IEICE Proceeding Series

Simultaneous prediction of temperature and demand for electric power by non-linear models with cross-correlations

Hiromasa Iwayama, Jun-Ichi Inoue

Vol. 1 pp. 320-323
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012
NOLTA2012

Simultaneous prediction of temperature and demand for electric power by non-linear models with cross-correlations

Hiromasa Iwayama† and Jun-Ichi Inoue‡

Complex Systems Engineering, Graduate School of Information Science and Technology
Hokkaido University, N14-W9, Kita-ku, Sapporo 060-0814, Japan
Email: †`iwayama@complex.ist.hokudai.ac.jp`, ‡`j_inoue@complex.ist.hokudai.ac.jp`

Abstract—We investigate the cross-correlation between temperature in Sapporo city and demand for electric power in our Hokkaido University. In order to predict the empirical time series, the so-called *PUCK (Potential of Unbalanced Complex Kinetics) model*, which is based on a potential-driven dynamics of Langevin-type, is modified in terms of information about cross-correlations. The modification enables us to predict the temperature and demand simultaneously. We will examine the statistical performance of the prediction for empirical data sets collected by *Hokkaido University Office for a Sustainable Campus* (`http://www.osc.hokudai.ac.jp/`) for every thirty minutes in August 2011. The resulting Temperature-Demand (TD) diagram is compared with the empirical evidence. Effect of the cross-correlation on the time series prediction and its usefulness are extensively discussed.

1. Introduction

Since the severe earthquake on 11th March 2011, several nuclear power plants located in Fukushima have been suspended. As the results, the demand for electric power became close to the supply level in the middle of last summer. Consequently, to save the electricity to compensate for the nuclear energy is an urgent question for us to overcome. Needless to say, the largest customers of the electric power are industries of various commodities, however, even in academia such as university or college, economizing in power is by all means expected. Actually, in order to draw our attention to the power saving, the empirical data for the total amount of power used in Hokkaido University has been visualized and opened to the public through the web site [1]. The data is recorded for every thirty minutes and the information about the percentage of power use against the upper bound (limit) is available as a 'time series' (see Figure 1 as an example). In university, electric power demands consist of air conditioning and lightning for each room, experimental devices, and maintenance of various computers, servers and so on so forth. Therefore, from the view point of practical use, it is important for us to grasp the balance of the supply and demand, and if possible, we should make a prediction for the demand precisely.

It is also well-known that temperature in Sapporo is relatively lower than the other large cities in the main island of

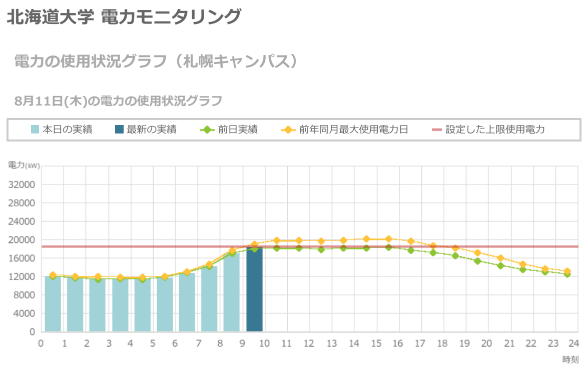

Figure 1: A typical time series of the percentage of power use against the upper bound (limit) in Hokkaido University (partially written in Japanese). The horizontal axis is time (the minimum unit is thirty minutes), whereas the vertical axis is the percentage of power use.

Japan. However, it might be a universal fact that we unconsciously turn on the air-conditioner when temperature in our room increases. As the result, the demand for electric power is expected to increase. Hence, we might naturally assume that there exists a finite cross-correlation between temperature of Sapporo city and the demand for electric power in Hokkaido University. This means that it is quite reasonable for us to use the information about the correlation to predict the temperature and demand simultaneously.

With these facts in mind, here we empirically analyze the cross-correlation between temperature and the demand for electric power at Hokkaido University in Sapporo city. We also attempt to construct the procedure to predict these two kinds of time series simultaneously. The empirical data for the benchmark test has been provided by Hokkaido University Office for a Sustainable Campus [2] for every thirty minutes in August 2011.

2. Empirical evidence

To construct the prediction model, here we will show several empirical evidences concerning temperature and demand for power at Hokkaido University in Sapporo city. We utilize the Pearson coefficient to quantify the correlation and visualize the distribution of the coefficient in each time zone in a day. The empirical data for the temperature in Sapporo city is provided from the web site of Japanese meteorological agency [3]. The temperature is available for every thirty minutes in the data set.

2.1. The pattern of time series

We first show the time-dependence of the demand for electric power in Hokkaido University D_t and temperature in Sapporo city T_t during 24 hours on a typical summer day (1st August 2011) in Figure 2. From these panels, we easily

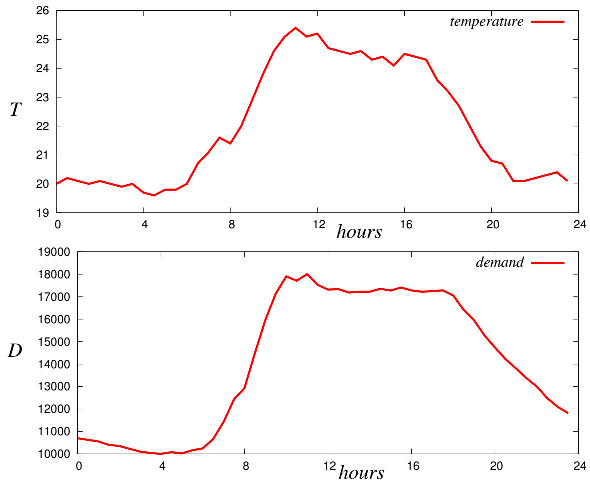

Figure 2: The change of the demand for electric power in Hokkaido University (the lower panel) and temperature in Sapporo city (the upper panel) in 24 hours on 1st August 2011. The units of vertical axis are 'Kilowatt' (the lower panel) and 'Celsius' (the upper panel), respectively.

find that the patterns of change for temperature and demand are quite similar. To grasp the correlation between these two time series more explicitly, we make two-dimensional scatter plots (T_t, D_t), $t = 1, \cdots, 48$ (what we call *TD diagram* [4]) in Figure 3. In this figure, the shape of *deformed*

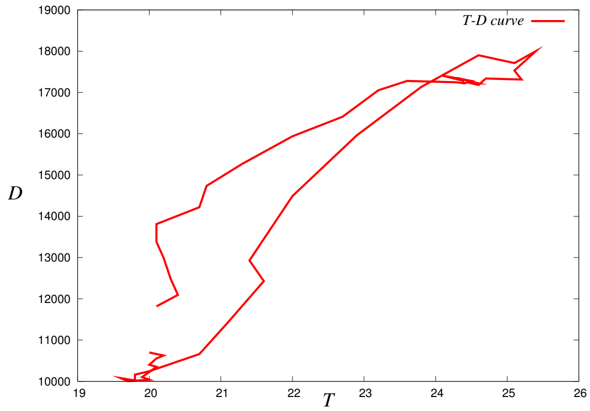

Figure 3: TD diagram for temperature in Sapporo and demand for power in Hokkaido University. The shape of *deformed 'eight (8)'* is apparently observed.

'eight (8)' is clearly observed in the TD diagram. This result implies that the correlation between temperature and demand for power at Hokkaido University in Sapporo city is not the special case but somehow 'universal'. The origin of the shape is intuitively understood as follows. Before noon, both demand and temperature increases, whereas afternoon, temperature decreases but the slope of decreases of demand is much looser than the slope of increase in the morning. This is because activity of office workers does not decreases even if temperature decreases. This fact makes the TD diagram a *deformed 'eight (8)'* shape.

2.2. Correlation between temperature and demand

To evaluate the cross-correlation between T_t and D_t, we shall use the *Pearson coefficient*. Let us first define the difference in the logarithmic scale of T_t, D_t between successive two time steps as $\Delta T_t = \log T_t - \log T_{t-1}, \Delta D_t = \log D_t - \log D_{t-1}$. Then, the moving average and the two-body correlation over the past M steps are given by

$$\overline{\Delta T_t} \equiv \frac{1}{M}\sum_{l=t-M+1}^{t} \Delta T_l, \quad \overline{\Delta T_t \Delta D_t} \equiv \frac{1}{M}\sum_{l=t-M+1}^{t} \Delta T_l \Delta D_l, \quad (1)$$

respectively. In terms of the above two quantities, the Pearson coefficient is calculated as

$$\rho_{TD}^{(t)} = \frac{\overline{\Delta T_t \Delta D_t} - (\overline{\Delta T_t})(\overline{\Delta D_t})}{\sqrt{[\overline{(\Delta T_t)^2} - (\overline{\Delta T_t})^2][\overline{(\Delta D_t)^2} - (\overline{\Delta D_t})^2]}}. \quad (2)$$

In Figure 4, we plot the relationship between the Pearson coefficient ρ_{TD} and the difference of demand ΔD. From these panels, we definitely find that the Pearson coefficient takes non-zero values and it is apparently 'biased'. The degree of the bias is dependent on the time zone in a day. For instance, the temperature and demand in 4:00-7:30 are strongly correlated, and from the fact $\Delta D > 0$, both temperature and demand increase in this time zone. Hence, we might use the information about these kinds of correlation to make an effective prediction model.

3. The PUCK model

To make an effective prediction model, we pick up the so-called PUCK model [5]. The PUCK model itself has been used in the prediction for tick-by-tick financial data. Nevertheless, here we apply it to the temperature-demand prediction as a basic model. In the PUCK model, the time series of temperature is predicted recursively as

$$T_{t+1} = T_t - \left.\frac{dU(T,t)}{dT}\right|_{T=T_t-T_t^{(M)}} + f_t \quad (3)$$

where $T_t^{(M)}$ stands for the moving average over the past M steps $T_t^{(M)} = \frac{1}{M}\sum_{l=0}^{M-1} T_{t-l}$, and the third term appearing in the left hand side of equation (3), that is f_t, denotes a white Gaussian noise satisfying $\langle f_t f_{t'} \rangle = \delta_{tt'}$. Namely, the temperature T_t is pushed toward the direction of gradient decent in the potential surface U on the analogy of classical mechanics. In other words, $-dU/dT$ denotes a sort of 'mechanical force' to push the temperature to the direction.

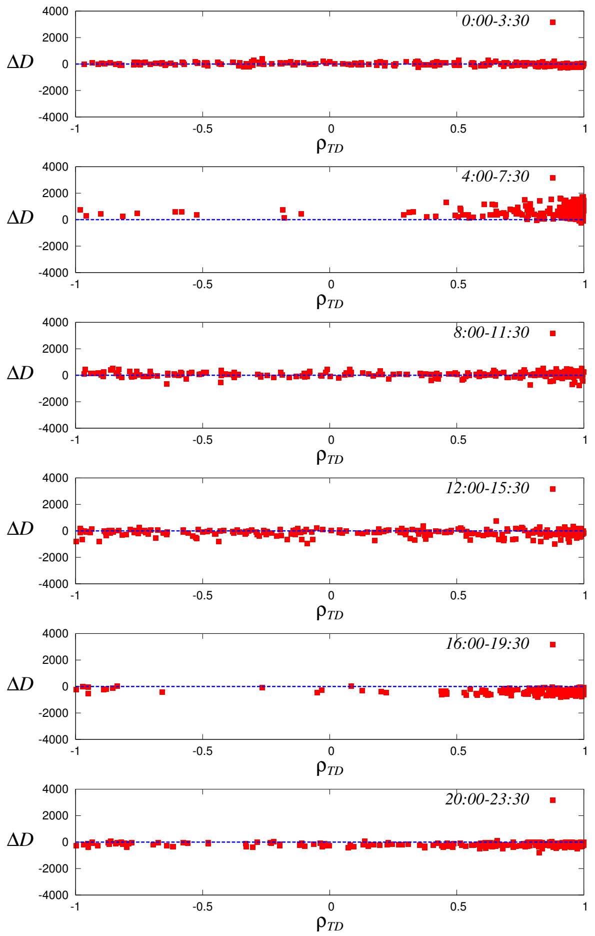

Figure 4: The relationship between the Pearson coefficient ρ_{TD} and the logarithmic difference of demand ΔD for successive two time steps. From the top panel to the bottom, the relationships in non-overlapping six time zones: 0:00-3:30, 4:00-7:30, 8:00-11:30, 12:00-15:30, 16:00-19:30 and 20:00-23:30 are shown.

For instance, the temperature is updated so as to move to the $T_t^{(M)}$ under the noise when the potential U is described as a parabola having the minimum at $T_t^{(M)}$. On the other hand, for the potential U of a cubic curve, the point $T_t = T_t^{(M)}$ is no longer stable and the instability makes the temperature to decrease (or increase) suddenly. Taking into account the above intuitive explanation, here we choose the parametric potential function as $U(T,t) = \sum_{k=1}^{K} \frac{b_k(t)}{k+1} T^{k+1}$. From the above definition, the potential function becomes a parabola for $b_1 \neq 0, b_2 = \cdots = b_K = 0$, whereas it becomes a cubic for $b_1, b_2 \neq 0, b_3 = \cdots = b_K = 0$.

3.1. The modification of the PUCK model

Here we will modify the PUCK model (3) by means of the Pearson coefficient. We rewrite equations for T_t and D_t by using $\rho_{TD}^{(t)}$ over the past Q steps as

$$T_{t+1} = T_t - \left.\frac{dU(T,t)}{dT}\right|_{T=T_t - T_t^{(M)}} + \sum_{q=0}^{Q} \lambda_q \rho_{TD}^{(t-q)} + f_t \qquad (4)$$

$$D_{t+1} = D_t - \left.\frac{dU(D,t)}{dD}\right|_{D=D_t - D_t^{(M)}} + \sum_{q=0}^{Q} \lambda_q \rho_{TD}^{(t-q)} + g_t \qquad (5)$$

where g_t and f_t denote uncorrelated white Gaussian noises satisfying $\langle g_t g_{t'} \rangle = \delta_{tt'}$ and $\langle f_t g_t \rangle = 0$. The potential function $U(D,t)$ is defined by the same way as $U(T,t)$, namely, $U(D,t) = \sum_{k=1}^{K} \frac{c_k(t)}{k+1} D^{k+1}$.

3.2. Model selection by means of the AIC

The parameters and the number K of them appearing in the potentials $U(T,t)$ and $U(D,t)$ are determined by minimization of the *Akaike Information Criterion (AIC)*: $AIC_T = -2\ln(l(\{b_k\},\{\lambda_q\},K,M)) + 2K$ and $AIC_D = -2\ln(l(\{c_k\},\{\lambda_q\},K,M)) + 2K$ with respect to the parameters and K, where we defined $\{b_k\} \equiv (b_1,\cdots,b_K)$, $\{c_k\} \equiv (c_1,\cdots,c_K)$ and $\{\lambda_q\} \equiv (\lambda_1,\cdots,\lambda_Q)$. In this paper, we shall fix the parameter Q as $Q = 1$ for simplicity. Accordingly, K stands for the number of parameters to be estimated and l means the log-likelihood for given two Gaussians $w[f_t], w[g_t]$. By taking into account the property of white Gaussian noise, we immediately have $l(\{b_k\},K,M) = \prod_{t=n+N-1}^{n} w[f_t]$ and $l(\{c_k\},K,M) = \prod_{t=n+N-1}^{n} w[g_t]$.

4. Numerical experiments

Let us examine our prediction model for empirical data of temperature and demand for electric power at Hokkaido University in Sapporo. To deal with the data in our model, we shall normalize the data to keep them in the range [0, 1] by dividing each value by the maximum.

4.1. The TD diagram

The resulting TD diagram is shown in Figure 5. From this figure, we clearly find that our prediction model reconstructs the similar TD diagram having the *deformed 'eight (8)'* shape to the empirical evidence shown in Figure 3.

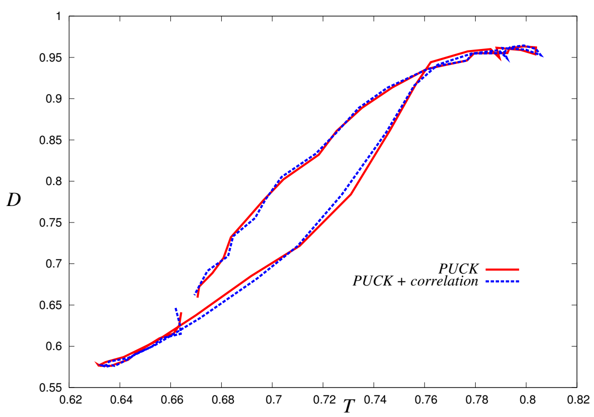

Figure 5: The resulting TD diagram. The line caption 'PUCK' denotes the result of the original PUCK, whereas 'PUCK + correlation' means the result by our proposed model. See Figure 3 for the counter evidence.

4.2. Average-case performance

We next utilize the mean-square error between the prediction T_t, D_t and the corresponding true values (observables) $\hat{T}_t, \hat{D}_t$ as a performance measurement

$$\varepsilon_1^{(T)}(t) = \mathbb{E}[(T_t - \hat{T}_t)^2], \quad \varepsilon_1^{(D)}(t) = \mathbb{E}[(D_t - \hat{D}_t)^2] \qquad (6)$$

where $\mathbb{E}[\cdots]$ stands for the average over the data sets for 23 days. We also introduce the mismatch measurement for the sign of 'up-down movement' in temperature and demand, that is to say,

$$\varepsilon_2^{(T)}(t) = \frac{1}{2}(1 - \mathbb{E}[\mathcal{T}_t \hat{\mathcal{T}}_t]), \; \varepsilon_2^{(D)}(t) = \frac{1}{2}(1 - \mathbb{E}[\mathcal{D}_t \hat{\mathcal{D}}_t]) \qquad (7)$$

for $\mathcal{T}_t \equiv \mathrm{sgn}(T_{t+1} - T_t)$, $\hat{\mathcal{T}}_t \equiv \mathrm{sgn}(\hat{T}_{t+1} - \hat{T}_t)$, $\mathcal{D}_t \equiv \mathrm{sgn}(D_{t+1} - D_t)$ and $\hat{\mathcal{D}}_t \equiv \mathrm{sgn}(\hat{D}_{t+1} - \hat{D}_t)$. We show the results in Figures 6 and 7. These numerical results obtained from limited

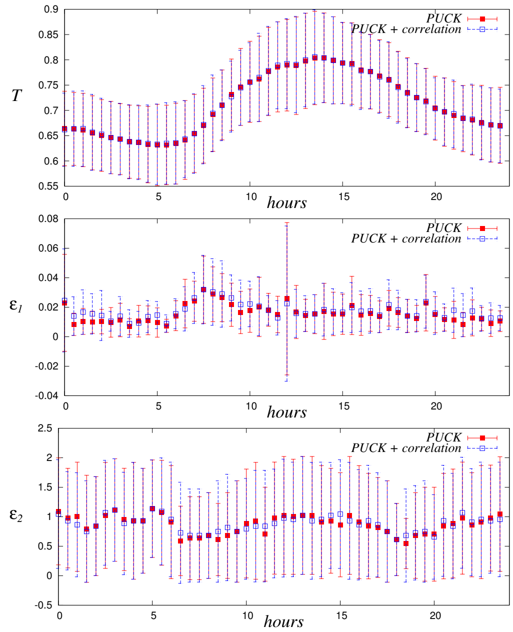

Figure 6: From the top panel to the bottom, the prediction curve for temperature, $\varepsilon_1^{(T)}, \varepsilon_2^{(T)}$ as a function of time are shown. The errorbars are calculated as standard deviations from the data sets for 23 days. The line caption 'PUCK' denotes the result of the original PUCK, whereas 'PUCK + correlation' means the result by our proposed model.

data sets are still at preliminary level, however, we find that the errors $\varepsilon_1, \varepsilon_2$ of our proposed prediction model becomes smaller than the conventional PUCK model without correlations on average. Especially, it works well in the time-zone from morning to noon and from evening to midnight at which both temperature and demand drastically change.

5. Concluding remark

In this paper, we investigated time-dependence of cross-correlations between temperature in Sapporo city and demand for electric power in Hokkaido University. In order

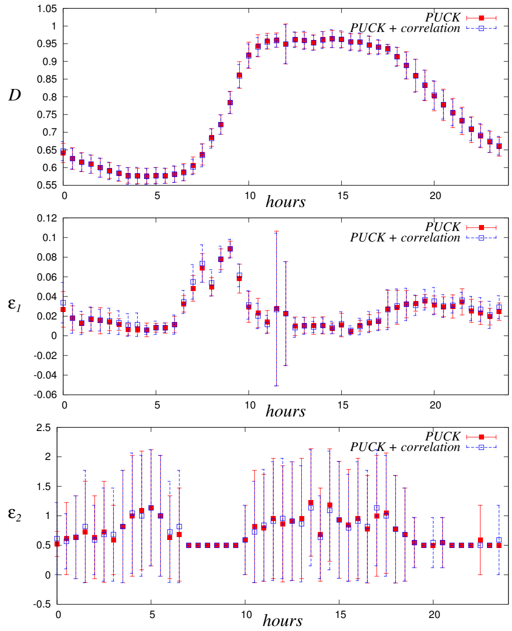

Figure 7: From the top panel to the bottom, the prediction curve for demand, $\varepsilon_1^{(D)}, \varepsilon_2^{(D)}$ as a function of time are shown. The errorbars are calculated as standard deviations from the data sets for 23 days. The line caption 'PUCK' denotes the result of the original PUCK, whereas 'PUCK + correlation' means the result by our proposed model.

to predict the time series of temperature and demand simultaneously, the PUCK model was modified in terms of information about cross-correlations. We numerically examined the statistical performance of the prediction for empirical data sets. We found that the resulting TD diagram is quite similar to the empirical evidence.

The details including much more extensive data analysis will be presented at the conference.

Acknowledgments

We would like to thank Tomohiro Morimoto and Takashi Yokoyama for providing us empirical data [2] for the demand for electric power in Hokkaido University. One of the authors (JI) was financially supported by Grant-in-Aid for Scientific Research (C) of Japan Society for the Promotion of Science, No. 22500195.

References

[1] http://www.hokudai.ac.jp/

[2] http://www.osc.hokudai.ac.jp/

[3] http://www.data.jma.go.jp/obd/stats/etrn/index.php

[4] Y. Nameki, *Proc. of the Operations Research Society Japan* **1-A-7**, (in Japanese) (1995).

[5] K. Watanabe, H. Takayasu and M. Takayasu, *Physical Review E* **80**, 056110 (2009).