Cross-correlation of output fluctuation and system-balancing cost in photovoltaic integration

Yuichi Ikeda[1], *Kazuhiko Ogimoto*[2]

[1]*Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto 606-8501, Japan*
[2]*Collaborative Research Center for Energy Engineering, Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan*
E-mail: ikeda.yuichi.2w@kyoto-u.ac.jp

Published in *The Journal of Engineering*; Received on 27th August 2014; Accepted on 29th August 2014

Abstract: The authors analysed the cross-correlation of photovoltaic (PV) output fluctuation for the actual PV output time series data in both the Tokyo area and the whole of Japan using the principal component analysis with the random matrix theory. Based on the obtained cross-correlation coefficients, the forecast error for PV output was estimated with/without considering the cross-correlations. Then the operation schedule of thermal plants is calculated to integrate PV output using the proposed unit commitment model with the estimated forecast error. The system-balancing cost of PV system was also estimated with or without demand response. Finally, validity of the concept of 'local production for local consumption of renewable energy' and alternative policy implications were discussed.

1 Introduction

Restructuring of the electric utility industry and large-scale grid-integration of photovoltaic (PV) systems have been intensively discussed since the East Japan Earthquake of 2011. The former includes separation of electrical power generation from power distribution and transmission, and establishment of the retail power market and revitalisation of the wholesale power markets. Although the institutional design of the power markets is still an open question in Japan, the market has to be designed so as to have optimal operation schedule, which is obtained using a unit commitment calculation, through competitions between generation companies.

The large-scale grid-integration of PV systems brings another kind of problem, namely, the PV output fluctuation, into the power system operation. The planned installation capacity of PV systems in the whole of Japan will be 100 GW in 2030 [1] and about one third of the capacity will be in the Tokyo area. The major fraction of the PV system will be installed on the rooftop of the consumer's residential houses and office buildings, which are widely distributed in the Tokyo area. Therefore the forecast of PV output with high spatial resolution is an important problem to be considered, and the cross-correlations of the PV outputs will be key quantities to estimate the forecast error of PV output.

In relation to the above discussion, the concept of 'local production for local consumption of renewable energy' has been proposed in Japan. Since electric power is in large demand in the Tokyo area, the area price could be high enough to be close to the feed-in tariff price for PV power. For this reason, the concept of 'local production for local consumption of renewable energy' of PV power is considered to be economically feasible [2]. This concept is also advantageous because of the mitigation of transmission loss. However, it is to be noted that this concept needs careful consideration for PV and wind power because of the inherent nature of output fluctuation, even though it is suitable for geothermal and biomass energies [3].

In this paper, we analysed the cross-correlation of PV output fluctuation for the actual PV output time series data [4] in both Tokyo area and the whole of Japan using the principal component analysis with the random matrix theory. Based on the obtained cross-correlation coefficients, the forecast error for PV output was estimated for some extreme cases. Then the operation schedule of thermal plants was calculated to integrate PV output using our unit commitment model [5, 6] with the estimated forecast error. The system-balancing cost of PV system was also estimated with or without demand response. Finally, the validity of the concept of 'local production for local consumption of renewable energy' and alternative policy implications were also discussed.

2 Cross-correlation of PV output fluctuation

2.1 System-wide output fluctuation

The forecast of system-wide PV output is decomposed as

$$\mathrm{pv}_t^{(f)} \equiv X(t) = \sum_{i=1}^{N} x_i(t) = \sum_{i=1}^{N} c_i y_i(t) \tag{1}$$

where $y_i(t) = x_i(t)/c_i$ and c_i are the forecast of PV output per installed capacity (load factor) and the installed capacity in the ith site, respectively. Our unit commitment model [5, 6] requires the PV output forecast time series and the forecast error to estimate the optimal operation schedule with consideration of the PV output fluctuation. If both accuracy and spatial resolution of the PV forecasting are high, the forecasted time series is similar to a moving average of actual PV output for each PV site, and consequently the cross-correlation of residual time series, which is equal to subtracting the actual output from forecast output at each time point, is expected to be a white noise. Thus, the forecast error of system-wide PV output σ_X is

$$\sigma_p^2 = \sum_{i=1}^{N} \left(\frac{\partial X}{\partial y_i}\right)^2 \sigma_i^2 = \sum_{i=1}^{N} c_i^2 \sigma_i^2 \tag{2}$$

where σ_i is the forecast error of PV output per installed capacity in the ith site. On the other hand, if the spatial resolution of the forecast is low and, for example, we have just a few forecasted sites in the Tokyo area, the residual time series includes the cross-correlation between the various PV sites located in different places. In this case, we have a larger forecasting error because of the cross-correlations. The forecast error of system-wide PV output σ_X is written as

$$\begin{aligned}\sigma_p^2 &= \sum_{i=1}^{N} \left(\frac{\partial X}{\partial y_i}\right)^2 \sigma_i^2 + 2 \sum_{i=2}^{N} \sum_{j<i} \left(\frac{\partial X}{\partial y_i}\right)\left(\frac{\partial X}{\partial y_j}\right)\sigma_{ij} \\ &= \sum_{i=1}^{N} c_i^2 \sigma_i^2 + 2 \sum_{i=2}^{N} \sum_{j<i} c_i c_j \sigma_{ij}\end{aligned} \tag{3}$$

$$\sigma_{ij} = \sigma_i \sigma_j \rho_{ij} \tag{4}$$

by including covariance among different sites σ_{ij}. Here, ρ_{ij} is the cross-correlation coefficient among different sites. Generally, it is expected that the number of forecasted sites is smaller than that of the installed sites N. For instance, we cannot forecast PV output for each roof-top PV of all the residential houses and office buildings with high accuracy in the Tokyo area because of both technological and economical reasons. Therefore it is required to consider the cross-correlation σ_{ij} to estimate the forecast error of system-wide PV output σ_X.

2.2 Random matrix theory

We analysed the de-trended PV output $z_i(t)$ obtained by filtering the actual PV output time series per installed capacity $y_i(t)$ using the Fourier series expansion. In general, it is expected that correlation coefficients are associated with random noise for a fluctuating time series such as PV output. The correlation coefficient between points i and j is calculated by

$$C_{ij} = \frac{\left\langle \left(z_i(t) - \langle z_i \rangle\right)\left(z_j(t) - \langle z_j \rangle\right)\right\rangle}{\sqrt{\left(\langle z_i^2 \rangle - \langle z_i \rangle^2\right)\left(\langle z_j^2 \rangle - \langle z_j \rangle^2\right)}} \tag{5}$$

where $z_i(t)$ is the de-trended PV output at the site $i(=1, \ldots, N)$ and time $t(=1, \ldots, L)$ and $\langle \cdot \rangle$ indicates the time average for the time series.

Now we consider the eigen-value problem

$$\boldsymbol{C}|\boldsymbol{\alpha}\rangle = \lambda_\alpha |\boldsymbol{\alpha}\rangle \tag{6}$$

for the correlation matrix $\boldsymbol{C}$. λ_α and $|\boldsymbol{\alpha}\rangle$ are the eigen-value and the corresponding eigen-vector, respectively. Here, we assume that the eigen-values are arranged in decreasing order ($\alpha = 0, \ldots, N-1$). Once the eigen-values are calculated using (5) and (6), the distribution of eigen-value $\rho(\lambda)^{\mathrm{E}}$ is obtained.

According to the random matrix theory [7–10], distribution of the eigen-value for the matrix $\frac{1}{T}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$, where all elements of the matrix $\boldsymbol{H}$ are given as a random number $N(0, \sigma^2)$, is given by

$$\rho(\lambda)^{\mathrm{T}} = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda} \tag{7}$$

where

$$Q = \frac{L}{N} \tag{8}$$

$$\lambda = [\lambda_{\min}, \lambda_{\max}] \tag{9}$$

$$\lambda_{\min} = \left(1 - \frac{1}{\sqrt{Q}}\right)^2 \text{ and} \tag{10}$$

$$\lambda_{\max} = \left(1 + \frac{1}{\sqrt{Q}}\right)^2 \tag{11}$$

Equation (7) is exact at the limit $N, L \to \infty$. For a randomly fluctuating time series such as PV output, it is expected that the distribution $\rho(\lambda)^{\mathrm{E}}$ obtained by data analysis agrees with the distribution $\rho(\lambda)^{\mathrm{T}}$ calculated using (7)–(11) for $\lambda \le \lambda_{\max}$. Therefore, only a small number of eigen-values for $\lambda > \lambda_{\max}$ have the information of genuine correlation.

To extract the genuine correlation, we rewrite the correlation matrix $\boldsymbol{C}$ using eigen-value λ_α and the corresponding eigen-vector $|\boldsymbol{\alpha}\rangle$ [11]. First we define the complex conjugate vector of the eigen-vector $|\boldsymbol{\alpha}\rangle$ by

$$\langle \boldsymbol{\alpha}| = |\boldsymbol{\alpha}^*\rangle^t \tag{12}$$

For the real symmetric matrix, such as the correlation matrix $\boldsymbol{C}$, all elements of the eigen-vector $|\boldsymbol{\alpha}\rangle$ are real, and thus the complex conjugate denotes the transpose t.

Then the correlation matrix $\boldsymbol{C}$ is rewritten as

$$\boldsymbol{C} = \sum_{\alpha=0}^{N-1} \lambda_\alpha |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \tag{13}$$

by multiplying (6) with the transposed vector $\langle\boldsymbol{\alpha}|$ from the left-hand side and taking summation over α. Here, the property of the projection operator $|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|$

$$\sum_{\alpha=0}^{N-1} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| = 1 \tag{14}$$

was used. As a result, the correlation matrix $\boldsymbol{C}$ of (13) is divided into the following components

$$\boldsymbol{C} = \boldsymbol{C}^t + \boldsymbol{C}^r = \sum_{\alpha=0}^{N_t} \lambda_\alpha |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| + \sum_{\alpha=N_t+1}^{N-1} \lambda_\alpha |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \tag{15}$$

The first term $\boldsymbol{C}^t$ corresponds to the genuine correlation component ($\lambda > \lambda_{\max}$). The second term $\boldsymbol{C}^r$ corresponds to the random component ($\lambda \le \lambda_{\max}$). The term $\lambda_0|0\rangle\langle 0|$ is interpreted as the change of a whole system, such as the weather change.

We introduce the vector $|\boldsymbol{z}(t)\rangle$, which consists of the time series of PV output $z_i(t)(i = 1, \ldots, N)$. Then the vector $|\boldsymbol{z}(t)\rangle$ is expanded on the basis of the eigen-vectors $|\boldsymbol{\alpha}\rangle$ [11]

$$|\boldsymbol{z}(t)\rangle = \sum_{\alpha=0}^{N-1} a_\alpha(t)|\boldsymbol{\alpha}\rangle \tag{16}$$

The expansion coefficient $a_\alpha(t)$ is obtained using the orthogonality of the eigen-vectors

$$a_\alpha(t) = \langle\boldsymbol{\alpha}|\boldsymbol{z}(t)\rangle \tag{17}$$

The time series corresponding to the genuine correlation C^t is extracted by truncating the summation up to N_t in (16)

$$|\boldsymbol{z}(t)\rangle = \sum_{\alpha=0}^{N_t} a_\alpha(t)|\boldsymbol{\alpha}\rangle \tag{18}$$

2.3 Data analysis

The genuine components of cross-correlation of the de-trended PV output per installed capacity were studied using the random matrix theory. The analysed data are the output time series acquired every 1 h for each prefecture [4]. Before analysing the data, two preprocessings were made. First, the data during night time was removed. Then, the trend was removed from the time series by filtering out the components with a period longer than 6 h using the Fourier series expansion. Therefore, only the short-term fluctuation component remained in the time series. The auto-correlation function and fluctuation distribution for Tokyo in May are shown in Fig. 1. The memory in the auto-correlation function is lost within a few hours. This means that the trend component is removed. The kurtosis of the fluctuation distribution is 5.0849, which is significantly larger than the value expected for the normal distribution, that is, 3. This means that the actual fluctuation distribution has a longer tail than the normal distribution. The two different types of functional forms of fluctuation distribution are shown in Fig. 2. If the fluctuation is distributed according to the normal distribution, the probability density function is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{19}$$

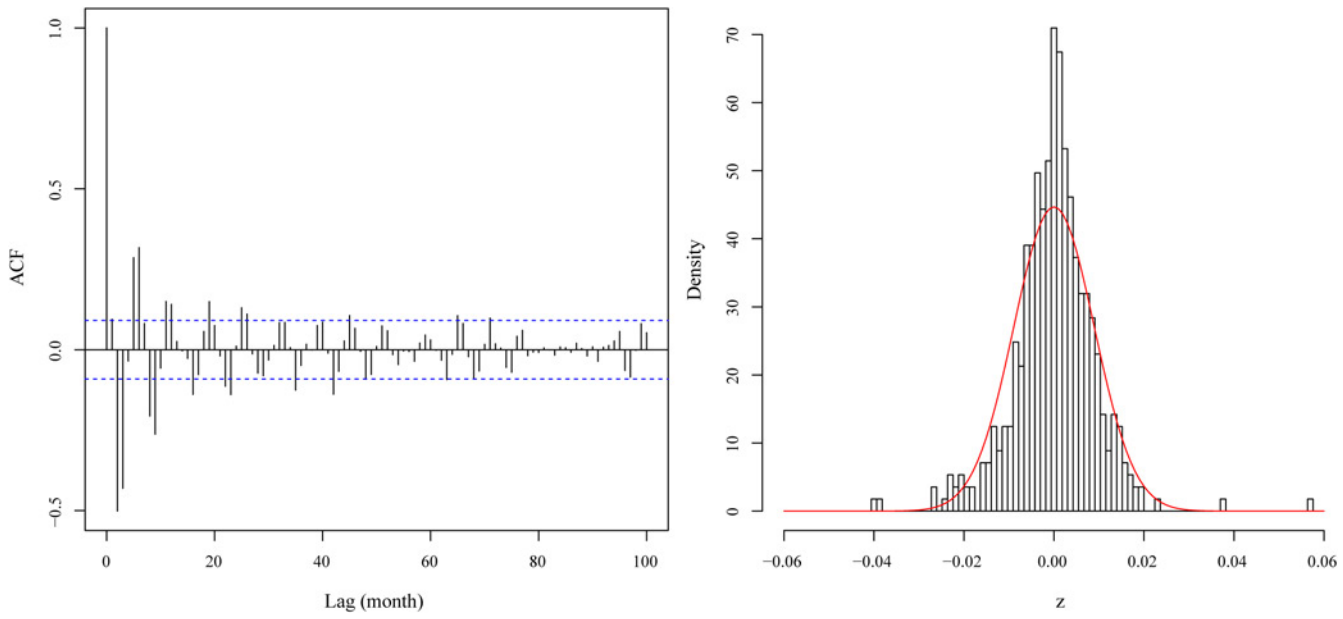

Fig. 1 *Auto-correlation function (left) and fluctuation distribution (right) for Tokyo in May*

and the cumulative distribution function is written using the error function erf[·] as

$$\phi(x) = \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{x-\mu}{\sqrt{2\sigma^2}}\right]\right) \tag{20}$$

where μ and σ are the mean and standard deviation, respectively. However, if the probability density function $p(x)$ is a Laplace distribution

$$p(x) = \frac{1}{2b}\exp\left[-\frac{|x-\mu|}{b}\right] \tag{21}$$

then the cumulative distribution function $\phi(x)$ is

$$\phi(x) = \frac{1}{2}\left(1 + \mathrm{sgn}(x-\mu)\left(1 - \exp\left[-\frac{|x-\mu|}{b}\right]\right)\right) \tag{22}$$

Here, a standard deviation is given by $\sigma = \sqrt{2}b$ and $\mathrm{sgn}(x-\mu) = +(x \geq \mu),\ -(x < \mu)$. The functional forms for these distributions are depicted for $\mu = 0$ and $\sigma = 1$ in Fig. 2. It is to be noted here that the Laplace distribution shows a distribution tail longer than the normal distribution.

Eigen-value distribution for the Tokyo area and the whole of Japan in May is shown in Fig. 3. For the Tokyo area, we calculate $\lambda_{\max} = 1.35$ using (11) with $N = 9$ and $L = 420$. The upper panel of Fig. 3 depicts that only the largest eigen-value is larger than $\lambda_{\max}$. On the other hand, for the whole of Japan, we calculate $\lambda_{\max} = 1.88$ with $N = 47$ and $L = 420$. The lower panel of Fig. 3 depicts that the five largest eigen-values are larger than $\lambda_{\max}$.

We show the distribution of genuine correlation coefficients calculated for the de-trended PV output time series in both the Tokyo area and the whole of Japan. The cross-correlation coefficients for the Tokyo area are shown in Fig. 4. Panels (a) and (c) are genuine correlation C^t and panels (b) and (d) are the random components C^r. The genuine correlation C^t was calculated using only the largest eigen-value and the corresponding eigen-vector. Fig. 4 depicts that the genuine correlation C^t has positive correlation and, on the other hand, the random components C^r distributes around 0. The first eigen-vector for the Tokyo area is shown in Fig. 5. The nine components correspond to eight prefectures, and Tokyo was included in the Tokyo area. It was noted that all vector components had the same sign. This means that the PV output fluctuates simultaneously in the same direction for all prefectures in the Tokyo area.

The cross-correlation coefficients for the whole of Japan are shown in Fig. 6. Panels (a) and (c) are genuine correlations C^t and panels (b) and (d) are the random components C^r. The genuine correlation C^t was calculated using only the five largest eigen-values and corresponding eigen-vectors. Fig. 6 depicts that the genuine correlation C^t has positive correlation and, on the other hand, the random components C^r distributes around 0. The

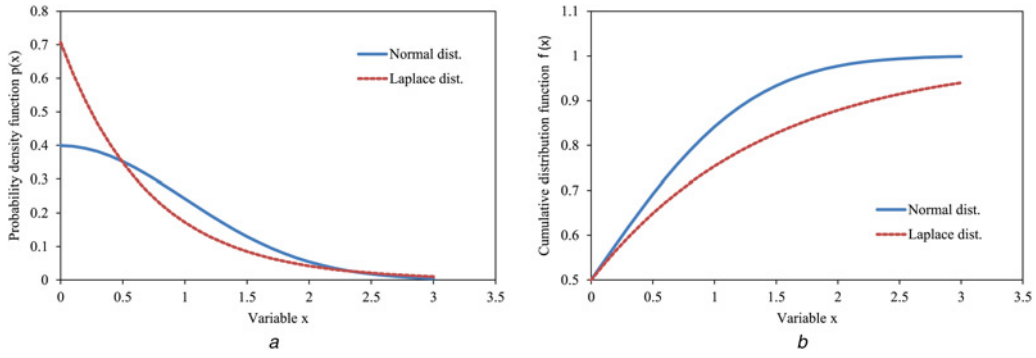

Fig. 2 *Functional form of fluctuation distribution*
a Probability density function
b Cumulative distribution function

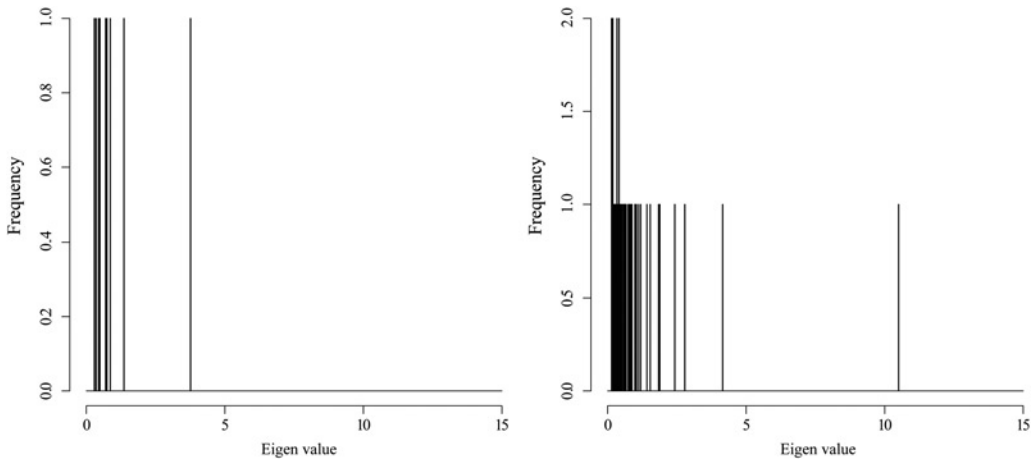

Fig. 3 *Eigen-value distribution for the Tokyo area (left) and whole of Japan (right) in May*

cross-correlation of PV output fluctuation in the Tokyo area was larger than the cross-correlation in the whole of Japan throughout the year. The first to third eigen-vectors for the whole of Japan are shown in Fig. 7. Forty seven components correspond to all prefectures from Hokkaido to Okinawa in the whole of Japan. The first eigen-vector has all components with the same sign. This means that the PV output fluctuates simultaneously in the same direction for all prefectures in the whole of Japan. The characteristic of the first eigen-vector in the whole of Japan is similar to the Tokyo area. The second eigen-vector shows the weather change between eastern and western Japan. The third eigen-vector is more complicated. These characteristics of the second to fifth eigen-vectors correspond to the smaller correlation coefficients in the whole of Japan. However, it is noted that the coefficient of variation of PV output does not decrease proportionally to $N^{-1/2}$ as the number of PV sites N increased because of the observed cross-

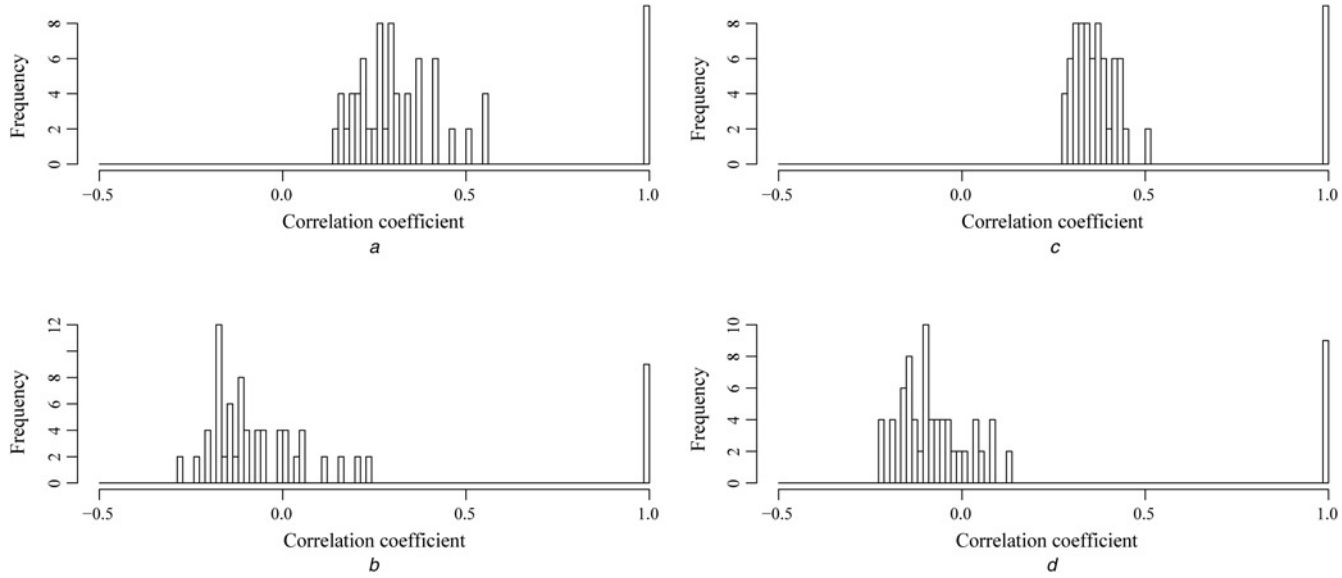

Fig. 4 *Cross-correlation coefficients for the Tokyo area in January and July*
a Genuine correlation in January
b Random component in January
c Genuine correlation in July
d Random component in July

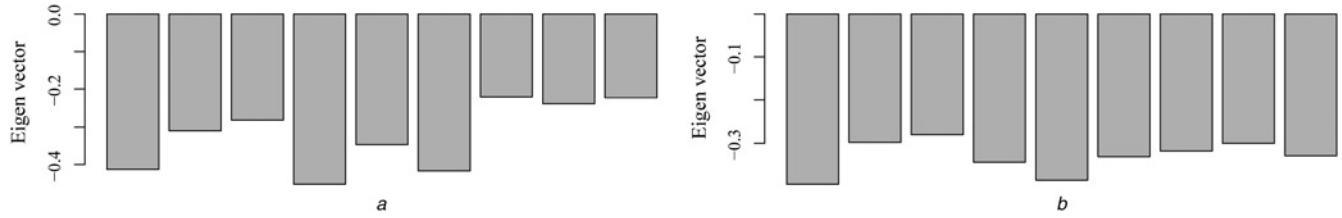

Fig. 5 *First eigen-vector for the Tokyo area in January and July*
a First eigen vector in January
b First eigen vector in July

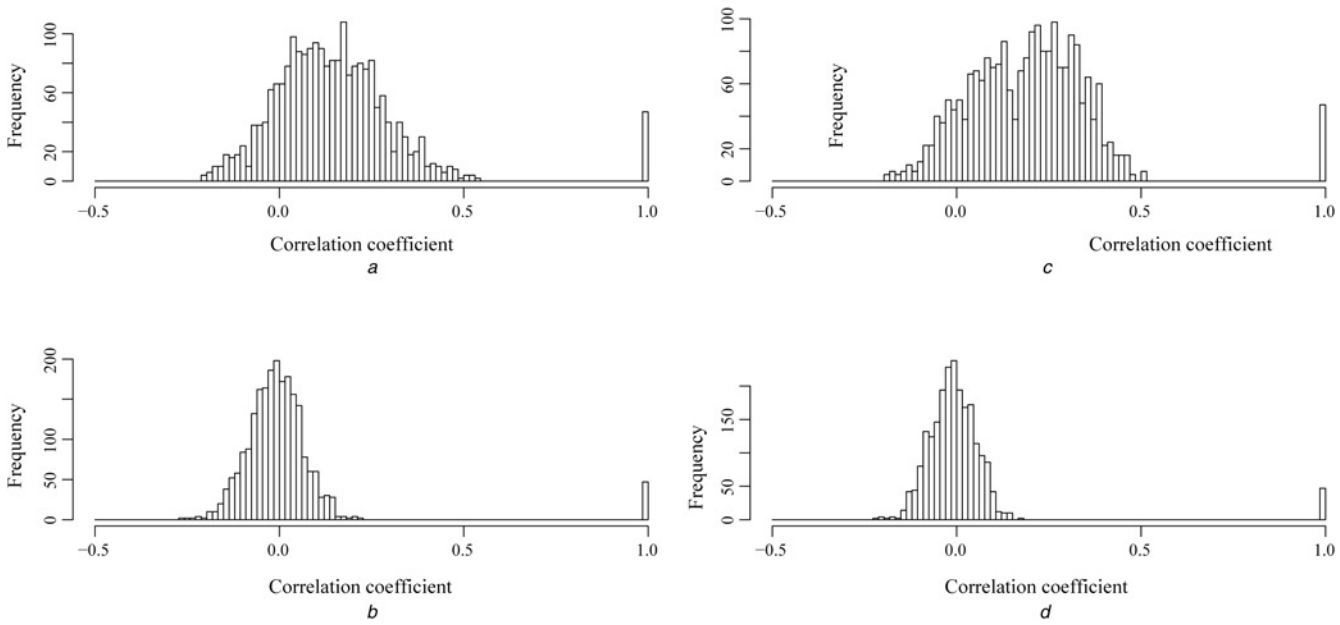

Fig. 6 *Cross-correlation coefficients for the whole of Japan in January and July*
a Genuine correlation in January
b Random component in January
c Genuine correlation in July
d Random component in July

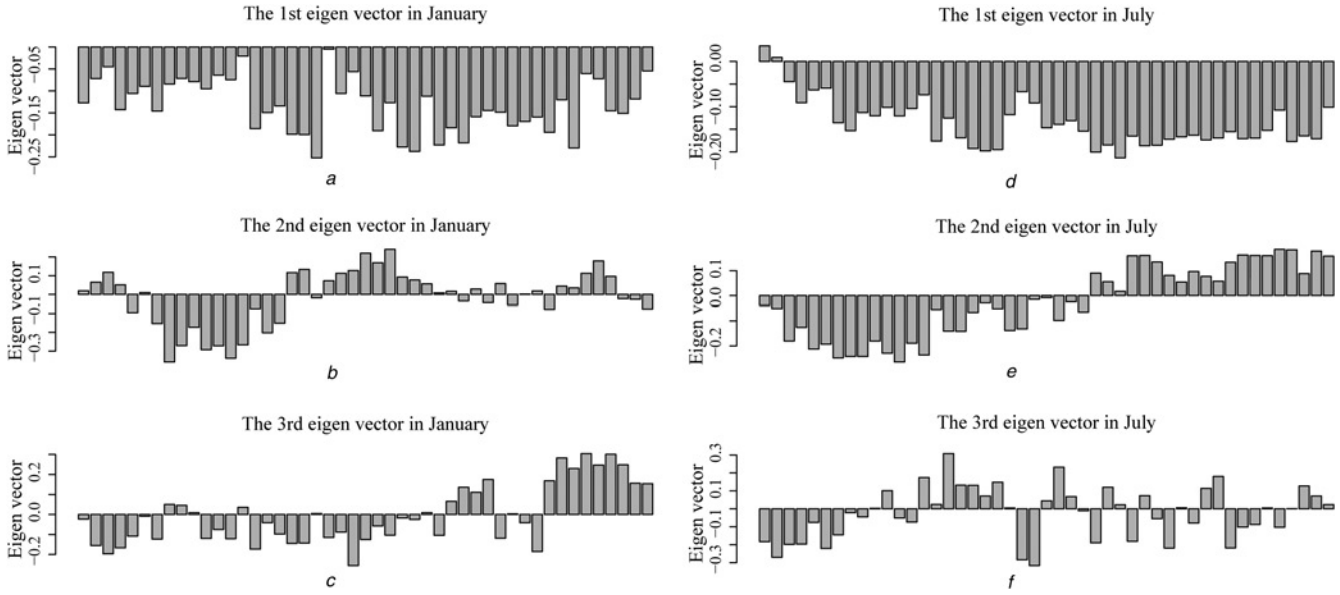

Fig. 7 *First to third eigen-vector for the whole of Japan in January and July*
a First eigen vector in January
b Second eigen vector in January
c Third eigen vector in Januray
d First eigen vector in July
e Second eigen vector in July
f Third eigen vector in July

correlation among the sites. Thus, the so-called smoothing effect is expected to be smaller compared with the ideal case without cross-correlation.

2.4 Estimation of forecast error

We estimated the lower limit of the system-wide forecast error using the cross-correlation coefficients of the output fluctuation described in the previous section. Recently, numerical weather forecasting has gained higher accuracy, because of meteorological information acquired by weather radars and meteorological satellites and the advancement of high-performance computers. Ultimately, as the forecast accuracy becomes higher, the forecast of the PV output time series at each site converges on the moving average trend of the site. Thus, we expect that the short-term fluctuation will be the main component of the forecast error, because short-term fluctuation cannot be forecasted. Therefore we assume here that the lower limit of the forecast error is identical to the short-term fluctuation.

doi: 10.1049/joe.2014.0235

Table 1 Lower limit of the system-wide forecast errors in The Tokyo area

Months	Error w/o cor	Var w/o cor	Error w cor	Var w cor
January	101.26	0.0168	181.36	0.0302
February	107.53	0.0167	206.78	0.0321
March	145.89	0.0208	296.55	0.0423
April	132.21	0.0195	246.95	0.0365
May	136.88	0.0194	262.23	0.0373
June	128.02	0.0226	221.59	0.0391
July	136.47	0.0222	256.34	0.0417
August	128.47	0.0187	236.19	0.0344
September	120.06	0.0202	239.98	0.0405
October	103.20	0.0191	183.09	0.0338
November	108.15	0.0222	236.34	0.0485
December	77.499	0.0144	154.74	0.0288

If the number of forecast sites is small, for example, just one site in each prefecture, the system-wide forecast error involves the cross-correlation between the sites and consequently the system-wide error becomes large. On the other hand, if the number of forecast sites is large, the system-wide forecast error does not involve the cross-correlation between the sites and consequently the system-wide error becomes small. If we consider that in the near future installed PV systems are widely distributed in various places, the actual system-wide forecast error is expected to be between the above two extreme cases.

We estimated the lower limit of the system-wide forecast errors and the coefficients of variation with/without considering the cross-correlations of the PV output fluctuation, using (3) and (4) with the genuine cross-correlation coefficient ρ_{ij} shown in Figs. 4 and 6. The installed capacity of PV systems in 2030 was estimated by dividing the 100 GW capacity in the whole of Japan [1] proportionally to the demand of each prefecture. The estimations of errors and variation coefficients in the Tokyo area and the whole of Japan are shown in Tables 1 and 2, respectively. The second to fifth columns of the tables represent error without correlation, coefficient of variation without correlation, error with correlation and coefficient of variation with correlation, respectively. Both the system-wide forecast errors and the coefficients of variation are increased by considering the cross-correlation of the fluctuation. The lower limit of the coefficients of variation in the Tokyo area is larger than the lower limit of the coefficients in the whole of Japan throughout the year.

3 Cost estimation for PV integration

3.1 Unit commitment model

The purpose of our unit commitment model [5, 6] was to plan the operation schedule of thermal power plants so as to maximise the profit of an electric power utility by taking into account both the forecast of output and its error for renewable energies and the demand response of consumers on the change of electricity prices. The essence of the model is described briefly as follows.

Table 2 Lower limit of the system-wide forecast errors in the whole of Japan

Months	Errors w/o cor	Var w/o cor	Error w cor	Var w cor
January	148.36	0.0093	352.85	0.0223
February	164.13	0.0090	408.97	0.0225
March	218.45	0.0103	655.99	0.0309
April	207.07	0.0094	532.03	0.0242
May	202.87	0.0089	579.35	0.0255
June	196.43	0.0108	449.46	0.0247
July	205.63	0.0104	567.38	0.0288
August	191.47	0.0087	532.09	0.0244
September	184.43	0.0099	557.29	0.0299
October	160.52	0.0092	409.99	0.0235
November	165.50	0.0113	558.46	0.0383
December	124.18	0.0083	313.99	0.0210

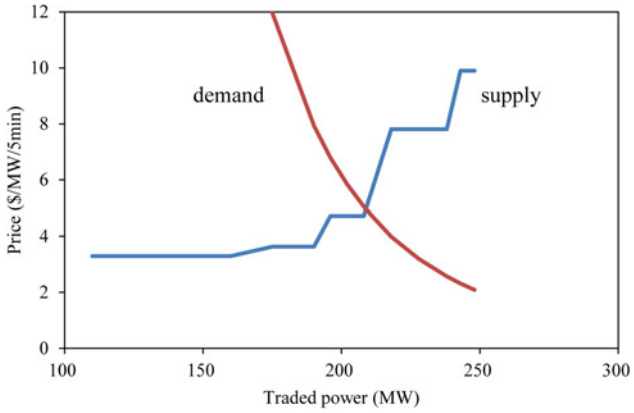

Fig. 8 *Demand and supply*

(1) *Objective function:* The time series of the operational state of thermal power plant $i (i = 1, \ldots, N)$ is obtained by minimising the objective function

$$F(p_t^i,\ u_t^i,\ z_t^i,\ w_t^l) = \sum_{t=1}^{T}\sum_{i=1}^{N}[b_i p_t^i + S_i z_t^i] \tag{23}$$

This objective function represents the operation cost of an electric power utility. Here N, T and L are the number of thermal power plants, time horizon and number of price levels, respectively. Continuous variables p_t^i are the output power variable of thermal power plant i and integer variables u_t^i and z_t^i are the status production variable of thermal power plant i (1 = committed, 0 = decommitted) and start-up variable of thermal power plant i (1 = start up, 0 = others), respectively. Parameters S_i and b_i represent the start-up cost of thermal power plant i and the fuel cost of the thermal power plant i, respectively.

Other parameters related to the demand response $\bar{r}$, r^l and ϵ_{d} are the average electricity price, price level and price elasticity of demand, respectively. If the electricity price r deviates from the average price $\bar{r}$, the demand d is changed from the average demand $\bar{d}$ as follows

$$\frac{d}{\bar{d}} = \left(\frac{r}{\bar{r}}\right)^{\epsilon_{\mathrm{d}}} \tag{24}$$

The dependence of demand d on price r is depicted in Fig. 8.

(2) *Global constraints:* The sum of the demand response variable w_t^l (1 = selected, 0 = not selected) has to satisfy the constraint

$$\sum_{l=1}^{L} w_t^l = 1 \tag{25}$$

to ensure that only a single price level r^l is selected. In addition to this constraint, the average of the selected price r^l has to be equal to the average price $\bar{r}$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{L} w_t^l r^l \le \bar{r} \tag{26}$$

doi: 10.1049/joe.2014.0235

Moreover, the total demand has to be unchanged by the demand response

$$\sum_{t=1}^{T} d_t^{(f)} = \sum_{t=1}^{T} \widetilde{d_t^{(f)}} \tag{27}$$

$$\widetilde{d_t^{(f)}} = d_t^{(f)} \sum_{l=1}^{L} w_t^l \left(\frac{r^l}{\bar{r}}\right)^{\epsilon_d} \tag{28}$$

The forecasted demand and its error are indicated by $d_t^{(f)}$ and σ_d, respectively. Here, (f) stands for forecasting. Importantly, the sum of supply has to be greater than the demand

$$\sum_{i=1}^{N} p_t^i + \mathrm{wd}_t^{(f)} + \mathrm{pv}_t^{(f)} + g_t - h_t \geq \widetilde{d_t^{(f)}} \tag{29}$$

where $\mathrm{wd}_t^{(f)}$, $\mathrm{pv}_t^{(f)}$, g_t and h_t are the forecasted wind power generation, forecasted PV generation, discharged power from pumped hydro power and charging to pumped hydro power, respectively.

If we consider the forecast error of demand σ_d, forecast error of wind power σ_w and forecast error of PV σ_p, the constraint in (29) can be rewritten as

$$\frac{\sum_{i=1}^{N} p_t^i + \mathrm{wd}_t^{(f)} + \mathrm{pv}_t^{(f)} + g_t - h_t - \widetilde{d_t^{(f)}}}{\sqrt{\sigma_d^2 + \sigma_w^2 + \sigma_p^2}} \geq \phi^{-1}(\alpha) \tag{30}$$

where α and $\phi(\cdot)$ are the probability to ensure the supply-demand balance and the cumulative distribution function, respectively. In this paper, we assumed that the system-wide error distribution is the normal distribution.

Pumped hydro power has to satisfy the constraints

$$v_t c^{\min} \leq g_t \leq v_t c^{\max} \tag{31}$$

$$(1 - v_t) c^{\min} \leq h_t \leq (1 - v_t) c^{\max} \tag{32}$$

$$R^{\min} \leq \sum_{s=1}^{t} (h_s \eta - g_s) \Delta_t \leq R^{\max} \tag{33}$$

where v_t, $c^{\min}$, $c^{\max}$, $R^{\min}$, $R^{\max}$, η and Δ_t are the state variable of the pumped hydro power (1 = discharge, 0 = charge), minimum discharge power, maximum discharge power, minimum stored energy, maximum stored energy, efficiency and time step, respectively.

Furthermore, the constraint for reserve capacity is

$$\sqrt{\varsigma_d^2 + \varsigma_w^2 + \varsigma_p^2} \leq \sum_{i=1}^{N} 0.05 u_t^i \bar{p}_{\max}^i + 0.05 c^{\max} \tag{34}$$

where ς_d, ς_w and ς_p are the standard deviations of short-term fluctuation for demand, wind power generation and PV power generation, respectively. We assumed here that 5% of the maximum output $\bar{p}_{\max}^i$ of thermal power plant i and 5% of the maximum discharge power $c^{\max}$ are used as reserve. It is noted that the modelling is crude for the point that thermal power plant i is able to be operated at the maximum output $\bar{p}_{\max}^i$ while 5% of the maximum output $\bar{p}_{\max}^i$ of thermal power plant i is used as reserve.

(3) *Local constraints for thermal power plants:* The following constraints are used for each thermal power plant as typical constraints in a unit commitment model.

- *Generation capacity:* The output power p_t^i has to be between the maximum output power $\bar{p}_{\max}^i$ and the minimum output power $\bar{p}_{\min}^i$ when the operation is in steady state

$$\bar{p}_{\min}^i u_t^i \leq p_t^i \leq \bar{p}_{\max}^i u_t^i \tag{35}$$

- *Minimum up-time constraint:* Thermal power plant i has to be operated longer than the minimum up-time requirement τ_+^i, once the unit is up

$$\begin{aligned} u_t^i &\geq u_s^i - u_{s-1}^i \\ s &\in [t - \tau_+^i,\ t - 1] \end{aligned} \tag{36}$$

- *Minimum down-time constraint:* Thermal power plant i has to be stopped longer than the minimum down-time requirement τ_-^i, once the unit is down

$$\begin{aligned} u_t^i &\leq 1 + u_s^i - u_{s-1}^i \\ s &\in [t - \tau_-^i,\ t - 1] \end{aligned} \tag{37}$$

- *Constraint on the start-up variable:* The start-up variable z_t^i has to satisfy the following constraints by definition

$$\begin{aligned} z_1^i &\geq u_1^i \\ z_t^i &\geq u_t^i - u_{t-1}^i (t > 2) \end{aligned} \tag{38}$$

3.2 System-balancing cost

We estimated the effect of the forecast error and the demand response on the system-balancing cost of PV output power for the Tokyo area in 2030 by using the unit commitment model described in the previous section.

In the unit commitment calculation of the power system in the Tokyo area in 2030, we used the estimated system-wide forecast

Table 3 Parameters used in the unit commitment calculation

Symbols	Parameters
N	91
T	1440 (30 days)
Δ_t	30 min
L	20
$\bar{r}$	16 JPY/kWh
r^1	14 JPY/kWh
r^L	18 JPY/kWh
ϵ_d	0 (without DR), −0.3 (with DR)
α	1.28 (90% CL for the normal distribution)
σ_w	10% of the wind output power
σ_d	5% of the system load
$c^{\min}$	0 MW
$c^{\max}$	11 800 MW
$R^{\min}$	0 MWh
$R^{\max}$	118 000 MWh
η	0.7
σ_p	10% of PV output power
σ_w	10% of wind output power
ς_p	5% of PV output power
ς_w	5% of wind output power
σ_d	5% of system load
ς_d	2.5% of system load
baseload	21 000 MW

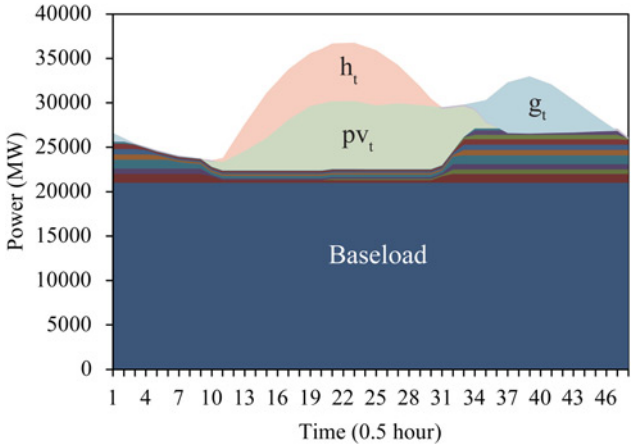

Fig. 9 *Power from various power plants to satisfy the demand in the early May 2030*

error and the model parameters shown in Table 3. The forecast error of PV output power σ_p is almost double the lower limit of the system-wide forecast errors in the Tokyo area shown in Table 1. The installed capacities of PV system and wind system in 2030 in the Tokyo area are assumed to be 26 481 and 1009 MW, respectively. The forecasted PV generation $pv_t^{(f)}$ and forecasted wind power generation $wd_t^{(f)}$ are assumed using the actual output [6]. The forecasted demand $d_t^{(f)}$ in 2030 in the Tokyo area is assumed to be equal to the actual demand in 2010. The baseload power is the sum of nuclear and hydro power other than pumped hydro powers. The pumped hydro powers are aggregated and considered as a single electric storage in (31)–(33).

In Fig. 9, power from various power plants to satisfy the demand in early May 2030 is shown for the Tokyo area. We had the smallest demand and largest output from PV systems in this season. Therefore, the condition for PV integration is the toughest throughout the year. Fig. 9 depicts that many thermal power plants stand by for the PV output fluctuation.

The system-balancing cost ϵ of the PV system per unit output energy was defined by

$$\epsilon = \frac{C(\sigma_p) - C(\sigma_p = 0)}{\sum_{t=1}^{T} pv_t \Delta_t} \tag{39}$$

where $C(\sigma_p)$ and $C(\sigma_p = 0)$ are the operation cost with forecast error and operation cost without forecast error, respectively. PV power pv_t in the denominator of (39) is integrated for each month and is shown in Fig. 10. Decrease in the generated electricity from the

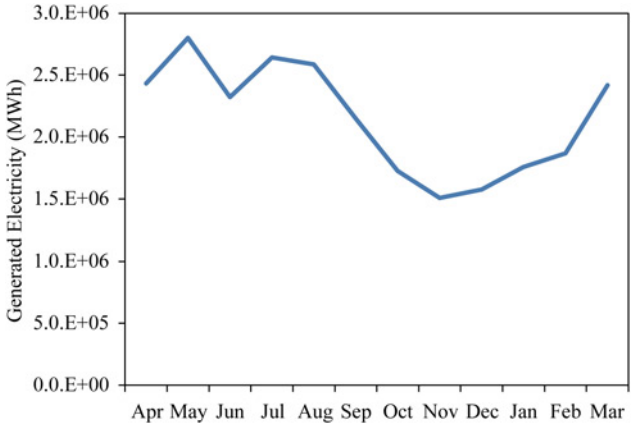

Fig. 10 *Generated electricity in the PV system in 2030*

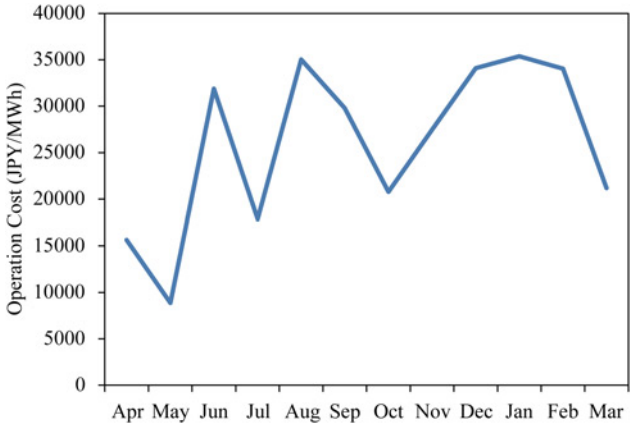

Fig. 11 *Operation cost of thermal power plants in 2030*

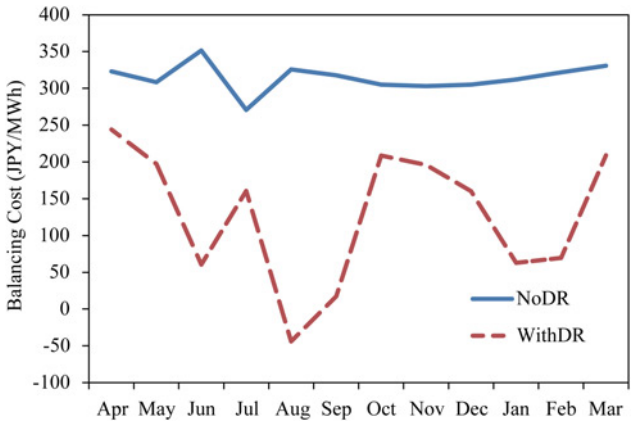

Fig. 12 *System-balancing cost to integrate PV output into the power system in 2030*

PV system is evident. Fig. 11 depicts the operation cost of the thermal power plants in the Tokyo area estimated using the unit commitment model with the forecast error of the PV output power shown in Table 3. The estimated system-balancing cost using (39) is shown in Fig. 12. The solid line and the broken line depict the system-balancing cost without and with the demand response, respectively. The system-balancing costs averaged over the year are 315 and 129 JPY/MWh for the cases without and

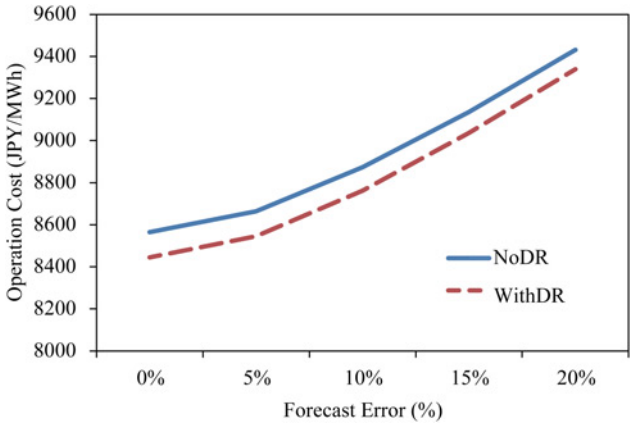

Fig. 13 *Effect of the demand response and the forecast error on the system-balancing cost of PV output power for the Tokyo area in May 2030*

with the demand response, respectively. Fig. 12 depicts that the demand response reduced the system-balancing cost by half throughout the year. The effect of the demand response and the forecast error on the operation cost of PV output power for the Tokyo area in May is shown in Fig. 13. The solid line and the broken line depict the operation cost without and with the demand response, respectively.

It is noted here that the system-balancing cost is a few per cent of the average electricity price $\bar{r}$. The demand response effectively reduced the system-balancing cost. It is naturally expected that the system-balancing cost per unit output energy is the smaller for the whole of Japan because of the smaller coefficient of variation.

4 Discussion

The validity of the concept of 'local production for local consumption of renewable energy' and alternative policy implications are discussed.

When the installed capacity of the PV system is small, the 'local production for local consumption of renewable energy' is economically feasible. The area price in the Tokyo area could be high enough to be close to the feed-in tariff price for PV power and transmission loss is mitigated because of limited transmission inside the Tokyo area. According to the installation plan, the capacity of PV system will become large, parallel with the vitalisation of the power market. In this phase, the lack of balancing capability becomes obvious and the system-balancing cost increases towards the current electricity price. Consequently, the 'local production for local consumption of renewable energy' concept becomes infeasible.

In the near future, we will expand the capacity of the interconnections between the Hokkaido and Tohoku area, the Tohoku and Tokyo area and the 50 Hz/60 Hz boundary in the West of the Tokyo area. A new role is expected for the transmission systems. That is the reduction of the coefficient of variation of the PV output. For instance, the development of the high-voltage direct current line in the whole of Japan [12] will reduce the requirement for the balancing capability. A novel methodology is desired in order to plan an optimal transmission system.

5 Conclusion

We analysed the cross-correlation of PV output fluctuation for the actual PV output time series data in both the Tokyo area and the whole of Japan using the principal component analysis with the random matrix theory. Based on the obtained cross-correlation coefficients, the lower limit of the forecast error for PV output was estimated with/without considering the cross-correlations. Both the system-wide forecast errors and the coefficients of variation were increased by considering the cross-correlation of the fluctuation. The lower limit of the coefficients of variation in the Tokyo area was larger compared with that of the whole of Japan throughout the year. Then, the operation schedules of thermal plants were calculated to integrate PV output using our unit commitment model with the estimated forecast errors. The system-balancing cost of PV system was also estimated with or without demand response. The estimated system-balancing cost is a few per cent of the average electricity price $\bar{r}$ and the demand response effectively reduced the system-balancing cost. Finally, the validity of the concept of 'local production for local consumption of renewable energy' and alternative policy implications were discussed. It is naturally expected that the system-balancing cost per unit output energy is smaller for the whole of Japan because of the small coefficient of variation. This means that the concept of 'local production for local consumption of renewable energy' is not economically feasible when the capacity of PV system becomes large. The development of the transmission lines planned in the near future will reduce the balancing capability required for PV integration.

6 References

[1] Japan Photovoltaic Energy Association: 'JPEA PV Outlook 2030', 2012. (in Japanese)

[2] Hatta T.: 'Denryoku sisutemu kaikaku wo dou susumeruka (How do we step up the reform of the power system?)' (Nikkei Publishing Inc., Tokyo, 2012), ch. 5. (in Japanese)

[3] Niitsuma H., Nakata T.: 'EIMY (energy in my yard)? A concept for practical usage of renewable energy from local sources'. European Geothermal Conf. 2003, Szeged, Hungary, 2003

[4] Oozeki T., Fonseca J., Takashima T., Ogimoto K.: 'Dataset of photovoltaic system output for power system analysis'. IEEJ, FTE11, 2011, pp. 1–5. (in Japanese)

[5] Ikeda Y., Ikegami T., Kataoka K., Ogimoto K.: 'A unit commitment model with demand response for the integration of renewable energies'. IEEE Power & Energy Society General Meeting 2012, 2012, vol. 1, no. 7, pp. 22–26. Available at http://www.arXiv:1112.4909v1 [cs.SY], accessed 21 December 2011

[6] Ikeda Y., Ogimoto K.: 'A unit commitment model with demand response and electric storage device for the integration of variable renewable energies', *IEEJ Trans. Power Energy*, 2013 (in Japanese), **133**, (7), pp. 598–605

[7] Metha M.L.: 'Random matrices' (Academic Press, Boston, 1991)

[8] Laloux L., Cizeau P., Bouchaud J.-P., Potters M.: Noise dressing of financial correlation matrices, *Phys. Rev. Lett.*, 1999, **83**, p. 1467

[9] Pleroux V., Gopikrishnam P., Rosenow B., Amaral L.A.N., Stanley H.E.: Universal and nonuniversal properties of cross correlations in financial time series, *Phys. Rev. Lett.*, 1999, **83**, p. 1471

[10] Pleroux V., Gopikrishnam P., Rosenow B., Amaral L.A.N., Guhr T., Stanley H.E.: Random matrix approach to cross correlations in financial data, *Phys. Rev. E*, 2002, **65**, p. 066126

[11] Iyetomi H., Nakayama Y., Yoshikawa H., *ET AL.*: 'What causes business cycles? Analysis of the Japanese industrial production data', *J. Jpn. Int. Econ.*, 2011, **25**, (3), pp. 246–272

[12] Tanaka N.: 'Energy security & sustainability for Asia in the 21st Century'. The 1st Energy Policy Roundtable 2012, Policy Alternatives Research Institute, University of Tokyo, 2012. Available at http://www.pari.u-tokyo.ac.jp/event/smp120221_tanka.pdf