that the maximum of PV pow-

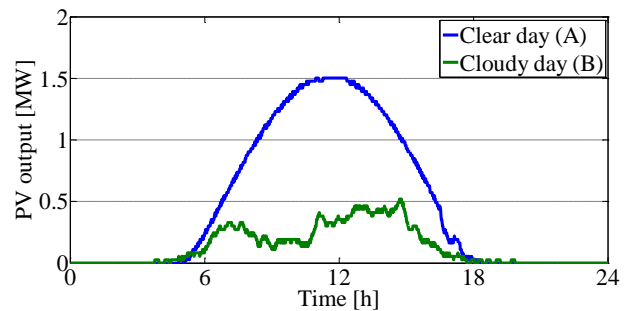


Figure11: Examples of actual PV output in a day with small value of negentropy.

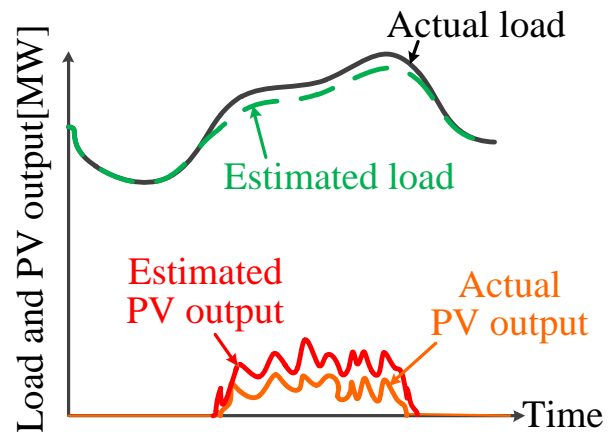


Figure12: Image of a day with high accuracy though small value of negentropy.

er output is small, SD of SCC becomes small logically. As a result, it is expected that the estimation error of PV power output does not have much influence on load MAER as seen from Figure 12 because PV power output is small. Therefore, it is seen that MAER for the actual load becomes small in spite of the fact that the value of negentropy is small.

5 CONCLUSION

This paper presented a new technique for separating actual load and PV power output in order to avoid feeder overloading caused in restoration operations. The proposed method was based on application of ICA and ways for solving permutation and scaling problems which were inherent question on the ICA were proposed. As a consequence on numerical analysis, positive outcomes could be obtained through offline simulation. On the other hand, there is still room for improvement in the case of small fluctuations of PV output. For the future, when the load power factor is variable and the actual load is estimated online, we have to try to check whether positive outcomes are derived as well as offline estimation.

REFERENCES

- [1] Y. Hayashi, J. Matsuki, "Loss Minimum Configuration of Distribution System Considering N-1 Securi-

- ty of Dispersed Generators”, IEEE Transactions on Power Systems, Vol.19, No1, pp.636-642, February 2004
- [2] Y. Y. Hong, S. Y. Ho, “Determination of Network Configuration Considering Multiobjective in Distribution Systems Using Genetic Algorithms”, IEEE Transactions on Power Systems, Vol.20, No.2, pp.1062-1069, May 2005
- [3] S. Naka, T. Genji, T. Yura, Y. Fukuyama, “A Hybrid Particle Swarm Optimization for Distribution State Estimation”, IEEE Transactions on Power Systems, Vol.18, No.1, pp.60-68, February 2003
- [4] A. Hyvarinen, J. Karhunen, E. Oja, “Independent Component Analysis”, John Wiley & Sons, 2001
- [5] S. Ikeda, N. Murata, “A method of ICA in time-frequency domain ”, In Proc. Int. Workshop on Independent Component Analysis and Signal Separation, pp.365-370, 1999
- [6] R. Vigario, J. Sarela, V. Jousmiki, M. Hamalainen, E. Oja, “Independent Component Approach to the Analysis of EEG MEG Recordings”, IEEE Transactions on Biomedical Engineering, Vol.47, Issue:5, pp.589-593, May 2000
- [7] E. Gursoy and D. Niebur, “On-line Estimation of Electric Power System Active Loads”, International Joint Conference on Neural Networks, pp.1689-1694, 2006
- [8] A. Hyvarinen, “Fast and robust fixed-point algorithms for independent component analysis”, IEEE Transactions on Neural Networks, Vol.10 Issue:3, pp.626-634, May 1999
- [9] A. J. Bell, T. J. Sejnowski, “An Information-Maximization Approach to Blind Separation and Blind Deconvolution”, Neural Computation, Vol.7, No.6, pp.1129-1159, November 1995
- [10] J. F. Cardoso, A. Souloumiac, “Blind beamforming for non-Gaussian signals”, IEE Proceedings F, Vol.140, Issue:6, pp.362-370, December 1993
- [11] Suburban system model, The institute of Electrical Engineers of Japan, <http://www2.iee.or.jp/ver2/pes/>
- [12] Japan Meteorological Business Support Center, <http://www.jmbssc.or.jp/>
- [13] Kandenko Co., Ltd., FY2002 NEDO report “Demonstrative research on clustered PV systems”, 2003